Exploring Metabolomic Flux and Achieving Prediction Capability in Cocoa Bean Fermentation using Model Systems

by

Warren Andrew John

a Thesis submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy in Biochemistry

Approved Dissertation Committee

Prof. Dr. Matthias S. Ullrich
Jacobs University Bremen

Prof. Dr. Nikolai Kuhnert
Jacobs University Bremen

Dr. Roy D'Souza
Jacobs University Bremen

Dr. -Ing. Maximilian Greiner
Barry Callebaut, Wieze, Belgium

Date of Defense: 29 April 2019

Department of Life Sciences and Chemistry

Statutory Declaration

Family Name, Given/First Name	John, Warren
Matriculation Number	20329558
What kind of thesis are you submitting: Bachelor-, Master-, or PhD-Thesis	PhD Thesis

English: Declaration of Authorship

I hereby declare that the thesis submitted was created and written solely by myself without any external support. Any sources, direct or indirect, are marked as such. I am aware of the fact that the contents of the thesis in digital form may be revised with regard to usage of unauthorized aid as well as whether the whole or parts of it may be identified as plagiarism. I do agree my work to be entered into a database for it to be compared with existing sources, where it will remain in order to enable further comparisons with future theses. This does not grant any rights of reproduction and usage, however.
The Thesis has been written independently and has not been submitted at any other university for the conferral of a PhD degree; neither has the thesis been previously published in full.

German: Erklärung der Autorenschaft (Urheberschaft)

Ich erkläre hiermit, dass die vorliegende Arbeit ohne fremde Hilfe ausschließlich von mir erstellt und geschrieben worden ist. Jedwede verwendeten Quellen, direkter oder indirekter Art, sind als solche kenntlich gemacht worden. Mir ist die Tatsache bewusst, dass der Inhalt der Thesis in digitaler Form geprüft werden kann im Hinblick darauf, ob es sich ganz oder in Teilen um ein Plagiat handelt. Ich bin damit einverstanden, dass meine Arbeit in einer Datenbank eingegeben werden kann, um mit bereits bestehenden Quellen verglichen zu werden und dort auch verbleibt, um mit zukünftigen Arbeiten verglichen werden zu können. Dies berechtigt jedoch nicht zur Verwendung oder Vervielfältigung.
Diese Arbeit wurde in der vorliegenden Form weder einer anderen Prüfungsbehörde vorgelegt noch wurde das Gesamtdokument bisher veröffentlicht.

.....
Date, Signature

Table of Contents

Acknowledgments

The thesis would not have been possible without the inspiring guidance and mentorship of my academic supervisor, Prof. Dr. Matthias Ullrich. His constant support has not just kept me striving for excellence these past 5 years, but also encouraged me to think creatively and given me a passion for the world of molecular biology. I thank him for the opportunities of supervising Bachelor and Master students, and especially getting involved in the microbiology lab course, and the summer camp which were both great teaching experiences. His approachability, patience and openness have made him the ideal supervisor and I can't think of anyone better to have worked under!

The experiments not only needed lab expertise and equipment but also the coordinated and laborious procurement of cocoa bean material which was only possible through the constant support and funding by Barry Callebaut, for whom I am very grateful. Special thanks goes to Marcello Corno, Audrey Bergounhou, Kouame Jean Koffi and Gino Vrancken for their guidance and support throughout the project. I also would like to acknowledge Jacobs University Bremen for providing me with the platform to carry out my studies.

I would like to thank the members of my thesis committee, Prof. Dr. Nikolai Kuhnert, Dr. Roy D'Souza and Dr. Maximilian Greiner for their presence on the committee, evaluation of my thesis and critical feedback. The members of the COMETA project have also been exceptionally resourceful in providing feedback and helpful suggestions in moving my research forward. Special thanks goes to Dr. Roy D'Souza who has been instrumental in the progression of my research by keeping me thinking critically and guiding me along with the way, especially through the statistical and analytical work he has done.

Work in the lab would not have been as fun without the company of my "lab-mates" in the AG Ullrich Workgroup: Maike, Nina, Candice, Veronika, Neha, Sujit, Song, Ben, Inam and Khaled. Thank you for all the lunch time conversations, "Torten Tuesdays", birthday celebrations, dinner outings, *etc*. You made coming to work every week such an enjoyable experience! Special thanks goes out to Maike and Nina who tirelessly helped me with the overwhelming amount of processing work that needed to be done!

There were plenty of side projects also being run by the "army" of students working under me. I am very grateful to Dilantha, Cise, Merilin, Isha, Laura and Daniel whose perseverance, despite incredibly frustrating moments, has contributed to the wealth of knowledge that has made this thesis possible.

Unlike most people, my usual day consisted of both the morning shift AND the evening shift. It has been a wonderful experience to be a part of the best college at Jacobs, and to have worked with such an amazing team of people. A hearty thank you goes out to Laura, Shola, Florence, Martin, Adilah, Patrice, Nickie, Farhat, Qasim, the kids and the members of the KCO team for making my life at home in Krupp College such an exciting one! There has never been a dull weekend.

I am so blessed to have gotten to know the Böttchers: Nina, Jörg, Jule and Michel, who have adopted me and looked after me as if I were one of their own. Thank you for teaching me what it means to be a true Bremer (although I have yet to go to a WB game) and for all the dedicated support you have given me in teaching me to drive, taking me home after surgery, helping me move (twice), looking after the musketeers, Christmas presents, cookouts, trips, *etc*. I

am eternally grateful and can assuredly say I have a family in Germany because of you.

I can't even begin to acknowledge my brothers and sisters in Christ, on campus in the Christian Fellowship, and off campus at the IBC Bremen. It has been an incredible blessing to have people to pray with, and encourage me even in the hardest times of life. Thank you for the wonderful fellowship during morning prayers, Bible study, retreats and outings, worship nights, *etc*. You have spurred on my growth in the Lord and have proven time and again that selfless, sacrificial and unconditional love is indeed a reality through the cross!

I am undoubtedly grateful to my parents who have sacrificed much for me to come to Germany and pursue my studies at Jacobs University. They have been pivotal in making me the man I am today, raising me up in the knowledge of God and seeking God's guidance and direction in my life constantly and unwaveringly. Thank you for believing and trusting in me and in God's faithfulness throughout and being such a source of spiritual strength and encouragement! A big thank you also goes out to Shane and the rest of my family in Sri Lanka and Switzerland for their support.

The only reason I stand here today is through the sovereign grace of God, who has been my constant anchor throughout these nine years here at Jacobs, and given me incredible victory in the experiments that I have undertaken and the challenges I have faced. "Seek first the Kingdom of God and His righteousness, and all these things will be given to you," says Matthew 6:33. He has proven time and again, that as I set my gaze on Him, He is always with me, and His plans for me are far greater than I could have ever imagined.

Summary

Cocoa bean fermentation is an essential process in the chain of chocolate manufacture. As opposed to other methods of fermentation, it is a more empirical process that is typically carried out on the plantations on which the beans are harvested. The process, however, is crucial to the formation of fundamental flavour and aroma precursors that would contribute to the pleasant flavour and aroma characteristic of chocolate.

A standard fermentation encompasses the growth of successions of three groups of microorganisms on the bean that commandeer the microbiome in such a way that they produce microbial metabolites which diffuse into the bean. Consecutive growth of yeasts, lactic acid bacteria and acetic acid bacteria result in their respective biosynthesis and secretions of ethanol, lactic acid and acetic acid, and their metabolism leads to dramatic increases in temperature within the bean pile. A combination of all these events affect the bean structurally and biochemically and results in a plethora of reactions within. These reactions, primarily driven by changing pH and temperature, include proteolysis, polyphenol degradation, changes in lipid composition, breakdown of sugars, *etc*. Taken together, these events lead to the alleviation of bitterness and astringency through protein and polyphenol breakdown, and the generation of subunits of reducing sugars, amino acids and oligopeptides increases the tendency for pleasant aroma and flavour notes.

Fermentation is a process that is carried out spontaneously. This means that farmers generally rely on a perfect alignment of the factors responsible, which would result in the perfect fermentation. This would involve the right weather conditions, the right microbial input, the correct microbial dynamics, good bean quality, and no spoilage, among other conditions. Suffice to say, this does not happen all the time. Despite centuries of practice, cocoa bean fermentation still remains an uncontrolled process and is given over to the elements of nature. This means that there is tremendous variety in the quality of beans resulting from trials of fermentation, from beans that have high tendency to produce exceptional-quality chocolate to trials of fermentation that have gone severely awry.

Consequently, the chocolate industry, especially Western conglomerates that buy fermented beans from farmers, is left with a huge dilemma. The lack of control over the process means the rest of the chocolate production process needs constant optimization depending on the quality of the fermented beans entering the production line, and consequences can be as adverse as discarding whole batches of beans shipped thousands of miles, simply due to the fact that they have not been accurately identified as underfermented. Furthermore, there is a lack of markers

that are able to distinguish between good and bad trials of fermentation.

The experiments described in this thesis were done in order to more accurately study the reactions that underlie the process of fermentation, both within and surrounding the bean in order to find biochemical markers defining good-quality fermentations. This was done using model fermentation systems through chemical-driven and microbial-driven systems, to assess the contributions of each factor on the fermentation. Cocoa bean fermentation, assessed through bean pH, microbial dynamics and the secretions of microbial metabolites, was successfully reproduced in lab-scale quantities using starter cultures, temperature regimes and submerged incubations.

The main findings of this research were that changes in bean pH led to dramatic changes in terms of protein and polyphenol content and that an optimum pH needed to be reached in order to allow correct proteolysis and flavanol degradation for the formation of cocoa aroma and flavour. Therefore strict control over the microbial populations growing on the surface of the beans would help in standardizing the outcome of the fermentation and ensure better yields of well-fermented beans. Acid influx into the bean also exhibited a preservation effect upon flavanols, which could be beneficial in preserving the health properties of cocoa. A comprehensive understanding of the factors contributing to fermentation through design of experiments also enabled the prediction of outcomes of fermentation trials.

Abbreviations

2D-PAGE	2-Dimensional Polyacrylamide Gel Electrophoresis
AAB	Acetic Acid Bacteria
CFU	Colony Forming Unit
cpSNP	Chloroplastal Single Nucleotide Polymorphism
ESI	Electrospray Ionization
FRAP	Ferric Reducing Antioxidant Power
LAB	Lactic Acid Bacteria
MALDI	Matrix-Assisted Laser Desorption and Ionization
MS	Mass Spectrometry
PCA	Principle Component Analysis
pI	Isoelectric point
POP	Glyceride of Palmitic acid-Oleic acid-Palmitic acid
POS	Glyceride of Palmitic acid-Oleic acid-Stearic acid
qToF	Quad Time of Flight
SOS	Glyceride of Stearic acid-Oleic acid-Stearic acid
SSR	Single Strand Repeats
TAG	Triacyl Glyceride
TcAP1	*Theobroma cacao* Aspartic Proteinase 1
TcAP1	*Theobroma cacao* Aspartic Proteinase 2
TMP	Tetramethylpyrazine
ToF	Time of Flight
TPTZ	2,4,6-tris(2-pyridyl)-s-triazine
TrMP	Trimethylpyrazine
UHPLC	Ultra-High-Performance Liquid Chromatography
UHR	Ultra High Resolution

1. Introduction

1.1. History and Origins of *Theobroma cacao*

Cocoa has been a commodity that has been farmed and cultivated for millennia. Its origins can be archeologically traced back to the Western Amazonian basin, and it was considered to be sacred by civilizations such as the Mayans and Aztecs (Dreiss & Greenhill, 2008). It was already known back in the time to have healing properties and the so-called cocoa drink "cacaoatl" was considered to be divine in nature, only being attributed to the highest forms of human lifestyle such as marriage and divine worship (Crown & Hurst, 2009; Vail, 2009).

Cocoa was, and still remains, a tree that is solely cultivated in tropical regions and therefore contributed vastly to the socio-economy of the regions of central and south America in the middle ages. In its early days, it was primarily used in drinks. Later on, fermented and dried beans would be sold as currency, eventually leading to the quip "money grows on trees" (Clarence-Smith, 2000). Between the 16th and 17th centuries, Southern Mexico was renowned for its trade of cocoa beans. The mass cultivation and export of beans led to cocoa being a major commodity that was heavily shipped around the world in the post-Columbian era. Cocoa was introduced in the Caribbean islands, where Jamaica experienced a "cocoa boom" in the 1670s. Trinidad, Martinique and Haiti were among the nations in the Caribbean where cocoa grew popular (Wood & Lass, 2001). Colonialism took the crop to the far reaches of the world. Cocoa was introduced in the Asian regions of the Philippines, Sulawesi, Jawa, South India and Sri Lanka, and then in the countries of West Africa (Bartley, 2005; Wood & Lass, 2001). As the cultivation of cocoa expanded, the techniques for farming and post-harvest processes also diversified.

While the Mesoamerican drink known as "cacaoatl" or "chocoatl" was being loathed for its horrid, bitter taste, Europeans started making chocolate which was much more pleasant. Chocolate had radical claims of miraculous healing properties and inflaming passions and romance. The first records of chocolate were already in 1652 (Wadsworth, 1652) and chocolate was gradually making its way into the wealthy homes of London. Over the years, the cultivation of cocoa and its subsequent processing have been refined to the palatable texture and taste of the chocolate we have today. The recipes of dried and fermented cocoa cotyledon, sugar and excess added cocoa butter is a beloved product that is enjoyed around the world by both young and old

(Wood & Lass, 2001).

1.2. Varieties of Cocoa

The expansion of cocoa cultivation over the last 500 years has resulted in an extensive array of cocoa varieties. The majority of researchers distinguish between three morpho-geographical groups: Criollo, Forastero and Trinitario (Laurent, Risterucci, & Lanaud, 1994). There is still no concrete genetic classification method for the types of cocoa. There have been various attempts to genetically group the varieties of cocoa (D'Souza et al., 2017; Kumari et al., 2018; Laurent et al., 1994; Motamayor et al., 2008), however none becoming established classification methods.

1.2.1. Traditional cultivars

Criollo is translated as "native" and typically refers to the trees that are native to South America where cocoa was first grown. Pre-Columbian civilizations in Mesoamerica cultivated Criollo trees. This variety is particularly difficult to cultivate because of its susceptibility to diseases and pests and its lower yields. Despite this, it is the most premium cocoa in the market and is used to produce fine-flavour cocoa (Laurent et al., 1994). The seeds are white in colour and the resulting cocoa has prominent nutty and caramel notes with low bitterness (Ascrizzi, Flamini, Tessieri, & Pistelli, 2017). Criollo production only accounts for 3-5% of the world's entire cocoa (Afoakwa, Paterson, Fowler, & Ryan, 2008; Ascrizzi et al., 2017; Castro-Alayo, Idrogo-Vásquez, Siche, & Cardenas-Toro, 2019). Nacional, which is a Criollo variety native to Ecuador, has been classed as a traditional cultivar due to its reduced genetic diversity (Lerceteau, Robert, Pétiard, & Crouzillat, 1997).

1.2.2. Hybridized cultivars

Forastero is translated as "foreign"; in essence, a variety that is not indigenous to the region. It is believed that this name stemmed from the name given to crops that were a hybrid of the Calabacillo variety, grown in Trinidad, and Criollo types. (Bartley, 2005). Later on, the term Forastero was applied to any of the types that were not Criollo. Thus, this particular variety is extremely diverse in terms of its genetic variants. Today the Forastero variety is mainly comprised of the Amelonado variety in West Africa, which is described as being melon shaped

and much smoother compared to other varieties. Forastero trees are known for their strength, resistance to disease and their high rates of production. This makes them much easier to cultivate (Laurent et al., 1994). However the resulting cocoa is not as fine-flavoured as the Criollo variety. Beans of the Forastero variety also need to be fermented in order to develop cocoa aroma and flavour whereas this is not a necessity for Criollo cocoa (Wood & Lass, 2001).

The need for a variety of cocoa that would result in fine-flavour cocoa as well as contribute to higher yields of cocoa resulted in the Trinitario variety which is a hybrid of the Criollo and Forastero types. There are multiples reports as to the origins of this variety. Some hypothesize that it was the result of cross-breeding of the Criollo and Amelonado from the Guyanas or the Lower Amazon regions, while others propose a mechanism of natural hybridization as a result of the major cocoa crop failure of the Trinidadian Criollo plantations (Motamayor, Risterucci, Heath, & Lanaud, 2003). Recent research through SSR and cpSNP analysis has revealed the genesis of the Trinitario variety to be complex and a result of multiple events (Zhang & Motilal, 2016). As a result of the quality and robustness of this variety, it grew in popularity among cultivations, even out-competing Forastero varieties in the 19th century. It eventually became classified as a fine-flavour variety. Today, however, the quality of Forastero cocoa is as good as that of the Trinitario hybrids.

1.2.3. Biochemical variation in cocoa

There are characteristic differences between the varieties of cocoa, notably between Criollo and non-Criollo types. The purple colour of non-Criollo beans is a result of an excess amount of 3-β-galactosyl- and 3-α-L-arabinosyl-cyanidins (Tran et al., 2015). Criollo beans are also known to have lower concentrations of polyphenols. MS-fingerprinting of Peruvian cocoa has also revealed distinct clustering between Criollo and non-Criollo varieties on PCA plots (Tran et al., 2015). Tran et al. (2015) have found that Peruvian Criollo had a much higher abundance of TMP prior to roasting, which is gives rise to its characteristic cocoa flavour/aroma, as compared to its Forastero counterpart.

Criollo varieties have also been reported to contain higher concentrations of linalool which is attributed to flowery and tea-like flavour. Ziegleder (1990) has observed co rrelations of higher linalool content of fine-flavour cocoa originating from South and Central America. One limitation of this study, however, is that it has not accounted for geographical variance. This was

recently confirmed by Qin et al. (2017) who found significantly higher concentrations of β-linalool in Trinitario and Criollo varieties. Our own investigations using proteomics have also revealed the notable presence of the protein linalool synthase in Criollo and Nacional hybrids of Ecuador and Brazil (personal communication with Neha Kumari, 07.02.2019, Kumari 2018).

The recent study by Qin et al. (2017) also elaborated on the differences in the composition of volatiles between unfermented beans of Criollo, Forastero and Trinitario varieties obtained from the Cocoa Germplasm Repository of the Spice and Beverage Research Institute in China. Analysis by GC-MS revealed Criollo and Trinitario varieties to have immensely higher concentrations of volatiles compared to Forastero cocoa. It was also clear that Forastero lacked the abundances of alcohols, esters and furans that contribute to the flavours in fine-flavour cocoa (**Figure 1.1**). Statistical analysis was able to differentiate between Criollo, Forastero and Trinitario hybrids through the compounds that contributed to their variance (**Table 1.1**).

1.3. Botany of *Theobroma cacao*

Theobroma cacao is an evergreen plant belonging to the family *Sterculiaceae*. Carl Linnaeus named it as such due to its deistic association in Mesoamerican society, and a Greek translation of the name is “food of the gods”. The genus contains 22 species, of which, *T. cacao* is the most common followed by *T. grandiflorum* (Lopes & Pires, 2015). *T. cacao* is a diploid species and has a chromosome number of 20 (Toxopeus, 1985).

1.3.1. General Biology

Being tropical plants, leaves of *T. cacao* are broad to encourage photosynthesis and transpiration. Flowers form along the trunk and branches of the tree, and 4-7 months after pollination, mature into fruits (pods). The seeds can begin to germinate within the pods or on the ground after pods are opened by animal or human intervention. Upon germination, the rootlet grows first and the cotyledons are subsequently raised above the ground. This is followed by the growth of the shoot and leaves which continue to grow straight until they reach about 1 m in height. The central stem/trunk then starts to divide into branches. The stem continues in orthotropic (vertical) growth while the branches grow in a plagiotropic (lateral) manner to create a canopy (Bartley, 2005; Toxopeus, 1985). Cocoa trees can reach heights of about 5 m (**Figure 1.2A**). The tap root systems of cocoa can extend up to 2 m below the surface. The root structure

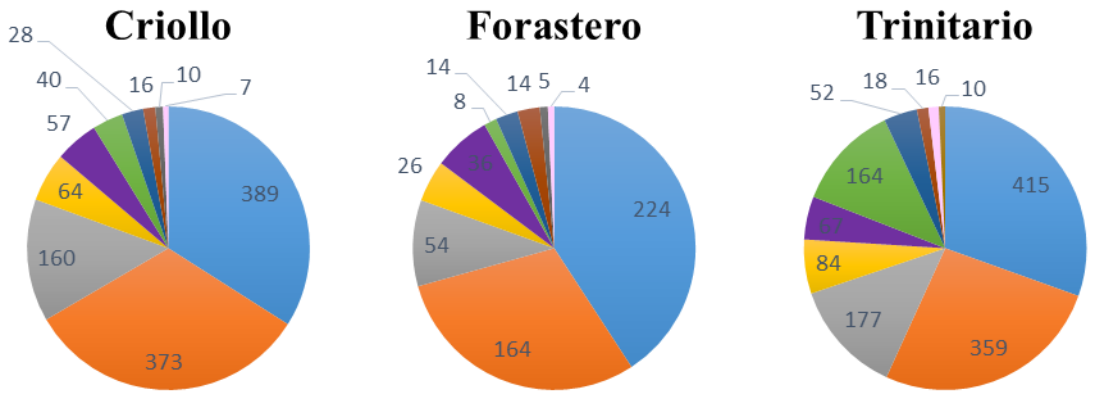


Figure 1.1 Variation of classes of volatile compounds between varieties of cocoa.
Relative amounts of **acids**, **hydrocarbons**, **alcohols**, **furans**, **phenols**, **esters**, **ketones**, **lactones**, **aldehydes** and **other volatiles compounds** as analysed by GC-MS in fresh *Criollo*, *Forastero* and *Trinitario* beans obtained from the Cocoa Germplasm Repository of the Spice and Beverage Research Institute in China. Adapted from Qin et al., (2017).

Table 1.1 Volatile compounds contributing to variance among cocoa varieties.
PCA analysis revealed the following volatile compounds to contribute to differentiating between the three main cocoa varieties. Analysis was done by GC-MS on fresh cocoa beans from the Cocoa Germplasm Repository of the Spice and Beverage Research Institute in China. Adapted from Qin et al., (2017).

Cocoa Variety	Compound
Criollo	Acetic acid, β-carophyllene, ethyl phenylacetate, 2-ethyl-1-hexanol, α-limonene, α-phellandrene, β-myrcene, phenethyl acetate, β-linalool, α-toluenol, furfural
Forastero	3-methylbutanoic acid, 2-(2-butoxyethoxy)ethanol, anethole, 2,4-pentanediol
Trinitario	2-phenylethanol, 2,3-butanediol, 1-pentanol, 2-pentanol, 2-heptanol, benzyl acetate, 1-phenylethanone, 2-nonanone, furaneol, nonanal, 2-acetylpyrrole, 3-carene

just below the surface is more lateral and can even stretch beyond the canopy of the tree, forming an intricately woven mesh that anchors the plant firmly to the ground while simultaneously taking up water.

Flowers are only produced on the woody parts of the tree that are at least two or three years

Figure 1.2 Plants of *Theobroma cacao* grown under greenhouse conditions in Bremen, Germany.
Seeds of *Theobroma cacao* were germinated and planted into pots on the campus of Jacobs University Bremen and grown under specific light, temperature and humidity conditions (A). Some of these plants were sent to the Botanika at the Rhododendron Park in Bremen, upon which they began to develop flowers (B); picture courtesy Daniel Papenberg. Pods (C) are the fruits of the cocoa tree containing the seeds within.

old, including the trunk and branches (**Figure 1.2B**). They are small in size consisting of two anthers. At a given time, a large number of flowers can be found because only 1-5% of the flowers are successfully pollinated to produce pods (Toxopeus, 1985). Pollination is naturally accomplished through insects. There have also been attempts at artificial pollination using mist-blowers which have more than doubled the yield in plantations in Brazil and Costa Rica. However, this was not possible on self-incompatible trees which cannot produce fruit with their own pollen (Toxopeus, 1985).

After fertilization occurs, the pod begins to grow. There are several growth stages involved, during which, the embryo develops, the ovule becomes filled with a jelly-like endosperm which is later consumed by the embryo, and there is a rapid accumulation of fat. It can be five to six months after pollination that the pods ripen (Toxopeus, 1985).

The pod (**Figure 1.2C**) consists of an outer fibrous husk, surrounding the seeds which are attached to a central placenta. Pods can differ in terms of size, colour and shape, depending on genetic and environmental factors (Bartley, 2005). Pod size varies from 12 to 30 cm in length and can contain 30-60 beans (Lopes & Pires, 2015; Toxopeus, 1985). The presence of anthocyanins leads to alluring shades of red splashed on to bright green backgrounds. This attracts animals that feed on the sweet, pulpy seeds within and start the cycle all over again.

1.3.2. Seeds

When referring to the cocoa bean in the production of chocolate, one is usually referring to the cotyledons, also known as nibs. However, the cocoa bean, in its entirety, is composed of the pulp, shell (testa), embryo and two cotyledons.

1.3.2.1. Pulp

The cocoa pulp is a white to yellow mucilaginous layer covering the bean with a moisture content of around 84%. The dry mass consists of the main sugars of glucose (4.5%) and fructose (4.8%), fat (1.6%), protein (0.2%), citric acid (1-3%), pectins (1-1.5%) and pentosans (2-3%) (Voigt & Lieberei, 2015). Endraiyani et al. (2017) have also found that cocoa pulp has heat-labile phenolic compounds, albeit in low abundance. Moreover, linalool, a constituent of fine-flavoured cocoa has been found by Ritter (1999) in the pulp of freshly harvested beans, that is assumed to diffuse into the bean during the course of fermentation.

The pulp is the main substrate for micro-organisms at the start of fermentation. Being rich in sugars and citric acid, virgin pulp favours the growth of yeasts. While yeasts break down the sugars in the pulp, their depectinization activity liquefies the pulp, making it drain off in the form of sweating. This effect not only encourages air ingress into the bean heap but also affects the seed shell, thereby altering the flux of substances to and from the bean (De Vuyst & Weckx, 2016a).

1.3.2.2. Shell (testa)

Directly surrounding the cotyledons of the bean is the protective testa. It contributes to as much as 14% of the dry weight and consists of two structural layers (Andersson, Koch, and Lieberei 2006; Voigt and Lieberei 2015). It is vital in regulating the movement of substances into and out of the bean and therefore contributes to differences between hybrids in terms of what flavour components are found within the bean at the end of fermentation. Additionally, the testa controls the entry and exit of water into the bean by means of the hilum, the inner contact zone of fruit pulp and seed coat, sclereid layer, hypostase, micropyle and endosperm cuticle structures. All these structures work together to maintain a turgor pressure within the bean that, in turn, regulates the flux of other substances (Andersson et al., 2006). Upon harvest, the placenta is detached from the individual beans, creating a small aperture at the location of the

hilum. This, together with the pressure-induced opening of the micropyle, become the main points of exposure of the cotyledons with their environment, especially for the diffusion of acid and other volatiles into the bean during fermentation (Voigt and Lieberei 2015). Since these events can be assumed to be important for the diffusion of acid into the bean, the termination of fermentation in Criollo beans after three or four days, for instance, would therefore mean that not as much acid would diffuse in, as compared to longer fermentations. Pre-drying of the beans can also reduce the diffusion of acid into the bean during fermentation and therefore influence the bean in different ways (Andersson et al., 2006).

1.3.2.3. Cotyledons

The main contributor to cocoa aroma or flavour in chocolate is the cotyledon, also known as nibs or kernels. *T. cacao* being a dicotyledonous plant, the two cotyledons surround the embryo and provide protection and nourishment during germination. There are two types of parenchyma cells that make up the beans: lipid-protein cells and polyphenol cells (Biehl, Passern, & Sagemann, 1982; Kim & Keeney, 1984). Lipid-protein cells have multiple, small vacuoles containing lipids, proteins and stage granules (Voigt and Lieberei 2015). Polyphenol cells, on the other hand, consist of a single large vacuole consisting of polyphenols and alkaloids. This suggests the wealth of lipids, proteins and polyphenols in cocoa beans.

During fermentation, acid breaks down cellular structures in the cotyledons and ruptures the cell membranes (Biehl et al., 1982). This results in the interaction of components within the bean and their ability to move around freely. The diffusion of acid and ethanol into the beans also results in the death of the embryo and a cascade of other reactions that begin to change the biochemistry of the bean. For the purpose of simplification, the cotyledons will hereon be referred to as beans or nibs.

1.4. Global Cultivation of Cocoa

Cocoa is commercially cultivated in the tropical regions of the world. Humid conditions of 80-90% are preferred along with ample rainfall (1,500 – 2,500 mm per year) or irrigation systems. Ideal soil pH is in the range of 5 to 7.5 (Fowler, 2009). Having originated from Mesoamerica, cocoa and its hybrids have spread through the regions of West Africa, South East Asia and South Asia. Currently, Ivory Coast is the world’s largest cocoa producer, accounting

for more than two-thirds of global cocoa production and around 70% of commercial cocoa is grown in West Africa (Fowler, 2009; Saltini, Akkerman, & Frosch, 2013). Around 95% of total cocoa cultivated is of the Forastero variety (Thompson, Miller, & Lopez, 2007).

Cocoa typically shares its farming region with other crops such as coffee, oil palm, rubber and banana. Trees are commonly planted in a density of 600 – 1200 per ha and usually done so in the shade to protect them from excessive sunlight (Lopes & Pires, 2015). That is why they are also intercropped and typically planted next to the taller and broad-leaved banana tree. This is especially critical in the first few years of its growth such that they grow to become healthy and robust trees (Obiri et al., 2007). The trees start yielding fruit between two or three years of age and are at their maximum productivity after six or seven years. Cocoa can keep bearing fruit for more than 30 years of its lifespan (Fowler, 2009).

The harvest of cocoa is performed during different seasons depending on the region. This is due to differences in genetics, disease outbreaks, as well as weather conditions such as temperature, rainfall, monsoon seasons, wind conditions, *etc*. Farmers distinguish between a main-crop season and a mid-crop season. Main-crop seasons involve greater yields and are also reported to have better fermentations as compared to mid-crop seasons where yields can fluctuate dramatically (Wood, 1985a). The sizes of the beans are also reported to vary between crop seasons. Crop season times can be diverse even within a continental geographical region. For instance, the main crop season in Brazil is from October to March, whereas those in Ecuador and Columbia are from March to June.

In addition to dealing with adverse weather conditions, cocoa trees also have to reckon with disease. An estimated 40% of cultivated cocoa is lost to disease (Lopes & Pires, 2015). Strains of Criollo are known to be more disease-susceptible as compared to the other variants. The so-called "blast", which wiped out a sizeable population of cocoa in the 17th century, is thought to be attributed to *Ceratocystis* wilt (Wood, 1985b). Today, black pod rot is the predominant disease-causing infection by *Phytophthora* spp. (Wood, 1985a). The symptoms of the disease include lesions on the pods, followed by pod rotting at the final stages where beans are completely damaged. West African cocoa also encounters *Badnavirus*, also known as the cocoa swollen shoot virus. The carrier of this virus is the mealybug and infection results in symptoms of leaf discolouration, swelling of the roots and shoots, and can even lead to complete death of the plant (Lopes & Pires, 2015). Witches' broom disease and frosty pod rot are

prevalent in the regions of Central and South America and caused by *Moniliophthora* spp. These also cause pod damage in the form of necrotic lesions and reduce the yield considerably (Lopes & Pires, 2015).

1.5. Post-harvest Processing

Following harvest, ripe cocoa pods go through a series of primary processing steps before the dry beans are shipped out to the factory for further processing. These steps are carried out directly on or near the plantations on which the trees are grown. Not all farmers follow a standardized procedure. Depending on the region, hybrid and plantation, various stages are omitted or undertaken in a different manner.

1.5.1. Pulp pre-conditioning

The pulp is the substrate for microbial activity during fermentation. Thus, changes in the properties of the pulp would lead to slight variations in the progression of fermentation. For instance, if yeast growth is inhibited at the beginning, this would also affect the growths of LAB and AAB and their subsequent metabolite secretions. Pulp pre-conditioning can be in the form of pod storage, de-pulping or bean spreading, and can result in changes to the total amount of pulp, as well as its moisture content, acidity and sugar content (Kongor et al., 2016).

1.5.1.1. De-pulping

Certain hybrids have an excess amount of pulp surrounding the beans which can lead to high levels of acid during fermentation. An example of this is in Brazilian cocoa, where the excess pulp is mechanically removed (Schwan and Wheals 2004). This can be achieved using presses, centrifuges or simply spreading beans out on a flat surface to stand for several hours (Kongor et al., 2016). There is also specialized de-pulping equipment used for making pulp juice which is common in Brazil.

Another mode of de-pulping is through the use of enzymes. Spraying a solution of pectinase on fresh cocoa beans is known to increase the quantity of extracted pulp by 23% as compared to the mechanical method (Schwan and Wheals 2004). However, enzymatic de-pulping can be a costly and prohibitive approach on large-scale production. Research is being done into improving the strains of yeast to enhance pectinolytic activity during fermentation.

1.5.1.2. Pod Storage

Much research has been carried out into the effects of pod storage on the fermentation and the quality of beans. Pod storage of between seven and 21 days has shown improvement in the quality of fermentation in Ghanaian cocoa as seen by the cut-test, as well as an acute influence over the chemical and physical characteristics of the beans (Afoakwa, Quao, Takrama, Budu, & Saalia, 2012, 2013; Kongor, Takrama, Budu, Mensah-brown, & Afoakwa, 2013). Nib acidification of Malaysian beans was also significantly reduced through this method, along with an increase in cocoa flavour (Meyer, Biehl, Said, & Samarakoddy, 1989).

1.5.2. Fermentation

Fermentation is arguably the most important of the post-harvest processes due to the complete morphing of the bean metabolome during this stage. It is a significant step up in the formation of cocoa flavour and aroma. Fermentation techniques are unique to regions and plantations across the globe. It generally entails the removal of beans from the pod which is usually manually done using machetes and knives. The beans are heaped up and left to stand for two to eight days, during which, microorganisms proliferate on the pulp of the beans and begin the process of fermentation (Thompson et al., 2007).

1.5.2.1. Types of Fermentation

One is able to categorize the techniques of fermentation in numerous ways. But for the sake of simplicity, fermentation will be considered to be either spontaneous or controlled.

Spontaneous fermentation is a naturally-driven process. Beans are usually piled up in boxes, drying platforms or baskets, or alternatively, heaped on top of and covered by banana leaves (Thompson et al., 2007). Bean heaps can range in mass from a few hundred to several thousand kilograms. Microbes, originating from the atmosphere, pod interior and surfaces, hands of personnel, and from the surfaces of the tools used for harvesting, fermenting and opening of pods, are transferred to the pulp of the beans. As the pulp is consumed, different groups of microbes dominate the microbiome. Yeast growth at the beginning of the fermentation starts to liquefy the pulp and produce sweating, which not only contains microbial CFUs, but also microbial and bean metabolites such as ethanol and organic acids. It is important that there is appropriate drainage for the sweating which would otherwise accumulate and ruin the

fermentation (Thompson et al., 2007).

Almost all trials of fermentation carried out on plantations are spontaneous. Consequently, there is a good number of fermentation trials that go awry, due to the lack of control over what sorts of microorganisms grow on the pulp, weather conditions, *etc*.

Controlled fermentation, therefore, aims to overcome the aspect of spontaneity through the use of starter cultures and microbial supplements. Investigations into the use of starter cultures in lab-scale, pilot-scale and commercial-scale trials of fermentation are manifold (De Vuyst & Weckx, 2016a). Certain trials have been conducted to simulate conventional fermentation in the lab for the study of fermentation in isolated systems (Moens, Lefeber, & De Vuyst, 2014; Pereira, Miguel, Ramos, & Schwan, 2012), while others have used microbial cocktails to achieve a desired output in terms of reliability or flavour profiles (Buamah, Dzogbefia, and Oldham 1997; Leal Jr et al. 2008; Lefeber et al. 2012; Meersman et al. 2016; Pereira, Soccol, and Soccol 2016; Sandhya et al. 2016; Schwan 1998). These studies have employed various strains of native and engineered yeast, LAB and AAB.

1.5.2.2. Microbiology

Yeasts begin to dominate the initial stages of fermentation, feeding off the sugar- and acid-rich pulp (**Figure 1.3**). The low pH of the pulp, encourages the growth of yeasts over other microbes. Yeasts can reach population of 10^8 CFUs/g pulp during the first two days (Camu et al. 2007; Schwan and Wheals 2004). Simultaneously, pectin is broken down in the pulp as yeasts secrete pectinolytic enzymes. As they multiply, oxygen within the heap is consumed and, at a certain point, anaerobic respiration takes over. This results in the accumulation of ethanol in the pulp which may also diffuse into the bean. Ethanol production is an exothermic process, and together with respiration releases heat which is distributed over the bean heap and can raise its temperature to around 10°C above ambient temperature (De Vuyst & Weckx, 2016a). Moreover, yeasts produce organic acids as well as volatile compounds such as esters and alcohols which diffuse into the bean and contribute to flavour (Schwan and Wheals 2004).

A vast repertoire of yeasts has been isolated from spontaneous fermentation trials. *Saccharomyces cerevisiae* and *Hanseniaspora opuntiae* are the most dominant species found (Schwan and Wheals 2004). Thermo-tolerant *Pichia kudriavzevii* dominate the latter stages of yeast growth. Also found are *Hanseniaspora guillermondii*, *Hanseniaspora thailandica*,

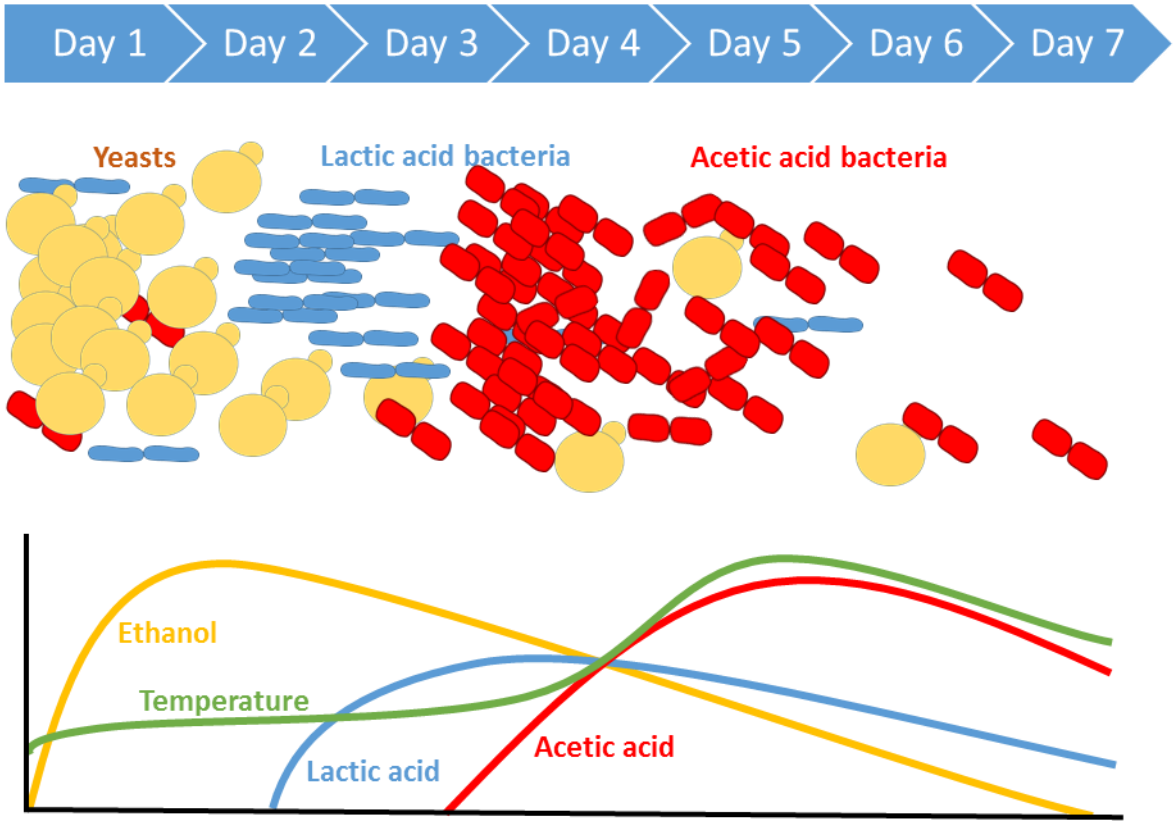


Figure 1.3 Microbial activity during cocoa bean fermentation.
A successive growth pattern of yeast, acetic acid bacteria and lactic acid bacteria populations grow on the pulp of the bean, simultaneously producing metabolites of ethanol, acetic acid and lactic acid, respectively, and causing rises in temperature within the bean pile.

Kluyveromyces marxianus, *Pichia anomala*, *Pichia fermentans*, *Pichia kluyveri*, *Pichia manshurica*, *Pichia membranifaciens*, *Candida* spp., *Schizosaccharomyces* spp., *Torulaspora* spp. and other *Saccharomyces* spp. among others (Daniel et al. 2009; Schwan and Wheals 2004; De Vuyst and Weckx 2016).

LAB start to thrive under the anaerobicity generated by yeast proliferation and are found to co-exist with yeasts. The strains of LAB are typically, microaerophilic, acid-tolerant, ethanol-tolerant and fructophilic bacteria (De Vuyst & Weckx, 2016a). LAB can either be homo- or hetero-fermentative and can ferment a wide range of saccharides. Lactic acid is the main secreted product but LAB can also metabolize sugars into acetic acid, ethanol and other organic acids (De Vuyst & Weckx, 2016b). Lactic acid is mainly produced through the Embden-Meyerhof pathway, and sometimes through the hexose monophosphate pathway (Schwan and Wheals 2004).

Various strains of *Lactobacillus* spp., *Leuconostoc* spp. *Fructobacillus* spp. and

Lactococcus spp. were found in spontaneous fermentation, with *Lactobacillus fermentum* being the most prevalent. It has been argued by some that lactic acid bacteria do not contribute to the flavour and aroma of cocoa and, if anything, rather ruin cocoa flavour by the excess production of lactic acid (Ho, Zhao, & Fleet, 2015).

AAB are the last of the main microbial influencers of cocoa aroma and flavour and start to flourish upon air ingress into the bean heap. This happens as a result of turning or mixing the beans within the heap, in combination with the drainage of the liquefied pulp, allowing a better air flow through the heap (De Vuyst & Weckx, 2016a). Populations of AAB can rise as high as 10^9 CFUs/g bean and the main species observed are *Acetobacter pasteurianus*, *Acetobacter tropicalis*, *Acetobacter syzygii*, *Acetobacter ghanensis*, *Acetobacter senegalensis* and *Gluconobacter* spp. (Camu et al., 2007; De Vuyst & Weckx, 2016a; Lefeber, Gobert, Vrancken, Camu, & De Vuyst, 2011). *Gluconobacter* spp. are less common and oxidize glucose into gluconic acid, causing another yeast growth phase that can result in off-flavours (De Vuyst & Weckx, 2016a). The favourable *Acetobacter* species oxidize the available ethanol and lactic acid into acetic acid and acetoin (Adler et al., 2014). The reactions involved in these bioconversions are highly exothermic and push the temperature of the bean pile to highs of 55°C (Schwan and Wheals 2004). Subsequently, this encourages diffusion of acetic acid into the beans as well as its evaporation, due to its volatility. Influx of acetic acid into the bean plays a key role in the development of flavour precursors. However, an overabundance of acetic acid within the bean is deleterious to cocoa flavour. Thus, it is vital that the growth of AAB stays within reasonable limits.

Other microorganisms also grow on the beans during fermentation. These include facultative anaerobic enterobacteria such as *Tatumella* spp., filamentous fungi, such as *Aspergillus* spp., *Fusarium* spp., *Mucor* spp. and *Penicillium* spp., and other bacteria such as *Bacillus* spp. (Schwan and Wheals 2004; De Vuyst and Weckx 2016a). Each of these also has the potential to influence cocoa flavour by either affecting the growth dynamics of yeast, LAB and AAB, or secreting secondary metabolites. Moreover, Salmonella has also been observed to grow considerably during fermentation and is therefore known to be one of the biggest hindrances to chocolate consumption (Nascimento et al., 2013; Werber et al., 2005).

1.5.2.3. Model fermentation systems

In addition to trials of spontaneous and controlled fermentation that occur in the field, researchers have also undertaken trials of fermentation under much more controlled conditions. These include trials that have been performed with beans on much smaller scales of 500 g – 3,000 g, under septic or aseptic conditions, or involving pulp simulation media.

Adler et al. (2014) and Lefeber et al. (2010, 2011) have employed pulp simulation media in order to investigate the growth dynamics of AAB and LAB and to understand the mechanism behind the formation of acetic acid in AAB. The experiments were performed in fermenters where pH, pressure and temperature were monitored. They were able to simulate the general conditions of a bean heap during fermentation and achieve a high degree of reliability due to the level of control achieved in growing the bacteria. Lee et al., (2018) have also made use of an artificial pulp system in order to re-ferment dried and unfermented beans.

Additionally, small-scale spontaneous fermentation has also been carried out using controlled temperature conditions. Romanens et al. (2018) conducted a trial of fermentation with 1-kg heaps in plastic buckets and exposed them to a temperature regime that would usually occur in a commercial-scale fermentation. Sweating was allowed to drain off through slits in the buckets and the beans were mixed every 24h. They were able to successfully replicate a large-scale fermentation in terms of microbial growth profiles and the consumption of substrates in the pulp.

Finally, aseptic incubation of cocoa beans in solutions have also been in use for the last 30 years and have revealed much about the biochemistry of cocoa and the metabolomic flux during fermentation. Studies in the 1980s were the first to employ such a system which has enhanced our understanding of proteolysis within the bean and how acidification affects it (Biehl, Brunner, Passern, Quesnel, & Adomako, 1985; Biehl et al., 1982; Biehl & Passern, 1982). Eyamo Evina et al. (2016) and Kadow et al. (2015) have also conducted similar trials in order to understand the behaviour of other components of the metabolome during fermentation and their effect on flavour.

1.5.3. Drying

Following fermentation, beans are dried immediately to prevent spoilage. Since cocoa plantations are in hot, tropical regions, drying is usually done by spreading beans out on to flat

surfaces and leaving them in the sun. The surfaces can include wooden platforms, concrete terraces or bamboo mats. The duration of drying depends heavily on the weather. A week of sunny weather is able to dry beans to the desired moisture content of 6% (Fowler, 2009). There are also artificial methods of drying which include wood fires below the drying platform and forced air dryers. Downsides to artificial drying techniques involve beans drying out too quickly and trapping acid within the bean, and smoke infusing into the beans resulting in off-flavours (Fowler, 2009).

1.6. Manufacture of Chocolate

Once the beans are dried, they are ready for the long journey, typically to factories in the Western world, where they will be processed into chocolate. Shipment logistics can involve long periods of storage. Hence, the moisture content needs to be within acceptable levels to prevent moulding.

Upon reaching the factory, all debris is separated out and the beans are de-shelled and broken using a bean breaker and winnower. The shells are separated out and only the broken beans (nibs) are processed further (**Figure 1.4**).

1.6.1. Alkalization

The process of alkalization is not standard in the manufacture of chocolate but is nevertheless utilized to a great extent. This is done to get rid of excess acidity within the beans. The degree of alkalization depends on the region the beans come from and the hybrid. Potassium carbonate is the usual alkaline solution that the beans are incubated in until a colour change is observed (Kamphuis, 2017). Alkalization reduces acidity and astringency in cocoa (Afoakwa et al., 2008).

1.6.2. Roasting

Roasting is a crucial process in flavour formation. It is at these high temperatures that the essential reactions for flavour formation take place. Roasting also further depletes volatile acids within the beans. Beans are roasted at temperatures of 120 – 150°C for periods of 5 – 120 min (Afoakwa et al., 2008). The time/temperature relationship of the roast designates whether it is a low, medium or high roast. The roast is dependent on the origin of the bean and can vary

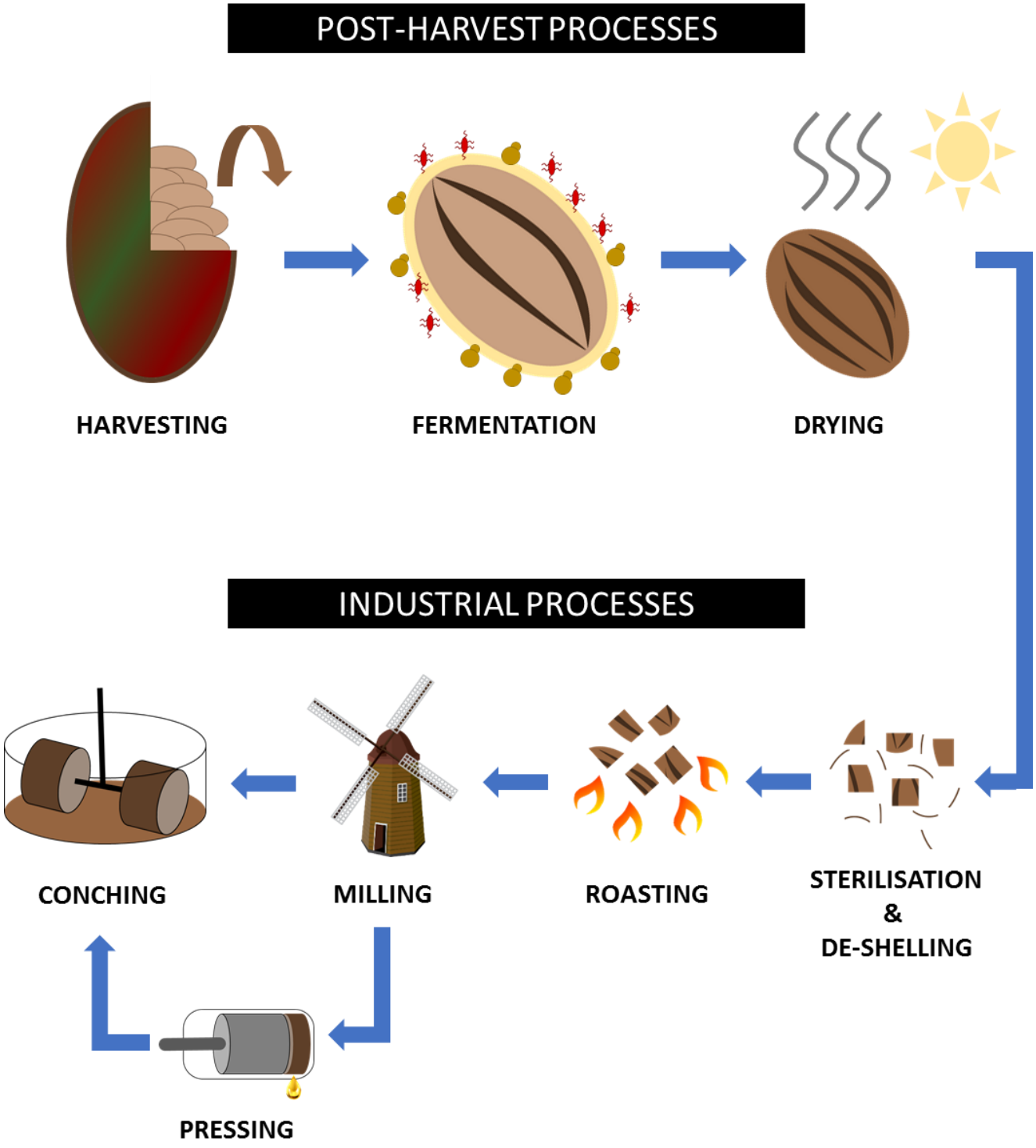


Figure 1.4 Processes involved in the manufacture of chocolate.
Following the harvest of beans, manufacture can be divided into post-harvest processes, which take place on the plantations where the beans are harvested, and industrial processes, which take place in the factories where chocolate is produced.

considerably between batches of beans.

Unlike fermentation, roasting is a more controlled and optimized process. There is ample research into different roasting specifications for desired taste and aroma attributes and the effects of each parameter on the formation of specific compounds (Arlorio et al., 2008; Farah, Zaibunnisa, & Misnawi, 2012; Rocha, Santana, Soares, & Bispo, 2017). This process, however, is an empirical one and parameters cannot be standardized through production pipelines. Following the roast, the nibs are milled in order to obtain cocoa liquor. Liquor is the final product of the main cocoa production chain, and is processed in a number of ways to obtain various products such as cocoa powder, cocoa butter and chocolate.

1.7. Biochemistry of Cocoa Beans and Flavour Formation

1.7.1. Lipids

Like most seeds, lipids constitute the greater proportion of cocoa beans, making up 46-57% of the dry weight (Borchers, Keen, Hannum, & Gershwin, 2000; Liendo, Padilla, & Quintana, 1997). The lipid fraction is also commonly referred to as cocoa butter, more so, after the lipid has been pressed out of the roasted powder. Saturated fats of stearic acid and palmitic acid dominate the lipid fraction, making up to 60% of cocoa lipid. Oleic acid is the main polyunsaturated fat and contributes to about 34% of the total fat content, with linoleic acid accounting for 2% (Borchers et al., 2000). These fatty acids exist in the form of TAGs, of which POP, POS and SOS are the main combinations of stearic, oleic and palmitic acids (Lipp et al., 2001). Fat content and composition can vary between hybrids and geographic origin, although within narrow limits. It has been observed that during the course of fermentation, the profiles of fatty acids and TAGs were altered. For instance, fermented and dried beans contained elevated levels of linoleic and palmitoleic acids (Sirbu, 2018). The degree of fatty acid unsaturation would also dictate the melting point of the resulting cocoa butter. The characteristics of cocoa butter are important to the texture of the manufactured chocolate. Despite lipids being the largest bean component, it does not contribute directly to taste and aroma and has therefore been disregarded in almost all studies relating to flavour characteristics of cocoa beans.

1.7.2. Polyphenols

Cocoa is incredibly rich in polyphenols, making up to 18% of the dry weight (Kim &

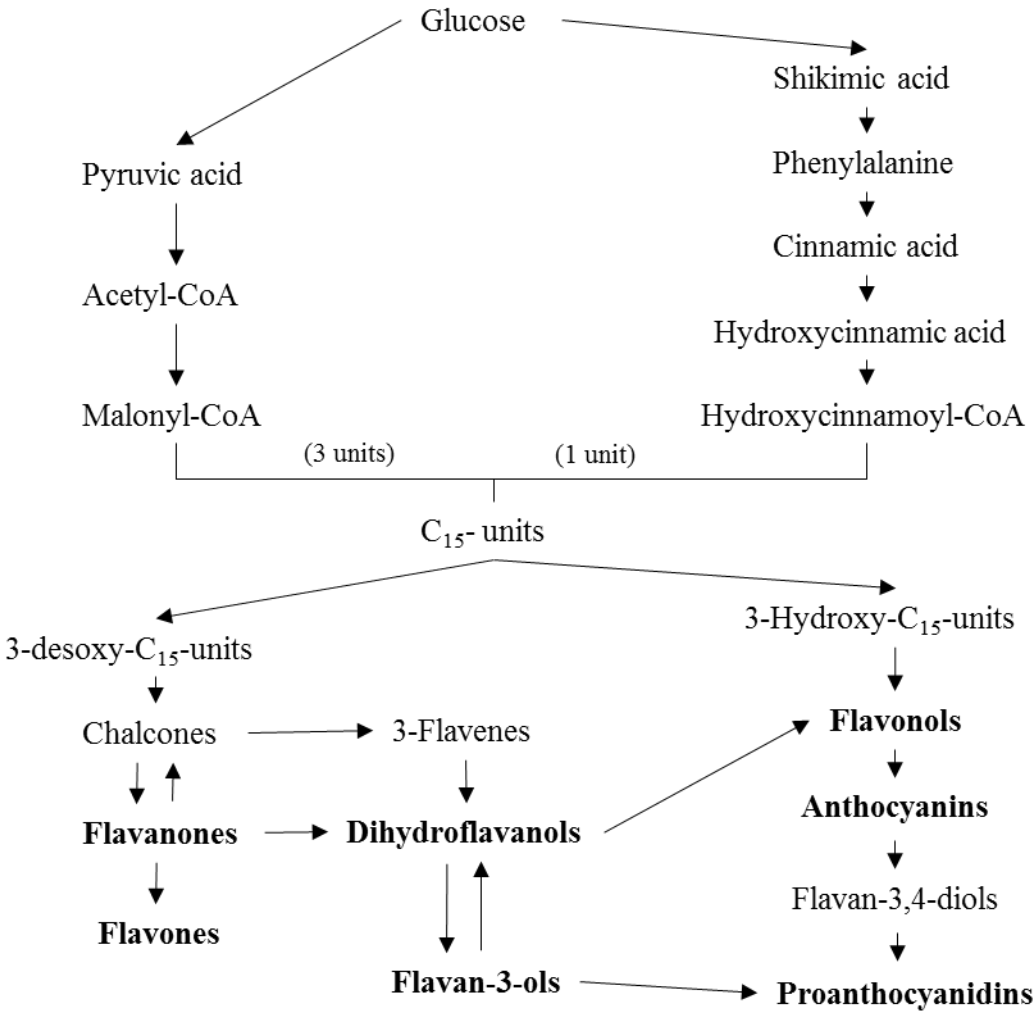


Figure 1.5 Biosynthesis of polyphenols in plants.
A schematic representation of the pathways involved in polyphenol biosynthesis in plants. Adapted from Wollgast & Anklam, (2000).

Keeney, 1984) and even contributing as much as 20% to the dry bean weight in Forastero hybrids (Wollgast & Anklam, 2000). While they are renowned for their extraordinary health benefits, they also contribute to bitterness and astringency of chocolate. Criollo beans are known to have two-thirds of the polyphenol content of Forastero beans. They are the products of secondary metabolites arising from two main biosynthetic pathways: the shikimate pathway and the acetate pathway (**Figure 1.5**) and stored in the polyphenol cells (Wollgast & Anklam, 2000). While there are 10 classes of polyphenols, the main ones found in cocoa beans are flavan-3-ols (catechins), anthocyanins and proanthocyanidins.

Flavan-3-ols account for 37% of total polyphenols in unfermented beans and there are four

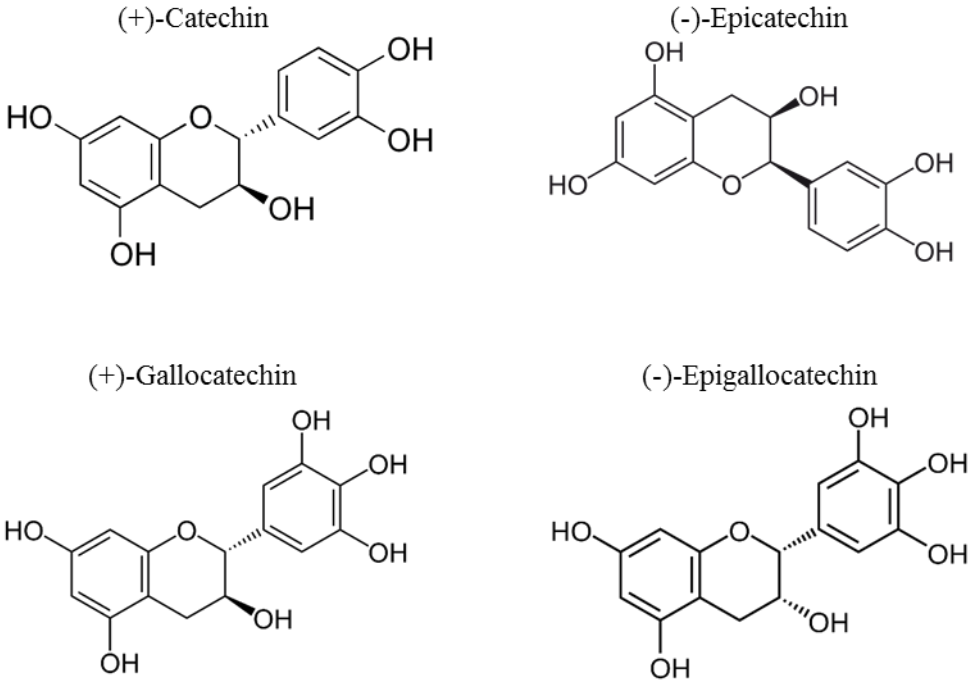


Figure 1.6 Structures of flavan-3-ol monomers in cocoa.
Catechin and epicatechin are among the most abundant polyphenols in cocoa and are classed as flavan-3-ols.

types in cocoa: (-)-epicatechin, (+)-catechin, (+)-gallocatechin and (-)-epigallocatechin (**Figure 1.6**). (-)-epicatechin is the most abundant of these (Jalil & Ismail, 2008). Most polyphenols in cocoa beans exist in the form of (-)-epicatechin and its oligomers, known as procyanidins, which together constitute a branch of proanthocyanidins. Procyanidins with oligomers of up to ten subunits have been reported (Borchers et al., 2000; D'Souza et al., 2017).

Anthocyanins give Forastero beans their characteristic purple colour and are therefore much less abundant in Criollo beans. Cyanidine-3-α-L-arabinoside and cyanidine-3-β-D-galactoside are predominant in this fraction (Voigt & Lieberei, 2015).

Other phenolic compounds such as quercetin, quercetin glycosides, clovamide, deoxyclovamide and caffeic, ferulic, gallic and ρ-coumaric acids have been also reported in unfermented beans (Borchers et al., 2000; D'Souza et al., 2017; Wollgast & Anklam, 2000), which are known to be bioactive molecules. The composition of phenolic compounds within cocoa beans has been observed to be geographically classifiable (D'Souza et al., 2017).

As the membranous structures within cocoa beans break down during fermentation (Biehl et al., 1982), polyphenols are released from their storage bodies and dispersed throughout the bean. They undergo oxidation to condensed, high molecular weight, insoluble tannins, which

could either be a non-enzymatic process, or one that is aided by the activity of polyphenol oxidase (Voigt & Lieberei, 2015). Polyphenol oxidase activity is drastically reduced during the first 48 hours of fermentation (Hansen, Del Olmo, & Burri, 1998). The oxidation of polyphenols also leads to a browning effect. Simultaneously, anthocyanins and flavonoid glycosides are hydrolysed by glycosidases leading to a loss in colour, and procyanidins are converted into more complex forms (Afoakwa et al., 2008; Aprotosoaie, Luca, & Miron, 2016). Polyphenols have also been reported to leach out of the beans in the sweating during the initial stages of fermentation as the pulp liquefies. The loss in polyphenol content during the course of post-harvest processing can be attributed to both genetic and environmental factors (Aprotosoaie et al., 2016). Certain studies have also reported preserved or elevated levels of polyphenols post-fermentation (Albertini et al., 2015; D'Souza et al., 2017; Romanens et al., 2018).

Polyphenols are thermo-labile structures. Drying and roasting stages therefore not only enhance their degradation but also lead to other reactions, the most significant of which is the Maillard reaction, fundamental to cocoa flavour development.

Maillard reactions, also known as non-enzymatic browning reactions, are the result of condensation reactions between amino groups and carbonyl groups, forming Schiff bases which would eventually rearrange to Amadori or Heyns products (Lund & Ray, 2017). The amino groups involved would typically be from peptides or amino acids and the carbonyl groups would stem from sugars or polyphenols. Amino acids form conjugates with sugars and polyphenols to form a repertoire of different compounds, some resulting in off-flavour and some, in fine-flavour (Ziegleder, 1991). These molecules are subsequently modified to α-dicarbonyl species which are extremely vulnerable to nucleophilic attack from compounds such as amines. Condensation reactions of these intermediates with free amino acids lead to the formation of Strecker aldehydes. The rate, extent and course of Maillard reactions are succumb to several factors such as temperature/time combinations, pH and water activity (Lund & Ray, 2017). The role of pH is vital as it influences which pathways the reactions follow.

Maillard reactions can be complex, to say the least, and the mechanisms involved at the final stages of the reactions are still not fully understood (**Figure 1.7**). Nevertheless, the key products leading to desirable flavour are Strecker aldehydes, furfural products, pyrazines and pyrroles (**Table 1.2**).

In all the stages of post-harvest processes, cocoa can lose up to 90% of polyphenol content

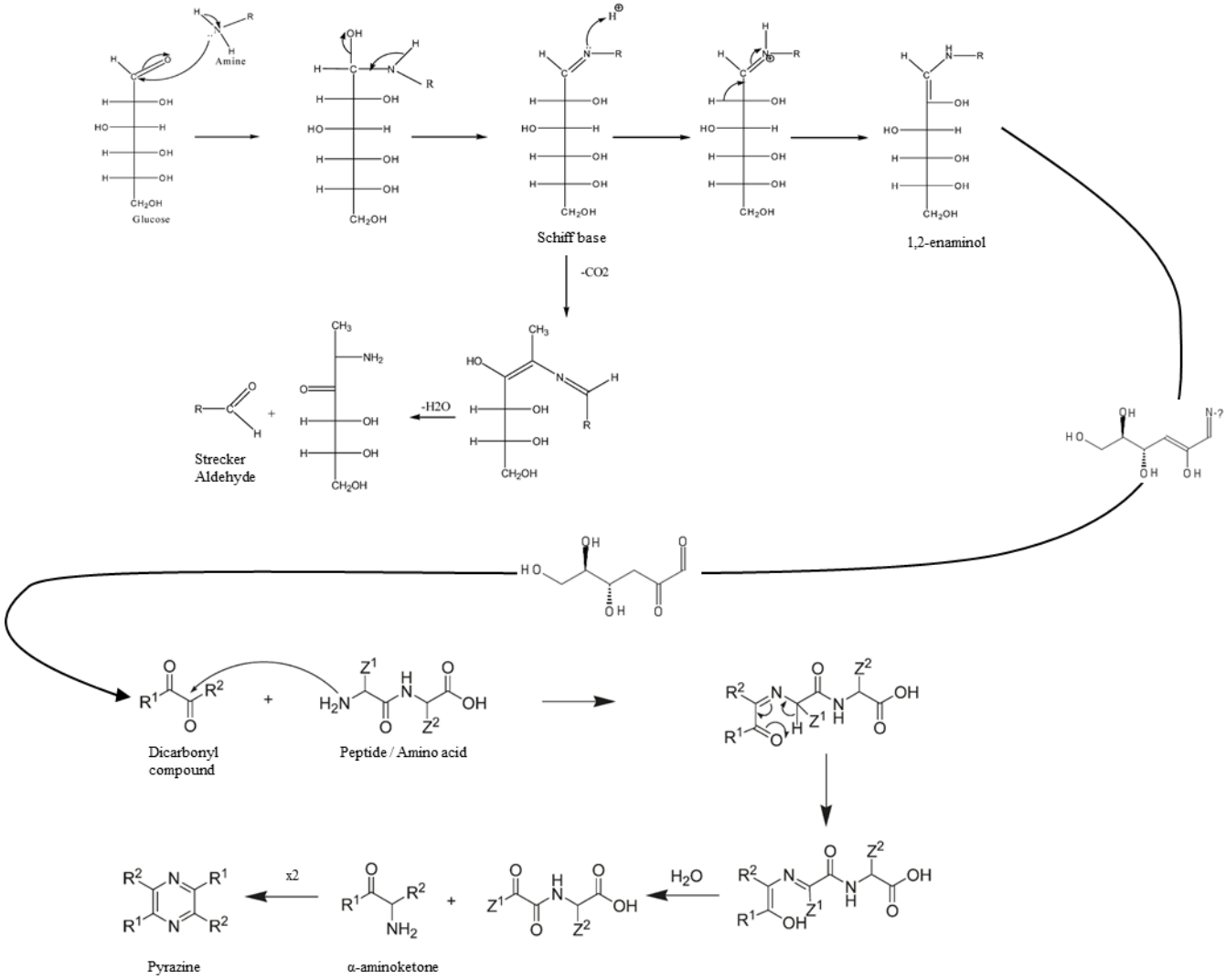


Figure 1.7 Proposed reaction scheme and mechanisms of Maillard reactions.
Maillard reactions that occur in cocoa are known to result in Strecker aldehydes and Pyrazines, produced from the sugars and amino acids and peptides that are formed after fermentation. Adapted from Afoakwa et al. (2008), Lund and Ray (2017), and Van Lancker, Adams, & De Kimpe, (2010).

and, hence, a lot of its health benefits. Currently, there are ideas of bypassing certain post-harvest processes such as fermentation and alkalisation in order to preserve polyphenol content and sell the resulting chocolate a health-based product (Bernaert, Blondeel, Allegaert, & Lohmueller, 2012).

1.7.3. Proteins

Proteins account for 10-16% of the dry weight of unfermented beans. Protein content has been seen to vary depending on geographical origin and can be a result of genetic or environme-

Table 1.2 Volatile compounds contributing to cocoa-specific flavour.
Volatile compounds generated during the course of post-harvest and industrial processes that contribute to cocoa flavour. These compounds also have additional aroma notes. Adapted from Aprotosoaie, Luca, & Miron, (2016) and Afoakwa et al., (2008).

Compound	Class of compound	Aroma	Flavour
1-propanol	Alcohol	Sweet, candy	Sweet chocolate
2,3-butanediol	Alcohol	Cocoa butter	Sweet chocolate
2-methylpropanal	Aldehyde	Chocolate	Not determined
3-methylbutanal	Aldehyde	Chocolate	Not determined
2-methylbutanal	Aldehyde	Chocolate	Not determined
2-phenyl-2-butenal	Aldehyde	Cocoa, sweet, roasted, rum	Not determined
ethyl cinnamate	Ester	Sweet, cinnamon-like	Sweet chocolate
5-methyl-2-furfural	Furfural	Sweet, caramel	Sweet chocolate
2-acetylfuran	Furan	Sweet, balsamic, slight coffee	Sweet chocolate
3-phenylfuran	Furan	Cocoa, green, mint	Not determined
2-acetylpyrrole	Pyrrole	Chocolate, hazelnut, cocoa, roasted	Sweet chocolate
furfurylpyrrole	Pyrrole	Roasted, chocolate, green	Not determined
3-ethyl-2,5-dimethyl-1H-pyrrole	Pyrrole	Cocoa, hazelnut, coffee, roasted	Not determined
2-methylpyrazine	Pyrazine	Nutty, chocolate, cocoa, roasted nuts	Sweet chocolate, nutty
2-ethyl-5(or 6)-methylpyrazine	Pyrazine	Cocoa, roasted, green	Not determined
2,5-dimethylpyrazine	Pyrazine	Cocoa, roasted nuts	Sweet chocolate, nutty
2,3-dimethylpyrazine	Pyrazine	Caramel, cocoa	Sweet chocolate
2,3,5-trimethylpyrazine	Pyrazine	Cocoa, roasted nuts, peanut	Sweet chocolate, nutty
2,3,5,6-tetramethylpyrazine	Pyrazine	Chocolate, cocoa, coffee	Sweet chocolate
2,3,5-trimethyl-6-ethylpyrazine	Pyrazine	Candy, sweet	Sweet chocolate
3,5(or 6)-diethyl-2-methyl-pyrazine	Pyrazine	Cocoa, chocolate, rum, sweet, roasted	Not determined

nmental factors (Kumari et al., 2018). Cocoa is rich in storage protein and was previously reported to contain a globulin fraction of 43% and an albumin fraction of 52% of the total protein content (Voigt & Biehl, 1993). However, we now know these values to be different, as elucidated recently, and are around 23% and 14% for the globulin and albumin fractions, respectively (Lerceteau, Rogers, Pétiard, & Crouzillat, 1999). The globulin fraction is represented by a vicilin (7S)-class globulin (hereon referred to as vicilin) which was found to be made up of three subunits: 47 kDa, 31 kDa and 15 kDa (Kratzer et al., 2009), and more recently, a 44.3 kDa subunit along with a pI-shifted variant of the 47 kDa subunit (Kumari et al., 2016). A multitude of other proteins have been resolved and identified using 2D-PAGE combined with MALDI-ToF-MS analysis, the most abundant of which were albumin (CAA39860), 47 kDa vicilin (EOY05738.1), peroxidase (XP_007019580), lipoxygenase (EOY28236.1), and basic chitinase (EOX97499.1) (D'Souza et al., 2018; Kumari et al., 2018).

Proteins are broken down into peptides and free amino acids which form the building blocks of cocoa aroma and flavour. The enzymatic proteolysis of vicilin is what is known to be most attributed to the formation of flavour and aroma precursors (Voigt et al. 1994, 2016). There are two types of proteases in effect: endoproteases and exoproteases. The main endoprotease

activity has been attributed to that of aspartic proteinase of which two variants have been observed (Laloi, Mccarthy, Morandi, Gysler, & Bucheli, 2002). Aspartic proteinase cleaves between hydrophobic amino acid residues (Bytof, Biehl, Heinrichs, & Voigt, 1995). The two aspartic proteinase variants of TcAP1 and TcAP2 have been seen to contain high sequence similarity to typical plant aspartic proteinase. Transcriptional analysis revealed TcAP2 to be the more expressed variant, and its mRNA is found predominantly in the early stages of bean development. This suggests that as the bean matures, less and less TcAP2 is produced. Plant proteases are naturally intended to degrade storage protein during germination such that the embryo can metabolize the free amino acids for protein synthesis. While transcriptional levels of TcAP2 remained fairly stable during germination, levels of TcAP1 began to rise (Laloi et al., 2002), highlighting the importance of both these enzymes in the life cycle of *T. cacao*. Recombinant TcAP2 and a purified complex of aspartic proteinase from cocoa beans were both observed to have a pH optimum of 3 using haemoglobin substrate *in vitro* (Guilloteau, Laloi, Michaux, Bucheli, & McCarthy, 2005; Laloi et al., 2002). Enzyme activity of the protein complex was optimal at a temperature range of 42-47°C.

Endoprotease activity is also aided by exoprotease activity, specifically for the cleavage into free amino acids. An aminopeptidase, cleaving from the N-terminus, and a carboxypeptidase, cleaving from the C-terminus work together to achieve this (Hansen et al., 1998). Purified extracts of the serine carboxypeptidase and the leucine aminopeptidase were found to have optimum working pH values of 5.5 and 7.0, respectively, using oligopeptides in *in vitro* reactions (Bytof et al., 1995; Hansen et al., 1998).

There has not been much study conducted on aminopeptidase activity, most likely because of its relatively high pH working range. However, studies into the combined activities *in vitro* of the aspartic proteinase and the carboxypeptidase on extracts of vicilin have brought to light the pH-dependant generation of cocoa- and nutty-specific aroma precursors (Voigt, Textoris-Taube, and Wöstemeyer 2018). It has also become clear that hydrophilic peptides and hydrophobic amino acids, generated by the *in concerto* activities of both enzymes, contribute immensely to cocoa-specific aroma components (Janek et al. 2016; Kirchhoff, Biehl, and Crone 1989; Voigt et al. 2016).

1.7.4. Carbohydrates

Carbohydrates in cocoa are present in the form of free sugars (2-4% of dry weight) and polysaccharides (12% of dry weight). The main free sugars are glucose, fructose, sucrose, mannitol, inositol, arabinose, xylose, galactose and sorbose, of which, sucrose represents 90% of the free sugars in unfermented beans. Polysaccharides exist in the forms of starch, pectins, cellulose, pentosans and mucilage (Aprotosoaie et al., 2016). Reineccius, Andersen, Kavanagh, & Keeney, (1972) have found that the composition of free sugars changes dramatically in the bean during fermentation. Sucrose is seemingly broken down into glucose and fructose, and levels of mannitol, inositol and sorbose rise during the early stages of fermentation. Glucose and fructose at the end of fermentation and drying would not only contribute to the sweetness of the manufactured cocoa, but would also go on to play a major role in flavour development during the process of roasting through Maillard reactions as mentioned in section 1.7.2.

1.7.5. Methylxanthines (alkaloids)

Unfermented cocoa beans contain roughly 4% of methylxanthines in dry weight. Caffeine, theobromine and theophylline are the major methylxanthines in cocoa, with theobromine being the most prevalent. Caffeine and theophylline are only present in traces but together with theobromine, contribute to the bitterness of cocoa. They are also stored in a single large vacuole in polyphenol cells and caffeine is known to form loose complexes with epicatechin (Aprotosoaie et al., 2016; Jalil & Ismail, 2008).

During fermentation, methylxanthines diffuse out of the bean, leading to a reduction of about 30%. There is a greater loss of caffeine (50-54%) as compared to theobromine (38-40%). Consequently, bitterness is reduced in the resulting cocoa. Furthermore, these alkaloids can also migrate into the lipid fractions of the bean, which is aided by membranal rupture during fermentation (Aprotosoaie et al., 2016).

Methylxanthines have been widely investigated for their adverse health effects. There are studies supporting the pro-oxidant activity of these compounds and their stimulative effects on the cardiovascular and central nervous systems (Borchers et al., 2000; Eteng, Eyong, Akpanyung, Agiang, & Aremu, 1997). In addition, caffeine and theobromine have been reported to be the cause of chocolate cravings (Smit & Blackburn, 2005).

1.7.6. Volatiles

There are a range of volatile compounds reportedly present in cocoa beans during the course of chocolate manufacture. These compounds are of the classes of alcohols, aldehydes, esters, ketones, fatty acids, organic acids, phenols, pyrazines and other compounds. The majority of volatiles develop over the course of post-harvest and industrial processes (Bonvehí, 2005; Rodriguez-Campos et al., 2012).

1.7.6.1. Alcohols and esters

The cotyledons of raw cocoa beans are lacking in significant quantities of volatile compounds. However, during fermentation, microbial activity on the pulp of the beans gives rise to the formation of various alcohols and esters, which, subsequently diffuse into the bean (Steensels & Verstrepen, 2014). Higher alcohols and esters confer fruity, floral and herbaceous notes. Their concentrations are seen to increase during the course of fermentation, and then decrease during drying and roasting of the beans due to their volatility (Rodriguez-Campos et al., 2012). Aside from linalool, benzyl alcohol and 2-phenylethanol are also found to increase within the bean during fermentation and they contribute to floral aromas. 2-methyl-1-propanol and 3-methyl-1-butanol were also found in relatively high abundance among alcohols and increased during the course of fermentation. Moreover, amyl alcohols are esterified during fermentation and, after statistical analysis, some argue that the degree of esterification can be used as a reliable fermentation index (Rodriguez-Campos, Escalona-Buendía, Orozco-Avila, Lugo-Cervantes, & Jaramillo-Flores, 2011). 2-phenylacetate is a floral- and honey-odorous ester, characteristic to Asian cocoa liquor (Aprotosoaie et al., 2016). Ethyl-2-methylbutanoate, which is also found in significant quantities, contributes to fruity notes.

The main microbial contributors to this phenomenon are yeasts which are able to secrete a range of volatile metabolites that can have a profound influence over flavour, as already seen in the wine industry (Lambrechts & Pretorius, 2000). Meersman et al., (2016) discovered that one is able to influence the flavour profile of chocolate just by varying the strain of yeast growing on the pulp. Using strains of *S. cerevisiae*, *C. fabianii* and *P. kluyveri*, they observed that different strains resulted in varied profiles of acetate esters, which are responsible for fruity notes in chocolate. Hybridization of parental strains led to a stable and greater production of acetate esters, but also compromised cocoa and roasted aromas. Moreover, they found that

concentrations of aldehydes and pyrazines were affected by different yeasts strains. Nevertheless, this could be attributed to indirect effects such as the amount of acidification in the bean, leading to proteolysis.

1.7.6.2. Aldehydes and ketones

Aldehydes are compounds crucial to cocoa flavour. They are mainly formed through the high-temperature-driven Strecker degradation reactions of free amino acids during roasting, but may also appear during stages of fermentation or drying. 2-methylbutanal and 3-methylbutanal, both, confer malty and chocolate aromas are found to increase in the bean during fermentation and are thought to originate from LAB using precursors of leucine and isoleucine (Rodriguez-Campos et al., 2012). Interestingly, the concentrations of these compounds are unaffected by the durations of fermentation and drying. Ketones such as acetoin and acetophenone also increase in abundance during fermentation and impart creamy and sweet/floral notes, respectively. Acetoin appears to be a precursor of TMP (Hashim, Selamat, Muhammad, & Ali, 1999). Longer fermentation trials involving higher temperatures were found to increase the concentrations of these favourable aldehydes and ketones (Rodriguez-Campos et al., 2012).

1.7.6.3. Volatile acids

During fermentation, the concentrations of volatile acids are also known to rise substantially, especially that of acetic acid, which is produced by AAB through the oxidation of ethanol, and diffuses into the bean (Bonvehí & Coll, 1997). Acetic acid is the most abundant volatile compound in cocoa beans at the end of fermentation and imputes a sour flavour (Rodriguez-Campos et al., 2012). Other short-chain carboxylic acids such as propanoic, isovaleric and isobutyric acids were found to increase over the course of fermentation. They are secreted by *Bacillus* spp. and result in undesirably pungent, rancid and sweaty flavour notes. Over-fermentation generally leads to the production of these compounds. Volatile acids are usually depleted at the end of roasting and conching (Rodriguez-Campos et al., 2012), except for acetic acid which persists to a moderate extent and contributes to the flavour of chocolate. Acetic acid is also responsible for pH modulation within the bean, leading to other morphological and biochemical changes within the bean, especially Maillard reactions that require low pH for the formation of pyrazines (Afoakwa et al., 2008).

1.7.6.4. Pyrazines

Pyrazines are heterocyclic compounds, formed at high temperatures during roasting, and are the key compounds responsible for cocoa aroma and flavour. Cocoa pyrazines can exist in a plethora of substituted forms, generally referred to as alkylpyrazines (Bonvehí, 2005). They are mainly formed from Strecker degradation in Maillard reactions (Mottram, 2007), but are also reported to be metabolites of bacteria, such as TMP production of *Bacillus subtilis* (Aprotosoaie, Luca, and Miron 2016; Schwan and Wheals 2004). The most direct route for their formation is through condensation reactions of α-aminoketones. Van Lancker et al. (2012) observed that, along with free amino acids, dipeptides and tripeptides also have the propensity to form pyrazines, and that the N-terminal amino acid determined the structure of the resulting pyrazine. This also highlights the role of proteolysis in the formation of pyrazines and how differing enzyme activity can lead to varying profiles of pyrazine compounds or their absence. Their results also revealed that free amino acids were more likely to yield unsubstituted pyrazine, and di- and tri-peptides yielded more substituted pyrazines.

The temperature and duration of the reactions are critical factors that influence the production of pyrazines (Aprotosoaie et al., 2016). TMP is easily formed during roasting and soon reaches a maximum level, whereas TrMP rises in a more gradual fashion. This enables a roasting index to be established between the ratio of TMP/TrMP and the perceptible level of cocoa aroma. Normal roasts are considered to have TMP/TrMP ratios of between 1.5 and 2.5 whereas over-roasted cocoa have ratios below 1.0 (Ziegleder 2009).

1.7.6.5. Other compounds

Aside from the main volatile classes of compounds discussed so far, there are other classes of compounds that are found in cocoa. These include phenols, thiazoles, oxazoles, pyrrole derivatives, pyridines, furans, furanones, pyrans, pyrenes, terpenoids, lactones, amines, amides, nitriles, purines and many others that have yet to be identified (Aprotosoaie et al., 2016; Bonvehí, 2005). With methods of gas chromatography becoming more and more elaborate, so too does the resolution and identification of volatile compounds in cocoa.

2. Research Goals

2.1. State of the Art

Despite the plethora of studies conducted on cocoa bean fermentation, there is a still a great deal of mystery surrounding the change in the cocoa bean metabolome during the process. The biochemistry of raw cocoa and the compounds that contribute to flavour and aroma have already been established. However, compelling links between the two events are still under investigation. This is mainly to do with the fact that the reactions that underlie the formation of flavour and aroma compounds hinge on a copious number of factors. Scientific attempts to characterize these reactions and their mechanisms have been thwarted by issues such as elusive bean hybrid information, lack of equipment on plantations, weather conditions, negligence and inaccuracy of sample handling and shipment, and treatment variations in post-harvest processes. Therefore the majority of studies conducted involve spontaneous trials of fermentation which relate the observations to the specific context in which the data have been gathered. The outcome of this is a superfluous amount of context-specific data (i.e. country or a type of bean) and therefore cannot be accurately compared with each other.

Furthermore, spontaneous fermentation, as the name suggests, is generally unpredictable in nature. This is also a great hindrance to reproducing results and achieving control over fermentation trials in order to guide the process in a desired direction. Controlled fermentation trials are able to harness slightly more control but are still subject to natural events such as spoilage, bad weather, *etc*. This makes it difficult for sourcing well-fermented beans. Chocolate manufacturers are faced with a number of problems when it comes to standardizing methods of fermentation and other post-harvest processes. As a result of such an unreliable process, a lot of money is wasted in shipping dried and fermented beans thousands of miles over to factories, only to discover that the beans are completely unusable.

The issue of unreliability and inconsistency also brings to light another problem in the industry. There is a lack of dependable markers through which one is able to identify a good batch of fermented beans. To date, there are two main methods of identifying well-fermented beans: the fermentation index and the cut-test (Kongor et al., 2013). In the cut-test, a sizeable portion of fermented and dried beans (typically around 100-300) are cut open lengthwise using a sharp knife or a special device. The beans are characterized according to their colour: “slaty”

denoting moulded or partially fermented beans; “fully purple” denoting unfermented beans; “fully brown” denoting well-fermented beans (Guehi et al., 2007). The proportion of each of these types of beans within the full set indicates the scale of fermentation. Determination of the fermentation index is done spectroscopically, using the ratio of the absorbencies of methanol extracts of cocoa at 460 nm and 530 nm. The advantage of these analyses is that they are fairly simple methods and can be employed directly on plantations. Both these analyses, however, can be very subjective and do not give a clear indication of what took place, biochemically, within the bean, including how much of each microbial metabolite has diffused into the bean. Determining the level of acidity within the bean, for example, can be essential in deciding how to further process the bean material.

2.2. General Assumptions

The research described in this thesis is based on the following assumptions:

- The vast majority of reactions that trigger the formation of flavour and aroma precursors in chocolate occurs during the fermentation process of cocoa beans.
- As long as the testa is intact, there is almost no microbial presence within the cotyledon of the bean.
- The genetics of each bean hybrid are significantly different and therefore result in different reaction profiles.

2.3. Hypothesis Statements

Based on the above assumptions, the following hypothesis statements were put forth. The aim of the research was to investigate the validity of these hypotheses.

- Statement 1: Of all the microbial activity during fermentation, the diffusion of the secreted metabolites of ethanol, acetic acid and lactic acid into the bean, in combination with the generated temperatures, have the greatest effects on the profiles of proteins, peptides and polyphenols.
- Statement 2: The dynamics of proteins, peptides and polyphenols of a commercial-scale spontaneous fermentation can be replicated using a lab-scale fermentation under controlled conditions.

2.4. Objectives and Methodology

In order to investigate the hypothesis statements elaborated in section 2.3, two systems of fermentation were developed. Artificial Fermentation is a procedure which involved the incubation of de-pulped beans in solutions containing the main microbial metabolites that beans would be exposed to during fermentation: ethanol, acetic acid and lactic acid. The entire system is maintained under sterile conditions in order to exclude any microbial influence upon bean biochemistry. Forced Fermentation is a "greenhouse" method of fermentation where relatively miniature heaps of beans (c.a. 300 g) were incubated under controlled environmental conditions involving temperature regimes, starter cultures, microbial supplements and air flow control. The purpose of this was to encourage microbial growth and guide the fermentation process.

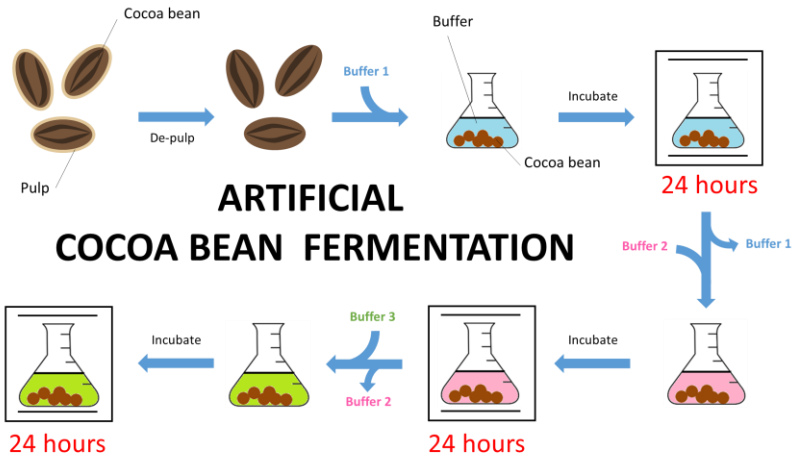
These two systems were exploited in the following manuscripts of sections 3-5 in order to:

1. gain better control over the fermentation process.
2. re-create fermentation in modern and well-equipped lab settings, far removed from the places of harvest.
3. study the events that lead to well- and poor-fermented batches of beans in order to find markers for good and bad fermentation trials.
4. discover and develop methods that can lead to refined and consistent trials of fermentation.

3. Results – Part I

Aseptic artificial fermentation of cocoa beans can be fashioned to replicate the peptide profile of commercial cocoa bean fermentation

Warren A. John, Neha Kumari, Nina L. Böttcher, Sergio Grimbs, Gino Vrancken, Roy N. D'Souza, Nikolai Kuhnert and Matthias S. Ullrich



Abstract:

The fermentation of cocoa beans is essential for the generation of flavour precursors that are required later on to form flavour components of chocolate. From the many different precursors that are generated, oligopeptides and free amino acids comprise a significant proportion as some of them form Maillard reaction products during the roasting process. Therefore, the diversity of peptides is an important contributing factor to the quality of a fermentation which is in turn controlled by proteolytic activity within the cocoa bean, and is driven by changes in the presence of fermentation by-products as a result of microbial activity outside the bean. Being able to control proteolytic activity within the bean using only the presence of fermentation by-products would prove a valuable tool in the study of these proteases and the processing of cocoa storage proteins. Thus, this tool would help elucidate key mechanisms that generate the components responsible for flavour. In this study, we describe an artificial fermentation system, free from microbial activity, which is able to replicate proteolytic degradation of protein as well as to generate similar peptide fragments as seen during a commercial fermentation.

3.1. Introduction

Farming of the cocoa tree (*Theobroma cacao*) has been carried out for over 3,500 years, having its roots in Mesoamerican societies and now having spread to all tropical regions of the world with Ivory Coast and Ghana contributing the most to cocoa cultivation (Dreiss & Greenhill, 2008; Ismail et al., 2009). Intense cocoa farming has resulted in vast numbers of cocoa hybrids designed for resistances to unfavourable weather conditions and pathogenicity as well as for chocolate flavour profiles (Motamayor et al., 2003; Wood & Lass, 2001) such that each farm or institution can have its unique cocoa to sell on the market.

The chocolate production process involves many stages that are crucial for the development of a pleasant flavour profile, among which are selection of the hybrid, bean fermentation and bean roasting (Afoakwa et al., 2008). Fermentation is a key process as it is involved in the development of flavour precursors from the cocoa bean. Over the course of the fermentation and on the exterior, a repertoire of yeasts, lactic acid bacteria, and acetic acid bacteria commandeer the process in respective succession (Pereira et al., 2012). The pulp, being rich in nutrients and low in pH, in combination with decreasing oxygen concentration, allow yeasts to thrive. The yeasts degrade the pulp using their pectinolytic enzymes, metabolizing sugar and citric acid and as a result, produce ethanol and raise pH (Papalexandratou and De Vuyst 2011; Rombouts 1953; Schwan and Wheals 2004). It is also known that yeasts contribute to flavour by the generation of volatile substances in the first stages of the fermentation (Schwan and Wheals 2004; Steensels and Verstrepen 2014). Anaerobic conditions and high CO_2 concentrations favour the growth of lactic acid bacteria, albeit for a brief period. After the beans are typically turned on the second day of fermentation, oxygen is reintroduced and populations of acetic acid bacteria take over, producing acetic acid from the remaining ethanol in an exothermic reaction thereby raising the bean pile's temperature (Camu et al., 2007). These fermentation products then diffuse into the bean over the course of the fermentation, causing the death of the embryo and a cascade of various enzymatic reactions and structural changes inside the cotyledons (Biehl et al., 1982; Lefeber, Gobert, et al., 2011)

The various fermentation-driven changes within the cocoa bean also involve the degradation of proteins (up to 20% of bean mass), the majority of which consists of 52% albumin (a 21-kDa storage protein) and 43% vicilin-like globular storage protein, hereon referred

to simply as vicilin (Voigt and Biehl 1993). The vicilin protein is preferred over albumin for breakdown by the native activities of a speculated aspartic endoprotease, aminopeptidase and a carboxypeptidase (Hansen, Del Olmo, and Burri 1998; Laloi, Mccarthy, and Morandi 2002; Lerceteau et al. 1999; Voigt et al. 1994). The activities of those enzymes are stimulated during the conditions generated during fermentation. It has been understood from the fermentation-like incubations carried out by Biehl *et al.* that as acetic acid diffuses into the bean during fermentation, there is a profound impact on subcellular structure within the cotyledons, affecting protein and lipid bodies and thus, also impacting proteolysis within the cotyledons (Biehl & Passern, 1982; Biehl et al., 1982). The endoprotease that is active in protein cleavage into peptides is known to have a pH optimum of 3.0, spurred on by the diffusion of acetic acid into the bean (Amin, Jinap, Jamilah, Harikrisna, & Biehl, 2002). These peptides and free amino acids can go on to react with sugars and polyphenols in Maillard reactions in the roasting process (Bonvehí & Coll, 2002; Mohr, Landschreiber, & Severin, 1976; Voigt et al., 2016; Ziegleder, 1991), thereby making them key precursors required for chocolate flavour.

The aim of the current study was to model cocoa fermentation in such a way that internal processes in the bean would be triggered solely by the presence of fermentation by-products in a succession of buffer incubations, without any microbial contribution. Changes in the environmental concentrations of these fermentation products were tracked and degradation products of proteins were then analysed with protein quantification assays and high-resolution HPLC-MS. The results were compared to those of a commercial spontaneous fermentation. Strikingly, we found that using an artificial system, we could replicate proteolysis observed in the commercial fermentation with similar peptide patterns being generated and that the concentrations of fermentation products significantly influenced peptide formation.

3.2. Materials and Methods

3.2.1. Chemicals and reagents

Tris-HCl (Pufferan®, ≥ 99.5%), yeast extract (bacteriology grade), sodium dodecyl sulphate (SDS, ≥ 99.5 % electrophoresis grade), HPLC-grade water (Rotisolv®), glycerol (86%), acrylamide (Rotiphorese® Gel 30: 37,5:1), acetonitrile (ACN, Rotisolv® HPLC ultra gradient grade), citric acid (≥ 99.5%), ethanol (≥ 96%, denatured with 1% MEK), lactic acid (90% synthetic grade), fructose (≥ 99.5% for biochemistry) and D-glucose monohydrate (≥ 99.5%)

were purchased from Carl Roth (Karlsruhe, Germany). Dithiothreitol (DTT, Biochemica), acetone (100%, Biochemica), TEMED, ammonium persulphate (analytical grade), sulphuric acid (90 – 91%, pure), acetic acid (100%, analytical grade) and methanol (100%, analytical grade) were purchased from Applichem (Darmstadt, Germany). Isopropanol (100%, analytical grade) and bromophenol blue sodium salt (research grade) were purchased from Serva Electrophoresis (Heidelberg, Germany). Formic acid (≥ 98% for mass spectrometry) was purchased from Sigma-Aldrich Chemie (Steinheim, Germany). 2D Quant Kit was purchased from GE Healthcare Biosciences (New Jersey, USA). Coomassie® Brilliant blue G-250 (electrophoresis grade) was purchased from Merck (Darmstadt, Germany).

3.2.2. Artificial Fermentation

All equipment used was either wiped with 70% ethanol or autoclaved in order to exclude any microbial contribution to the system. Cocoa beans of the hybrid C308XC5 from Ivory Coast were removed from the pods upon shipment and immediately de-pulped under high pressure water. Beans were rinsed in 70% ethanol for one minute to get rid of any remaining microbes on the surfaces of the beans. 100 g portions of beans were placed inside 500-mL bottles along with 90 mL buffer and the headspace was flushed with either nitrogen or compressed air through a 0.2-µm filter depending on the day of incubation. The bottles were placed in incubators and the buffers exchanged every 24 hours under sterile conditions inside a laminar flow work bench. Two trials of artificial fermentation were run along with a control involving low concentrations of the buffer constituents. **Table 3.1** shows an overview of the schemes of both trials. Sampling was done in triplicate after every 24 hours where three bottles representing three biological replicates were removed and the beans and buffers were collected and stored at -20°C. Shells of the beans were removed and the beans were ground to powder and stored at -20°C. The remaining bottles had their buffers exchanged, were flushed once again and placed at the respective incubation temperature. Sterility of the system was confirmed by plating on LB medium, acetic acid medium, glucose-yeast extract-carbonate (GYC medium) and malt extract agar to check specifically for the growth of yeasts, lactic acid bacteria, and acetic acid bacteria as well as any other microorganisms that could potentially thrive in these buffers.

Table 3.1: Buffers and their components present in incubations of artificial fermentation along with temperature conditions.
Beans were incubated in a succession of buffers at a constant temperature for 24-hour periods for each fermentation trial. *pH values that were controlled for before incubation.

Fermentation	Buffer	Citric acid [g/L]	Ethanol [g/L]	Acetic Acid [g/L]	Lactic Acid [g/L]	Temperature [°C]	pH
Trial 1	A1	15	80	0	0	37	2.3
	A2	10	80	15	30	45	2.0
	A3	10	40	15	30	30	2.0
Control with low concentrations	B1	1.2	0.79	0.17	0.31	37	2.7
	B2	0.9	1.18	0.40	1.2	45	2.6
	B3	0.1	1.58	0.53	1.7	30	2.7
	B4	0.1	0.79	0.53	1.7	30	2.7
	B5	0.1	0.40	0.53	1.7	30	2.7
Trial 2	C1	15	22.5	15	0	35	2.0*
	C2	10	22.5	15	0	35	3.0*
	C3	5	45	10	0	35	4.0*
	C4	0	45	0	0	35	5.0*
	C5	0	90	0	0	35	6.0*

3.2.3. Commercial Fermentation

Pods of the hybrid C308XC5 from Ivory Coast were harvested in December 2013 in the estate of Champ Semencier de Mafere where the average age of the trees was 30 years. The beans were removed from the pods within 12 hours of harvesting with placenta, broken beans and the pod husks being discarded. Beans were placed into a pile of 100 kg, and fermentation was carried out for a period of seven days. Beans in the pile were turned after 48 and 96 hours. Sampling was done from the middle of the pile in portions of 2 kg and were immediately stored at -20°C and shipped to Jacobs University Bremen, Germany on dry ice within three days. Shells of the beans were removed and the beans were ground to powder and stored at -20°C.

3.2.4. Protein Extraction

1 mL of protein extraction buffer, containing 100 mM Tris-HCl, 1% SDS and 1% DTT, was added to 250 mg of cocoa bean powder in a 2-mL centrifuge tube, vortexed briefly, and kept

on a gently turning head-over-head rotator at 4°C for one hour. The mix was vortexed again and centrifuged at 16,000 *g* for 20 min at 4°C. The resulting aqueous fraction was carefully pipetted out into a pre-equilibrated Vivaspin 2 column (Sartorius, Goettingen, Germany) containing a hydrosart membrane with a 10 kDa molecular weight cut-off carefully avoiding the transfer of any fat or bean material. The Vivaspin 2 column was pre-equilibrated by spinning down 1 mL HPLC-grade water at 7,500 *g* at 4°C for 5 min in a fixed-angle-rotor Eppendorf 5810R centrifuge (Eppendorf, Hamburg, Germany). The protein extract was applied to the column and afterwards spun down under identical conditions. The filtrate fraction at the bottom of the column, hereon referred to as the peptide fraction and containing molecules less than 10 kDa, was transferred into a 1.5-mL centrifuge tube and stored at -20°C. The residue fraction still remaining on top of the filter, hereon referred to as the crude protein fraction (CPF) and containing molecules larger than 10 kDa, was transferred into a 1.5-mL centrifuge tube for storage. An aliquot of the CPF was transferred into a new centrifuge tube and mixed with 100% acetone in a ratio one part CPF to four parts acetone. The solution was vortexed briefly and then stored overnight at -20°C for protein precipitation. The precipitate was spun down at 16,000 *g* for 10 min at 4°C, the supernatant removed, and the tubes placed in an Eppendorf 5301 vacuum centrifuge to remove any residual acetone. The protein pellet was later re-suspended in 50 mM Tris-HCl (pH 8.0), hereon referred to as the purified protein fraction (PPF), and stored at -20°C for further processing.

3.2.5. Protein Quantification

Quantification of purified protein fractions was done using the 2D-Quant Kit according to its protocol where 30 µL of PPF in 50 mM Tris-HCl (pH 8.0) were used. The samples were quantified in three technical replicates. A standard calibration curve was generated using bovine serum albumin (BSA) in the range of 0 – 2,000 µg/mL. Absorbance was measured at 480 nm using an Ultrospec 2100 pro UV/visible spectrophotometer (GE Healthcare Biosciences).

3.2.6. SDS Polyacrylamide Gel Electrophoresis

25 µL of PPF was mixed with 5 µL loading dye consisting of 0.35 M Tris-HCl (pH 6.8), 10.3% SDS, 37% glycerol, 600 mM DTT, and 1 mg/mL bromophenol blue. The sample was placed at 95°C for 5 min before it was loaded to a 12.5% polyacrylamide gel. Gels were run

using the Biorad Mini-Protean Tetra Cell gel electrophoresis system (Biorad Laboratories, CA, USA) in SDS running buffer (0.3% Tris base, 0.19 M glycine and 0.1% SDS) for 30 min at 110 V and for approximately 75 min at 130 V.

After electrophoresis, the gels were washed under running deionized water and placed in Coomassie® G-250 staining solution (2.5 g/L Coomassie® G-250, 10% acetic acid and 45% methanol) for 20 min on gentle agitation. The gels were then de-stained in four consecutive washes, the first and second washes of 20 min each, the third overnight and the fourth of 20 min, in de-staining solution (5% isopropanol and 10% acetic acid).

3.2.7. Peptide analysis with LC-MS

The Vivaspin 2 column peptide fractions were concentrated in a vacuum centrifuge at 30°C and the solid mass was re-suspended in 200 µL of 0.1% formic acid. The samples were applied to an AdvanceBio Peptide Mapping, 2.1 x 250 mm, 2.7 µm, LC column equipped with an AdvanceBio Peptide Mapping Fast Guard, size 2.1 x 5 mm guard column (Agilent Technologies, CA, USA). A sample injection volume of 2 µL was used and the sample was eluted with the ratios of solvent A (0.005% formic acid in water) and solvent B (100% ACN) from 8% to 90% solvent B in solvent A within 30 min at a flow rate of 250 µL/min. Tandem MS was performed with an Impact™ HD UHR-QToF with ESI-MS/MS (Bruker Daltonik, Bremen, Germany). The nebulizer pressure was kept at 2 bars with a dry gas flow rate of 8 L/min and a dry gas temperature of 200°C. The ToF was calibrated with 0.1 M sodium formate. The instrument was run with positive MSn mode with m/z detection range of 160 - 3000. A blank of HPLC grade water was run after every 5 samples to check for contamination from the previous samples. The spectra obtained were then run on the Mascot search engine (Matrix Science, Boston, USA). Parameters were set to search the NCBInr database for “Other green plants” with variable modifications of methionine oxidation. The mass tolerance MS was set to 100 ppm and MS/MS tolerance to 0.1 Da.

3.2.8. Fermentation product analysis with HPLC-RI

Buffer and sweating samples were spun down and filtered through a 0.45-µm pore size filter. 20 µL were passed through a Waters Breeze 2 HPLC system, comprising the following components: Waters 2414 refractive index detector, Waters 2489 ultraviolet detector, Waters

1525 pump system and Waters 2707 injector system, was running the Breeze 2 software v6.2 (Waters Corporation, MA, USA). An Aminex HPX-87H column (Biorad Laboratories) was used containing 8% cross-linked resin and a pH range of 1-3. Quantification was performed for citric acid, ethanol, acetic acid, lactic acid, glucose and fructose. An eluant of 5 mM sulphuric acid was used with a flow rate of 0.6 mL/min and a column temperature of 45°C. The readings were recorded and analysed using the Breeze™ 2 software.

3.3. Results

3.3.1. Influence of fermentation by-products on cocoa bean proteins

De-pulped cocoa beans of the hybrid C308XC5 were placed in a series of buffers consisting of citric acid, ethanol, acetic acid and lactic acid to mimic the environmental conditions outside the beans as they would occur during a fermentation (Trial 1). The concentrations of the buffer constituents were based on those measured in commercial fermentations (Camu et al., 2007), although increased to accentuate the effects of the buffer constituents on the beans (**Table 3.1**). **Figure 3.1** shows the changes in the levels of these fermentation products in each set of buffer incubation measured before and after 24-hour incubation with the beans. Of all the four buffer constituents, acetic acid and lactic acid seem to be most assimilated within the beans, since even after successive buffer incubations there is a decrease in the concentrations of these acids within the buffer.

Upon addition of acetic and lactic acids, and an increase in temperature to 45°C (**Table 3.1**), there is an almost complete breakdown of proteins within 24 hours (**Figure 3.2A**). Such rapid protein degradation has never been observed before in a commercial fermentation as proteins generally degrade gradually over the course of fermentation (**Figure 3.2B**). A control trial with much lower concentrations of buffer constituents but with similar temperature profile did not show as much protein degradation (**Figure 3.2C**) implying that one or more of the buffer constituents was indeed responsible for a rapid breakdown of proteins.

3.3.2. Peptide repertoires of degraded vicilin protein show resemblance

Qualitative peptide fragment analyses for the vicilin storage protein (**Figure 3.3**) showed a significant increase in the peptide repertoire after incubation for 48 hours, i.e., buffers 1 and 2 consecutively, which is accompanied by massive protein degradation occurring at the same point

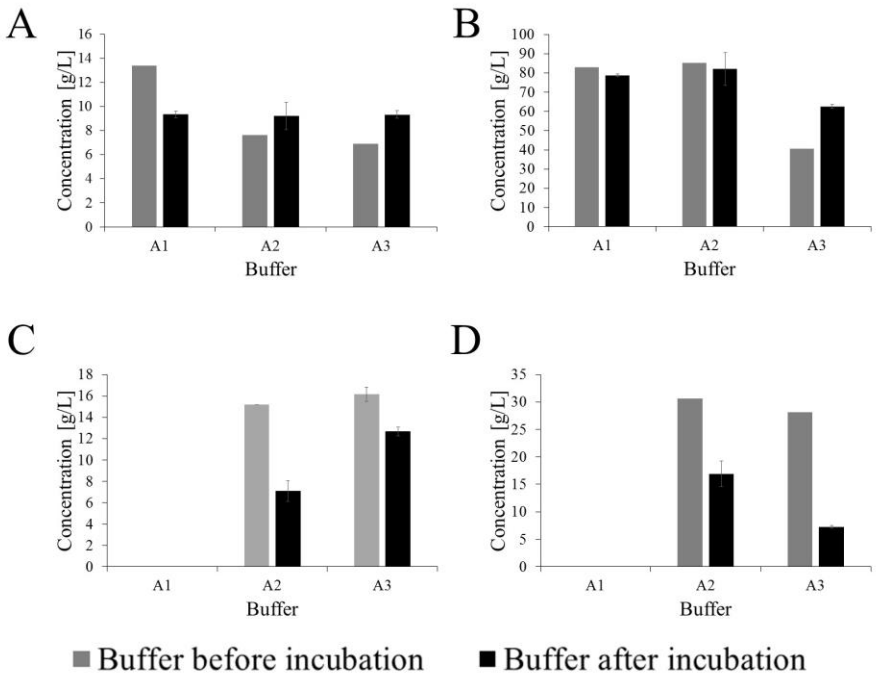


Figure 3.1: Fermentation product quantification of buffers during Trial 1 of artificial fermentation.
Buffers were analysed using HPLC with an RI detector for concentrations before the incubation (grey bars) and after 24 hours of incubation (black bars). Beans were incubated in a succession of buffers where beans were first incubated in buffer A1, followed by buffer A2, *etc* (Table 1). Quantification was done for citric acid (A), ethanol (B), acetic acid (C), and lactic acid (D). Standard deviation (error bars) is given for three replicates.

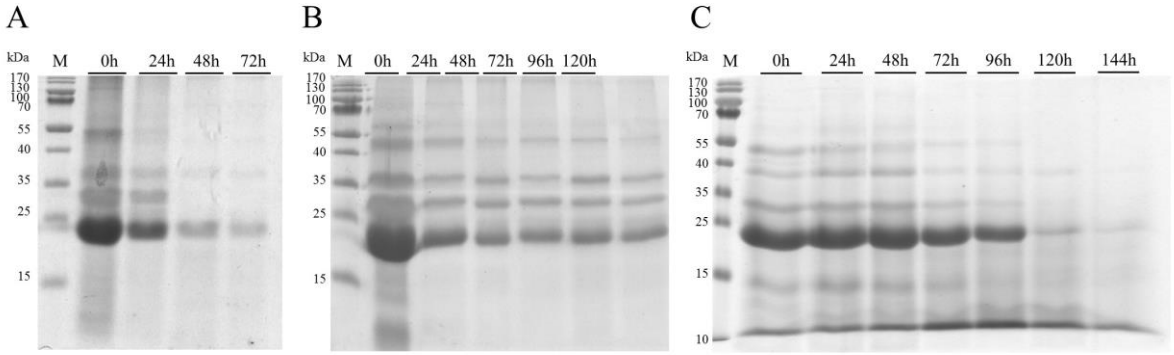


Figure 3.2: Protein profile visualization by SDS-PAGE analysis.
Protein fractions were run on 12.5% acrylamide gels stained with Coomassie G-250 for Trial 1 of artificial fermentation (A), commercial fermentation (B), and a control with low concentrations of buffer constituents (C).

of the artificial fermentation (**Figure 3.2A**). As it has been previously established that breakdown products from vicilin lead to the generation of more chocolate aroma (Voigt et al. 1994), peptides found during a 120-hour commercial fermentation were compared side-by-side

with peptides generated during the artificial fermentation (**Figure 3.3**). This comparison showed that artificial fermentation resulted in a greater diversity of the peptide repertoire. This may suggest that the artificial fermentation featured a greater degree of endoprotease activity and lower exopeptidase activity throughout. This would also validate data supporting an active exopeptidase with a working pH of about 5.5 (Voigt et al. 1994) which could be inhibited by the low pH incited by the acids present in the buffers used herein, and an active aspartic proteinase which has a working pH of 3.5 (Hansen et al., 1998).

As each peptide could later on form flavour or aroma compounds through Maillard reactions, all detected peptides were examined in terms of their lengths and sequences. Upon comparison of the vicilin peptides, it was found that 75 to 78% of the peptides observed after each time point of the commercial fermentation were present after 72 hours of incubation in the

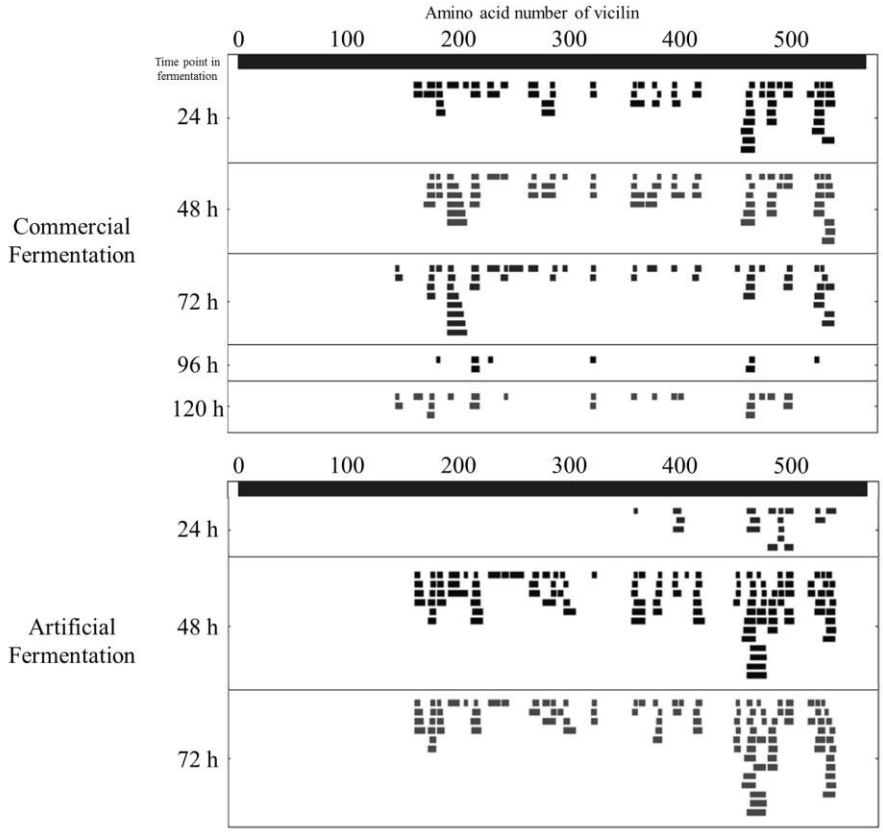


Figure 3.3: Peptide repertoire of *Theobroma cacao* vicilin degradation during fermentation resolved by UHPLC-MS/MS.
Fractions of extracted peptides with MW of less than 10 kDa separated by LC-MS and analysed via Mascot search. Peptides identified by the Mascot search (small black bars) are placed against the amino acid sequence of vicilin represented by the long black bar on top for 24, 48, 96, and 120 hours of a commercial fermentation, and for Trial 1 of artificial fermentation after 24, 48, and 72 hours of incubation.

artificial fermentation. This result demonstrated that native cocoa bean enzyme activity suffices in producing a vast proportion of the peptides that may eventually form flavour precursors. With the total number of peptides having decreased towards the latter stage of the commercial fermentation, there is a high likelihood of these peptides being flavour or aroma precursors.

3.3.3. Terminal amino acid residue analysis highlights specific enzyme activity

Scrutiny of the termini of the peptides revealed that their N- and C-terminal amino acid residues did not vary dramatically at the different time points for either the commercial fermentation or Trial 1 of the artificial fermentation, respectively. Consequently, the termini found for Trial 1 could be compared with terminal amino acyl residues of the peptides observed during a commercial fermentation (**Figure 3.4**). When comparing 163 peptides from the pool of the commercial fermentation with 220 peptides from the pool of the artificial fermentation, it was observed that the N-termini had an exceptional preference for aspartate in the artificial fermentation. However, the peptides derived from the commercial fermentation did not seem to indicate a preference for a particular terminal residue (**Figure 3.4A**). This result prompted us to speculate on activity of aspartic endoprotease(s) (D'Hondt et al., 1993; Marseglia et al., 2014) in Trial 1. Additionally, a preference for aspartate on the N-terminus of peptides in the artificial fermentation could suggest a higher degree of aminopeptidase activity in the commercial fermentation which in turn would result in the cleavage of aspartate residues at the P1 position of the peptides (Hansen et al., 1998).

Meanwhile, the C-termini of peptides had the highest preference for phenylalanine in both fermentations and second highest preference for tyrosine (**Figure 3.4B**) suggesting that the conditions were favourable in both cases for equally high carboxypeptidase activity as this enzyme group is known to be inhibited by phenylalanine (Bytof et al., 1995; Jacobsen & Bartlett, 1981). Both, phenylalanine and tyrosine are known to have an inhibitory effect on carboxypeptidase A (Auld & Vallee, 1970; Steitz, Ludwig, Quiocho, & Lipscomb, 1967). Bytof *et al* had argued that the carboxypeptidase activity in cocoa is similar to that of the carboxypeptidase A activity in porcine pancreas (Bytof et al., 1995).

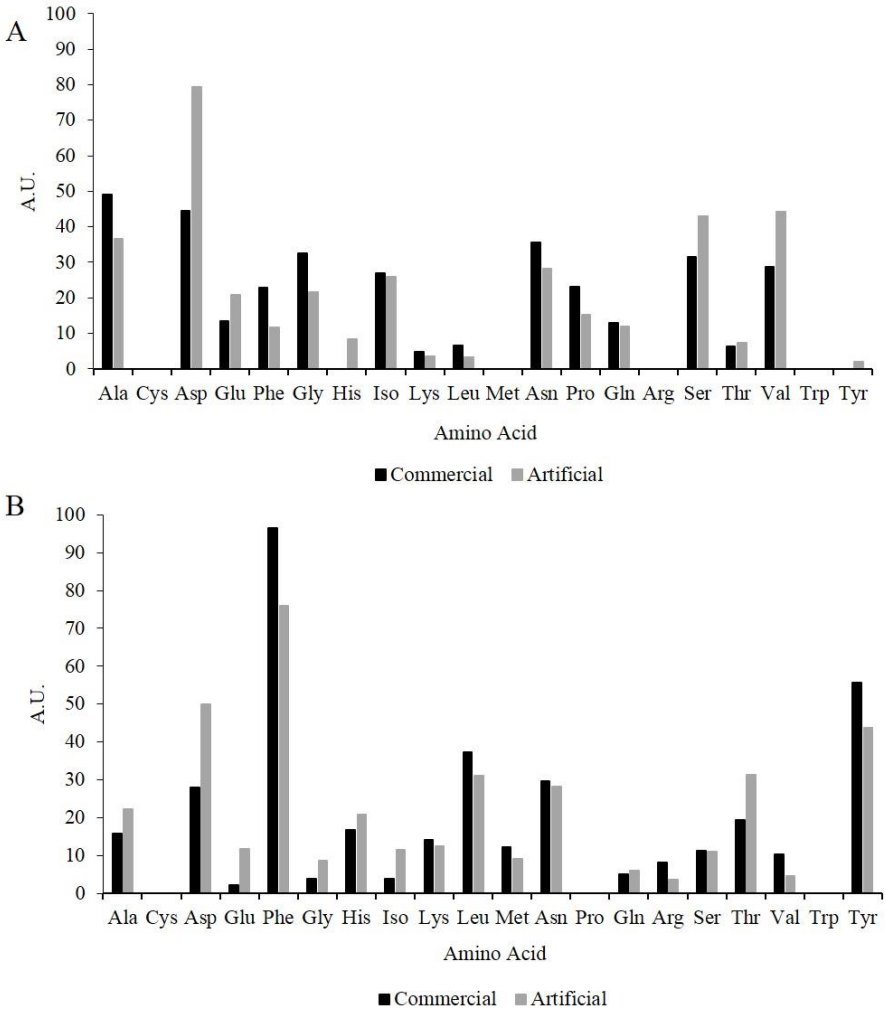


Figure 3.4: Terminal amino acyl residue analysis of peptides of vicilin resolved by UHPLC-MS/MS for a commercial fermentation and Trial 1 of artificial fermentation.
N-termini (A) and C-termini (B) of the characterized peptides normalized to the number of peptides per sample and the occurrence of the amino acid in the sequence of vicilin comparing all observed peptides of a commercial fermentation and Trial 1 of artificial fermentation. A.U: arbitrary units.

3.3.4. Altered buffer composition impacts protein degradation profile

Another artificial fermentation was set up (Trial 2) allowing the cocoa beans to be incubated at subsequently increasing pH conditions via decreasing citric and acetic acid concentrations in the added buffers. This pH increase is usually experienced on the bean's exterior in the course of a commercial fermentation, however, the components within the buffer were varied in a manner unlike in a normal fermentation (**Table 3.1**). The temperature was kept steady at 35°C in order to dissect effects caused by changing acidity from temperature effects.

The most rapid protein degradation occurred during the initial 24 hours of incubation in the first buffer with low pH and 15 g/L of acetic acid (**Figure 3.5**). This is most likely due to endoprotease activities stimulated by a dramatic drop in internal pH of the beans. An analysis of the buffer pH before and after bean incubation showed that the cocoa beans exhibited a buffering effect where the pH after bean incubation was brought to a range of between 3 and 4 (**Figure 3.6A**). This would most likely be a result of acetic acid having diffused into the bean in the first

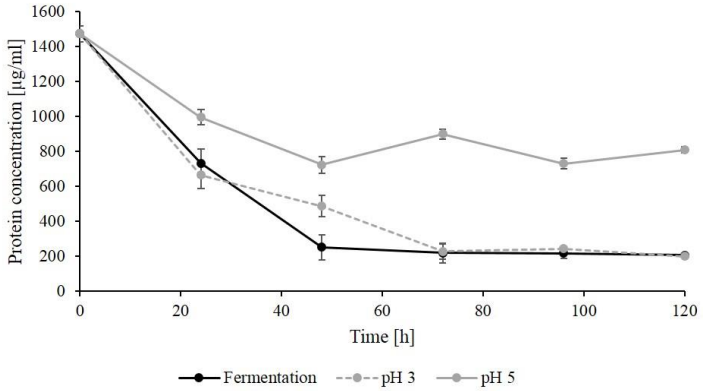


Figure 3.5: Protein concentration in cocoa beans during Trial 2 of artificial fermentation.
Protein fractions above 10 kDa were quantified. Buffers were exchanged in Trial 2 fermentation according to Table 1. Controls of incubations in one buffer of pH 3 and one of pH 5 were maintained for all five days with no buffer exchange. Error bars represent standard deviation for three biological replicates each having three technical replicates.

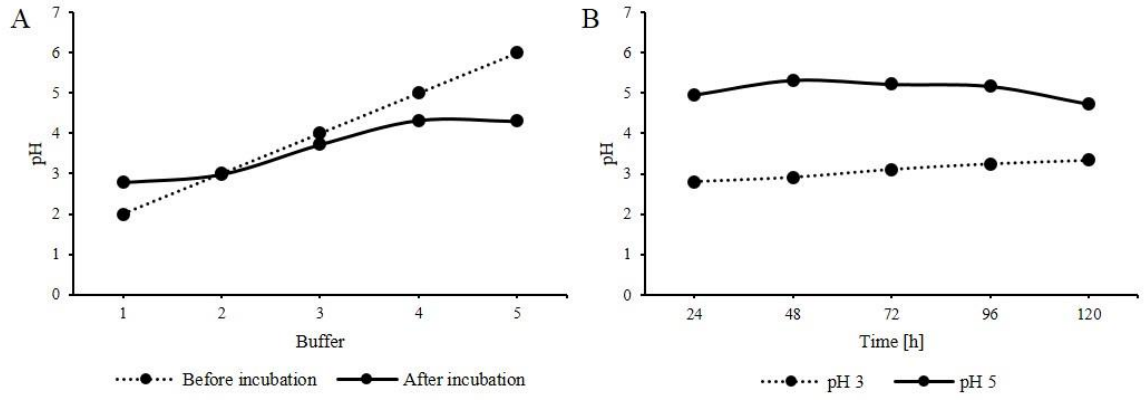


Figure 3.6: Measurement of the pH of buffers in Trial 2 of artificial fermentation.
The pH was measured before incubation and after incubation for the Trial 2 artificial fermentation (A) and only after incubation for the control incubations at pH 3 and pH 5 (B). Beans of the Trial 2 fermentation were incubated in a succession of buffers according to Table 1, whereas the buffers of the control runs were not exchanged. The standard deviation (error bars) is given for three biological replicates in A, and for one replicate in B.

two incubations and diffusing out towards the end (**Figure 3.7**).

The corresponding control incubations had cocoa beans incubated at pH 3, with the same buffer constituents as buffer 1 for five days without buffer exchange, and cocoa beans kept at pH 5, with the same buffer constituents as buffer 5 for five days without buffer exchange. Only slight variations in pH were observed over the course of the incubation (**Figure 3.6B**). Importantly, it was observed that the presence of acetic acid and low pH contributed vastly to protein degradation as evidenced by the control maintained at pH 3 showing a very similar decline in protein content as during the fermentation of Trial 2 (**Figure 3.5**). This provided evidence for more endoprotease activities at low pH, and lower activity with high pH as the proteins were not fully degraded.

Initially high proteolytic activity was confirmed also by all five time points of the fermentation of Trial 2, with peptide repertoires of the first three time points showing the greatest diversity (**Figure 3.8**). Comparing the peptides between commercial fermentation and Trial 2 of artificial fermentation, only 25 to 50% of the peptides of each time point commercial fermentation were found after 120 hours of the artificial fermentation. This is a manifest drop in

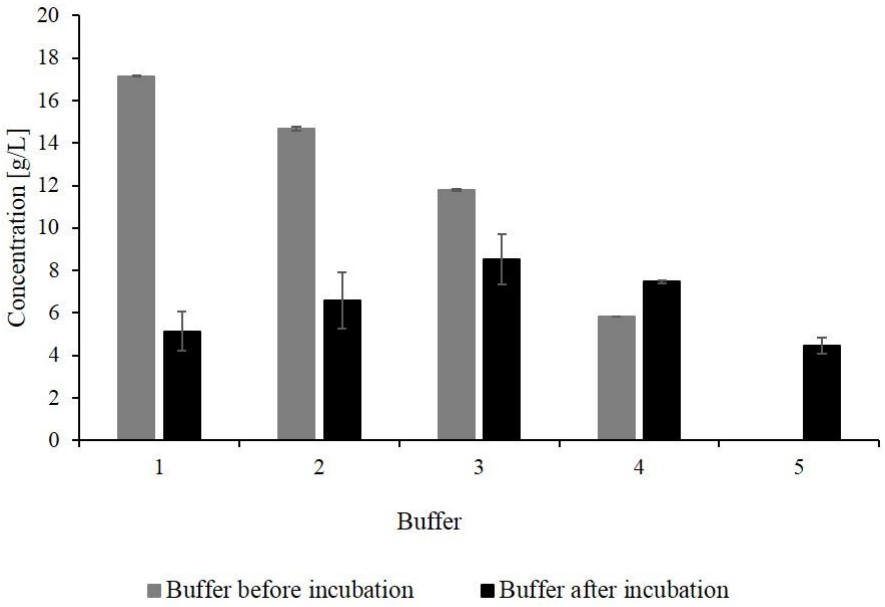


Figure 3.7: Quantification of acetic acid in buffers of Trial 2 of artificial fermentation.
Buffers were analysed using HPLC with an RI detector. Grey bars represent concentration of the compound being analysed before the incubation and black bars represent concentration of the compound being analysed after the 24 hour incubation. Beans were incubated in a succession of buffers where beans were first incubated in buffer 1, followed by buffer 2, *etc*. The standard deviation (error bars) is given for three replicates.

the similarity of the peptides between the fermentation.

As found in Trial 1, a preference for aspartate as the preferred N-terminal amino acyl residue of peptides as well as a significant preference for phenylalanine and tyrosine found at the C-termini of peptides suggested activities of aspartic proteinase(s) and carboxypeptidase(s), respectively, during fermentation of Trial 2 (**Figure 3.9**).

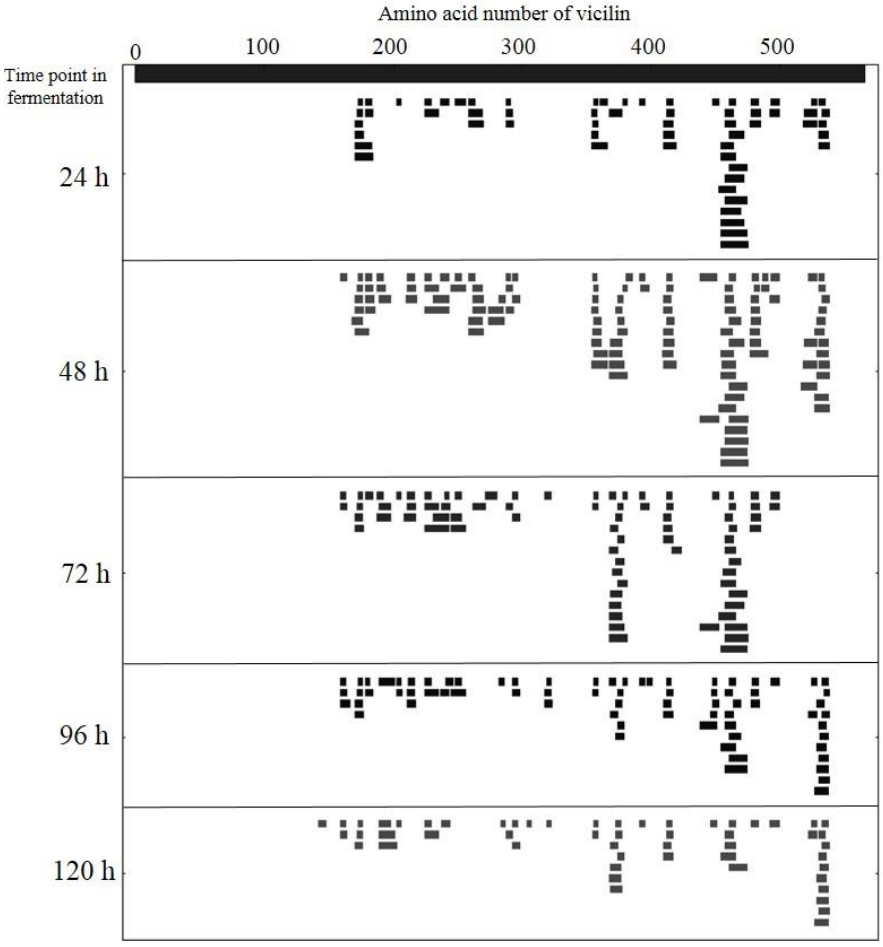


Figure 3.8: Peptide repertoire of vicilin protein resolved by UHPLC-MS/MS for Trial 2 of artificial fermentation.
Fractions of extracted peptides with MW of less than 10 kDa separated by LC-MS and analysed via Mascot search. Peptides identified by the Mascot search (small black bars) are placed against the amino acid sequence of vicilin represented by the long black bar on top for incubation after 24 hours (A) 48 hours (B), 72 hours (C), 96 hours (D), and 120 hours (E).

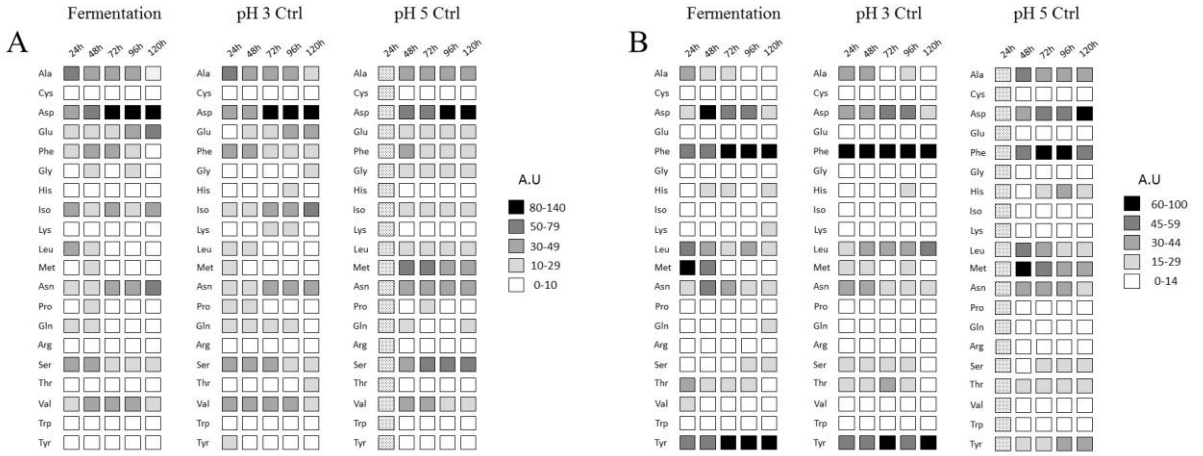


Figure 3.9: Terminal amino acyl residue analysis of vicilin peptides resolved by UHPLC-MS/MS for Trial 2 of artificial fermentation and for control incubations with fixed pH.
N-termini (A) and C-termini (B) of the characterized peptides normalized to the number of peptides per sample and the occurrence of the amino acid in the sequence of vicilin. The Trial 2 fermentation was run according that in Table 1 and a control at pH 3 and pH 5 were maintained. There were no peptides present after the first 24 hours of incubation in the control at pH 5. A.U: arbitrary units.

3.3.5. Cocoa bean endo- and exoproteases exhibit broad working pH ranges

In contrast to previous reports about lower activity of endoprotease(s) at higher pH (Laloi et al. 2002), our peptide fragment analyses revealed that there was persistent endoprotease activity as evidenced by the presence of peptides formed throughout the artificial fermentation of Trial 2 and even at pH 5 at 120 hours of incubation (**Figure 3.8**). The average size of the peptides decreased over the course of both, the fermentation and the control incubation at pH 3, which pointed to greater exopeptidase activity and/or lower endoprotease activity, possibly being a result of the rise in pH (**Figures 3.8** and **3.10**). Surprisingly, there was significant carboxypeptidase activity as evidenced by a high preference for phenylalanine at the peptides' C-termini (**Figure 3.9B**) even though the pH of the control incubation at pH 3 only reached a pH maximum of 3.5 after 120 hours of incubation, which was different from the presumed pH optimum of 5.5 for carboxypeptidase (Bytof et al. 1995; Hansen et al. 1998).

Even in the control incubation at pH 5, peptides were present throughout, however with the majority of them appearing later as compared to the Trial 2 fermentation (**Figure 3.11**). This would suggest that even at more neutral pH, endoprotease activity was high. Again,

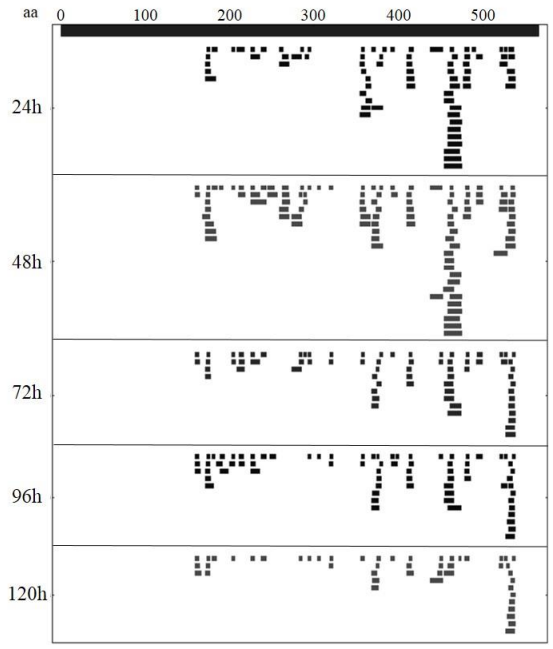


Figure 3.10: Peptide repertoire of vicilin resolved by UHPLC-MS/MS for the control run at pH 3 of Trial 2 of artificial fermentation.
Fractions of extracted peptides with MW of less than 10 kDa separated by LC-MS and analysed via Mascot search for vicilin. Peptide hits (small bars) are placed against the amino acid sequence of vicilin represented by the long black bar on top.

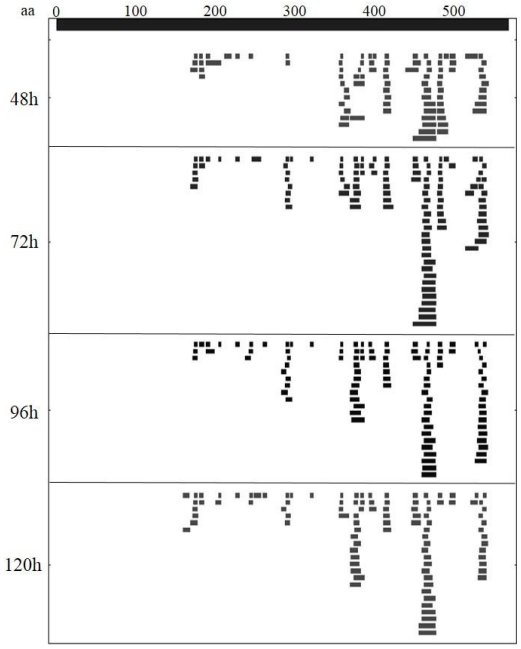


Figure 3.11: Peptide repertoire of vicilin resolved by UHPLC-MS/MS for control run at pH 5 of Trial 2 of artificial fermentation.
Fractions of extracted peptides with MW of less than 10 kDa separated by LC-MS and analysed via Mascot search for vicilin. Peptide hits (small bars) are placed against the amino acid sequence of vicilin represented by the long black bar on top.

endoprotease activity was corroborated by a significantly higher preference for aspartate on the N-termini of the peptides (**Figure 3.9A**) although this preference was less pronounced as compared to the control at pH 3 thus supporting previous reports on the optimal activity of aspartic proteinase at a pH of 3 (Laloi et al. 2002). Since the protein content in the control incubation at pH 5 did not decrease as much as in the Trial 2 fermentation or in the control run at pH 3, respectively, (**Figure 3.5**) and since there was greater peptide diversity present during the latter stages of the control run at pH 5 compared to the Trial 2 fermentation and the control at pH 3, one may conclude that there was still only mediocre exopeptidase activity even at a pH of 5 despite carboxypeptidase having a pH optimum of 5.5. The herein observed reduced preference for phenylalanine and tyrosine as C-terminal amino acyl residues of peptides in the run at pH 5 (**Figure 3.9B**) supported this notion.

As it is known that internal cocoa bean pH does not drop to less than 4.5 during a commercial fermentation (Thompson et al., 2007), it has to be assumed that there is endoprotease

activity even at this pH range. Only once processing of proteins by endoprotease has proceeded, exopeptidase activity would be at its highest due to the presence of digested rather than fully-folded proteins. This would consequently explain why bean pH has to drop to a slightly more acidic pH for complete protein breakdown into both peptides and free amino acids.

Interestingly, the control incubation at pH 5 as well as the Trial 2 fermentation exhibited initially high preference for cleavage after methionine residues (**Figure 3.9B**). This atypical cleavage pattern could be indicative for the activity of additional seed proteases.

3.3.6. Degradation of albumin shows potential for flavour precursor formation

Even though it had been noted previously that vicilin is selectively degraded during cocoa bean fermentation (Voigt and Biehl 1993), diverse peptide fragments derived from the other storage protein, albumin, were observed in the current study (**Figure 3.12**). However, there was noticeably less exopeptidase activity since most of this protein was still present in the form of long peptide fragments. The generation of a substantial diversity of peptides from albumin could indicate that it too might play a role in forming flavour or aroma precursors and should not be disregarded as such too swiftly. This observation is consistent with recent findings of others

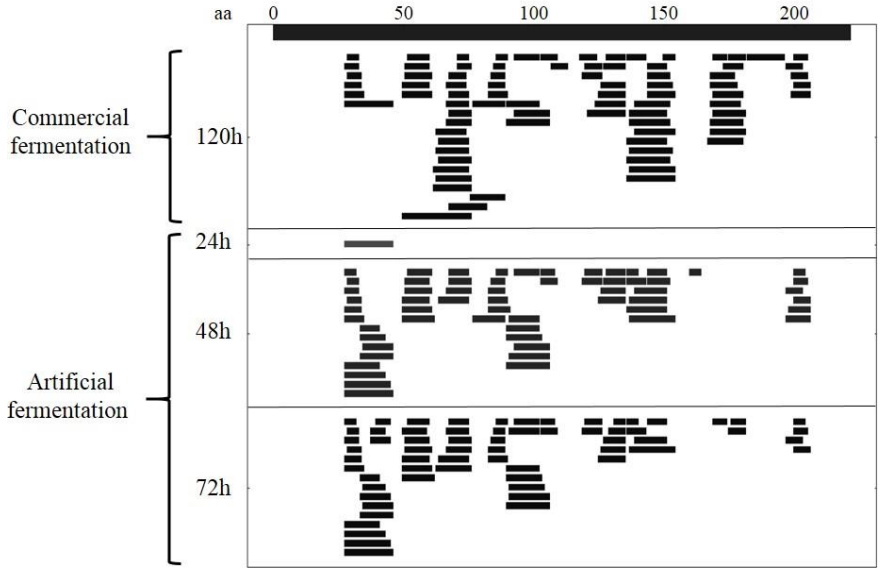


Figure 3.12: Peptide repertoire of albumin resolved by UHPLC-MS/MS for Trial 1 of artificial fermentation. Fractions of extracted peptides with MW of less than 10 kDa separated by LC-MS and analysed via Mascot search for cocoa albumin. Peptide hits (small bars) are placed against the amino acid sequence of albumin represented by the black bar on top. Peptides are presented for 120 hours of a commercial fermentation (top) and Trial 1 artificial fermentation after 24, 48, and 72 hours of incubation (bottom).

where peptides of albumin were found, although in a much smaller repertoire (Caligiani et al., 2016; Marseglia et al., 2014). With albumin being degraded at a much slower rate than vicilin (see the 21-kDa band in **Figure 3.2B**) together with a greater peptide repertoire present towards the latter stages of fermentation, it is plausible that along with free amino acids and peptides of vicilin, peptides of albumin could contribute to cocoa flavour as well.

3.4. Discussion

Until now, analyses of cocoa bean fermentation have been done with commercial fermentations on a large scale driven by microbial activity on the bean's exterior. As the sole exception, Biehl & Passern, 1982 had analysed cocoa bean biochemistry under aseptic conditions with full beans incubated in liquid solutions containing fermentation products. The incubation conditions described in their study are very similar to those described in this paper. The exceptions were extended incubation periods of three to five days in the current study as well as a more thorough decontamination procedure involving de-pulping with high pressure water followed by short rinse in 70% ethanol. The current study aimed to build on previous knowledge of protease activity under aseptic conditions based on the analysis of peptide fragments observed over the course of incubation. Thereby generated peptides are known to influence formation of flavour or aroma components during the roasting process as reported previously (Buyukpamukcu et al. 2001; Marseglia et al. 2014; Rashidah et al. 2007; Voigt et al. 2016).

In both trials of artificial fermentation described herein, a diverse repertoire of peptides was observed, some of which have not been observed before. It was concluded that due to a low pH being maintained throughout, there was considerably high endoprotease activity throughout the fermentation. For the commercial fermentation conducted in parallel, peptides of vicilin were only abundant after 120 hours of fermentation, after which they disappeared suggesting exopeptidase activity would later cleave these peptides into smaller oligopeptides and free amino acids and thereby reduce the peptide pool. Comparing Trial 1 of the artificial fermentation with the commercial fermentation, one may conclude that the bean incubations in buffers resulted in a more rapid degradation of protein as compared to the natural process. This rapid degradation seems to be a result of a swift temperature increase as well as the presence of acetic and lactic acids. In Trial 2 of artificial fermentation, one can observe a rapid decline in protein levels at the

very beginning which provides further evidence for the importance of acetic acid/low pH as contributing factors to stimulate proteolysis. Since there was no temperature alteration in this second trial, one has to question its role in promoting protein degradation during cocoa fermentation.

Proteolytic activity in both, Trials 1 and 2 of artificial fermentation, was the result of endoprotease activities followed by exoprotease activities. As it became evident that there is prominent endoprotease activity even at a relatively high pH of 5, a determining factor in fermentation might be that the internal cocoa bean pH should drop to a low enough pH range sufficient for generation of peptides. According to Trial 2, even though pH 5 showed considerable protease activity, there was not as much similarity of peptides as compared to Trial 1.

The herein observed remarkable similarity of peptides present in a commercial fermentation and in Trial 1 of the artificial fermentation highlighted the notion that native cocoa bean enzyme activity is sufficient for processing of storage proteins for the subsequent generation of flavour precursors. As this study focused on the processing of the two major storage proteins, it remains possible that other (even non-cocoa) proteins may still be substrates for the generation of additional flavour precursors. For example, previous studies with yeast cultures showed that the growth of different types of yeast during fermentation affected chocolate flavour (Leal Jr et al., 2008; Meersman et al., 2016). One could only speculate if the observed flavour differences could be attributed to the formation of volatile compounds or the diffusion of specific yeast-borne enzymes into the bean that could result in changes in the metabolome. However, our current study suggested that the majority of peptides is cocoa-borne and not of foreign nature. On the other hand, the observation that the peptides of Trial 2 of artificial showed much less resemblance to those of the commercial fermentation suggests that one is able to vary proteolytic activity significantly by varying the buffer constituents during artificial fermentation. It would be interesting to ascertain what flavour profiles could result from trials such as that of Trial 2 which do not conform to the normal pattern of fermentation by-product formation in cocoa bean fermentation.

Intriguingly, only peptides from vicilin and albumin could be detected with the method used herein irrespective of the more diverse assortment of peptides found in artificial fermentation as compared to commercial fermentation. This result could suggest that it is mainly

these peptides that are abundant and might represent important precursors for flavour development. In consequence, peptides derived from other cocoa proteins might not play a role in subsequent processes of aroma or taste development. Voigt et al. (2016) had managed to narrow down peptides generated from breakdown of vicilin that are potential flavour precursors for specific aroma notes. In the current study, one of these oligopeptides, *RNNPYYFPK*, as well as another with the sequence *NNQRIF* were found in Trial 1 of the artificial fermentation after 48 hours. In addition, several peptides that were reported to be present in well-fermented beans from different countries of origin (Marseglia et al., 2014) were also noted after 72 hours of incubation in Trial 1 of artificial fermentation. These peptides included those with the sequences *DEEGNFKI*, *NGKGTITF*, *QVKAPLSPGDVF*, and *ASKDQPLN* and consequently confirmed that the artificial fermentation system was able to generate peptides present in the commercial fermentation of various cocoa bean hybrids. We conclude that the proteolytic processing of proteins and the stability of resulting peptides found in an artificial fermentation can be made to resemble those detected in a commercial fermentation.

Since peptides from albumin were also detected during artificial and commercial fermentation, respectively, we can currently not exclude that precursors derived from albumin might contribute to flavour formation. The observed presence of intact albumin throughout the fermentation and the lower degree of exopeptidase activity on this storage protein would however suggest that albumin peptides are of rather minor importance for the later processes of flavour and aroma formation in chocolate production. By future manipulation of the herein described artificial fermentation system, one could possibly induce albumin breakdown and subsequently analyse taste and aroma of subsequently produced cocoa liquor using taste panel analysis.

Features of vicilin proteolysis observed in the current study confirmed the importance of aspartic proteinase in generating peptides thereby efficiently breaking down the storage proteins as previously shown and identified using its specific cleavage site. (D'Hondt et al. 1993; Voigt et al. 1994). Surprisingly, endoprotease activity was observed even at pH 5, suggesting that the endoprotease(s) of cocoa beans has a wider optimal pH range than previously observed and reported (Afoakwa et al. 2008; Guilloteau et al. 2005; Hansen et al. 1998; Laloi et al. 2002). However, previous experiments were conducted *in vitro* and Laloi *et al*. 2002 even used haemoglobin as a substrate. Voigt *et al*. (1994) observed proteolysis at pH 5.2 and demonstrated

that indeed aspartic proteinase was responsible for the observed endoproteolytic activity. In consequence, the *in vitro* detected high activity of cocoa endoprotease(s) at low pH might not fully represent the actual *in situ* activity of these enzymes. A wide working pH range of aspartic proteinase is further supported by the fact that the actual pH inside the cocoa bean does not decrease below 4.5 during a fermentation despite the acidic environment the bean pulp is exposed to (Doyle and Beuchat, 2007).

Preference of phenylalanine and tyrosine residues as C-termini of generated peptides suggested the activity of carboxypeptidase(s) (Auld & Vallee, 1970; Steitz et al., 1967). A significant preference for these two C-terminal residues in peptides found at the beginning of Trial 2 fermentation and the control run at pH 3 revealed that the carboxypeptidase(s) also displays broad range activity as in the case of the aspartic endoprotease. Albeit, the reduced preference for these two residues in the control run at pH 5 was astonishing, seeing as one would expect high carboxypeptidase activity closer to its optimal pH of 5.8 (Bytof et al. 1995; Hansen et al. 1998). We cannot however conclude that there was indeed less carboxypeptidase activity as the free amino acid concentration would also have to be determined to substantiate this claim accordingly. However, this phenomenon could still be explained by an optimal pH that prevented the inhibition of the carboxypeptidase by these aromatic amino acids.

The additionally observed preference for C-terminal methionine residues as well as the high endoprotease activity at pH 5 could hint towards the presence of one or more other (unknown) protease types in cocoa beans. This assumption is supported by the presence of at least 170 reported proteases in the *Theobroma cacao* genome according to the Cacao Matina1-6 Genome v1.1 genome sequencing project[1]. Alternatively, at altered pH the cleavage specificity of the aspartic proteinase(s) might change due to pH-mediated conformational changes that results in exposition of different protein domains (Everitt, Persson, & Wohlfart, 1988) such that it cleaves after methionine residues.

The herein conducted experiments demonstrated that the microbial fermentation-mimicking presence of fermentation by-products is driving native proteolytic activities within the cocoa bean leading to the generation of peptide-based flavour precursors. In consequence, we submit that the contribution of microbes to protein processing is rather minimal and restricted to providing inducing signals but not inducing proteolytic activities *per se*. The methods for

1 http://www.cacaogenomedb.org/search/search_by_kegg

conducting bench-scale artificial fermentations developed in this study will now be employed to compare compositional changes in other major cocoa metabolites such as polyphenols, carbohydrates, and lipids. These methods will furthermore be used to manipulate parameters and stimuli and test the effects of such manipulations on the composition of cocoa metabolites important for the flavour and aroma of cocoa.

3.5. Acknowledgments

The authors gratefully acknowledge Kouame Jean Koffi for his help in the procurement of pods from Ivory Coast as well as for performing trials of large scale commercial fermentation in Ivory Coast.

4. Results – Part II

Forcing Fermentation: Profiling Proteins, Peptides and Polyphenols in Lab-scale Cocoa Bean Fermentation

Warren A. John, Nina L. Böttcher, Maximilian Aßkamp, Audrey Bergounhou, Neha Kumari, Ping-Wei Ho, Roy N. D'Souza, Elke Nevoigt and Matthias S. Ullrich.

Abstract:

This study encompassed the lab-scale fermentation of cocoa beans in 300-g heaps under controlled laboratory conditions in order to replicate the microbial dynamics and metabolomic changes that usually occur in large-scale spontaneous fermentations. Remarkably, the growth profiles of yeast and acetic acid bacteria (AAB) with the native assortment of microbes as well as with the use of a starter culture were very similar to those observed in literature. Greater production of acetic acid by AAB not only led to more acidic-tasting liquor but also contributed to bitterness due to polyphenol preservation. We observed an initial increase in polyphenol content which is a phenomenon reported in literature but has neither been considered reproducible nor relevant. An increased production of acetic acid also brought about a drastic drop in pH leading to greater proteolytic activity as evidenced by the degradation of protein and the generation of a larger peptide repertoire. Peptides generated through proteolysis also showed incredible similarity to those reported in literature, in particular, those speculated to be involved in cocoa-specific flavour. A closer look at the naturally occurring peptide repertoires of our fermentation trials, generated by the breakdown of cocoa storage protein, pointed to a potential peptide responsible for cocoa-specific aroma.

4.1. Introduction

The process of cocoa bean fermentation is vital to the production of enjoyable chocolate and is practiced differently from one region to the other (Saltini et al., 2013). This process has been developed over centuries, originating in Mesoamerican society, and is typically conducted on plantations immediately after harvesting cocoa pods. Pods are manually cut open using knives or machetes, and the beans are removed by hand, piled up in heaps ranging from a meagre 100 kg to a massive 1,000 kg, and left to stand for a period of three to ten days with frequent turning of the beans. After fermentation, the beans are spread out to dry until they reach a moisture content of less than 7% and shipped off to the factory for roasting, milling and conching (Afoakwa, Paterson, Fowler, & Ryan, 2008; Schwan & Wheals, 2004).

A microscopic look at the surface of the beans during fermentation reveals consecutive successions of three groups of microorganisms, namely, yeasts, lactic acid bacteria (LAB) and acetic acid bacteria (AAB). Yeasts begin by degrading the pulp surrounding the bean, consuming available oxygen, sugars and citric acid. Citric acid is a non-volatile organic acid and its concentration remains fairly stable with slight decrease during the course of fermentation. Secreted pectinases from yeasts aid in the process of pectin breakdown. In turn, they generate an anoxic microenvironment, which favours the production of ethanol. The hereby commonly observed yeasts are *Saccharomyces cerevisiae*, *Hanseniaspora* sp., *Candida* sp., and *Pichia* sp. (Papalexandratou & De Vuyst, 2011; Rombouts, 1953). LAB such as *Lactobacillus* sp. and *Leuconostoc* sp. start to thrive under these anoxic conditions, until the beans are turned over by the farmer and oxygen is reintroduced into the system. The growth of LAB gives rise to lactic acid, which is another non-volatile organic acid, produced from the available ethanol. However, the survival of LAB is rather short. With oxygen brought back into play, AAB such as *Acetobacter* sp. and *Gluconobacter* sp. commandeer the microbiome and utilize the available ethanol and lactic acid to produce volatile acetic acid in a noticeably exothermic reaction, which starts to raise the temperature of the bean pile to highs of 50–60°C (Camu et al., 2007).

The diffusion of the secreted microbial metabolites into the cocoa bean cotyledon and the temperature increase give rise to structural and biochemical changes that eventually lead to the development of important flavour precursors (Biehl et al., 1982). It has been observed that there is breakdown of storage proteins into peptides and amino acids and losses in polyphenol content,

which in combination eventually go on to form essential flavour components, mainly during the roasting process (Albertini et al., 2015; Mohr et al., 1976; Wollgast & Anklam, 2000). The majority of the bean's storage protein exists in the form of two proteins, namely albumin (also referred to as 21-kDa seed storage protein) and vicilin. It has been described that peptides deriving from vicilin are the key contributors to quality-determining aroma and flavour precursors (Voigt, Heinrichs, Voigt, & Biehl, 1994). Extensive research has been carried out to elucidate the contribution of vicilin-derived peptides to flavour (Voigt et al., 2016, 2018).

Due to a plethora of factors that can influence fermentation, the degree of fermentation and flavour profile can vary considerably even within a single plantation. More recently, research has tapped into the use of chemically driven fermentation. This involved incubation of beans in solutions made to mimic the changing exposure to various microbial metabolites (Biehl & Passern, 1982; Eyamo Evina et al., 2016; John et al., 2016; Kadow et al., 2015). Likewise, multiple studies involving the use of starter cultures have been carried out on lab-scale and pilot-scale fermentations (Buamah et al., 1997; Lefeber, Gobert, et al., 2011; Moens et al., 2014; Pereira et al., 2012; Romanens et al., 2018; Sandhya et al., 2016; Steensels & Verstrepen, 2014). These attempts were done in order to achieve greater control over the fermentation, an accelerated post-harvest process and a desired flavour profile for the resulting chocolate. A major challenge of working with small-scale fermentations under laboratory conditions is to ensure that each trial results in well-fermented beans in terms of their metabolomic profiles. For instance, microbial growth can proceed in a number of ways even on a matrix of cocoa pulp, which has led to much difficulty in the study of fermentation in specialized research laboratories (Schwan & Wheals, 2004). Moreover, whilst the majority of current research mainly focuses on either volatile fingerprints (Afoakwa et al., 2008), microbial population studies (De Vuyst & Weckx, 2016), or specific biochemical foci (Voigt et al., 1994), very little research has been devoted to linking metabolomic profiles of fermentation trials to the resulting flavour profiles of the corresponding liquors or chocolates.

The following research describes small-scale cocoa bean fermentations performed under laboratory conditions and driven by external temperature regimes. The objectives of this research were to study the effects of a yeast extract-based microbial supplement and a starter culture on the bean metabolome during fermentation, and to correlate the profiles of proteins, peptides and polyphenols to the flavour profiles of the resulting liquors. Microbial population dynamics

studies were also conducted in order to corroborate the production of the main microbial metabolites and show how the bean metabolome responds correspondingly. Our research was based on the hypothesis that a small-scale fermentation under the influence of a microbial supplement or starter culture and temperature regime would show similar profiles of proteins, peptides and polyphenols as compared to those of a spontaneous fermentation. Three 7-day long trials of fermentation were undertaken consisting of bean aliquots of 300 g placed inside glass bottles. While trial A was run without any additives, trial B had the addition of a microbial supplement at the start of the fermentation, and trial C was inoculated with a starter culture consisting of *S. cerevisiae* and *Acetobacter pasteurianus* prior to fermentation.

4.2. Materials and Methods

4.2.1. Chemicals and reagents

Agar-agar (microbiology grade), calcium carbonate (99%), chloramphenicol (≥ 98.5%, for biochemistry), citric acid (≥ 99.5%), ethanol (≥ 96%, denatured with 1% methyl ethyl ketone), fructose (≥ 99.5% for biochemistry), D-glucose monohydrate (≥ 99.5%), glycerol (86%), HPLC-grade water (Rotisolv®), hydrochloric acid (Rotipuran®, 32%) lactic acid (90% synthetic grade), MRS-broth (microbiology grade), peptone (from casein, microbiology grade), sodium acetate (≥ 99%), sodium chloride (> 99.8%), sodium dodecyl sulphate (SDS, ≥ 99.5% electrophoresis grade), Tris-HCl (Pufferan®, ≥ 99.5%) and yeast extract (bacteriology grade) were purchased from Carl Roth (Karlsruhe, Germany). Acetic acid (100%, analytical grade), acetone (100%, Biochemica), acetonitrile (ACN, LC-MS grade), dichloromethane (stabilized with 20 ppm amylene, analytical grade), dithiothreitol (DTT, Biochemica), methanol (100%, LC-MS grade) and sulphuric acid (90 – 91%, pure) were purchased from Applichem (Darmstadt, Germany). 2,4,6-Tris(2-pyridyl)-s-triazine (TPTZ, ≥ 98%), cycloheximide, formic acid (≥ 98% for mass spectrometry), gallic acid and iron (iii) chloride hexahydrate (reagent grade) were purchased from Sigma-Aldrich Chemie (Steinheim, Germany). dNTP mix (10 mM), 10X PCR buffer (ThermoPol®), and *Taq* polymerase were purchased from New England Biolabs, MA, USA. Primers were purchased from Eurofins Genomics, Ebersberg, Germany. Ethanol (LiChrosolv® HPLC grade) as an HPLC standard was purchased from Merck, Darmstadt, Germany. Chromafil® PTFE filters were purchased from Macherey-Nagel, Düren, Germany.

4.2.2. Starter culture preparation

Loopfuls of *Saccharomyces cerevisiae* and *Acetobacter pasteurianus* were taken from their glycerol stocks stored at -80°C and placed in yeast extract-peptone-dextrose (YPD; 1% yeast extract, 2% peptone and 2% glucose) medium for *S. cerevisiae* and acetic acid medium (AAM; 1% glucose, 0.8% yeast extract, 0.5% ethanol, 0.3% acetic acid and 1.5% peptone) for *A. pasteurianus* (Lefeber, Gobert, et al., 2011). Strains were cultured for a period of three days at 30°C with shaking at 250 rpm in a New Brunswick Scientific Innova series shaker incubator (Eppendorf, Hamburg, Germany).

4.2.3. Forced Fermentation

All equipment used in the processing was sterilized with either 70% ethanol or by autoclaving at 120°C in order to exclude any unintended microbial contamination of the system. Slightly premature pods of the hybrid named "German", which is of the Forastero variety were harvested from Ebol-Nkok in Edzendouan, Cameroon (GPS coordinates: 4° 17' 33.654" N, 11° 44' 6.292" E) and shipped to Bremen, Germany. The pods were of a green or green-yellow or yellow-orange colour with a diameter of between 10-15 cm. Pods were surface-disinfected by first rinsing with tap water to remove grit and debris, followed by careful wiping with 70% ethanol and finally with Sterillium® (Bode Chemie, Hamburg, Germany). Swabs from the surfaces of random pods were plated on YPD-agar (YPD broth in 1.5% agar), De Man, Rogosa and Sharpe agar (MRS-agar; 52 g/L MRS broth, 2% agar) and glucose-yeast extract-calcium carbonate agar [GYC-agar; 10% glucose, 1% yeast extract, 2% calcium carbonate, 1.5% agar, pH 6.8; (Sharafi, Rasooli, & Beheshti-Maal, 2010)] for validation of sterilization. Pods were cut open using a clean, sterile knife inside a laminar flow sterile bench. Beans from pristine pods were selected and 16 kg in total were collected in a disinfected container while damaged or partially rotten pods were disregarded. The beans were treated in the container as follows in the three trials: Trial A (no treatment); Trial B (addition of 100 mL/kg beans 10 g/L yeast extract-based supplement); Trial C (addition of 100 mL/kg beans 0.25 OD_{600} *S. cerevisiae* and 0.05 OD_{600} *A. pasteurianus* re-suspended in 10 g/L yeast extract-based supplement).

Beans were divided into 300-g portions and placed inside sterile 1-L glass bottles, one bottle for every day of fermentation, except for the 72-h and 120-h time points, which had two bottles each. Bottles were sealed shut in order to prevent gas exchange and to allow the microbes

Table 4.1: Incubation temperatures of the forced fermentation system.
Beans were incubated in 1-L bottles where one replicate represents one bottle containing 300 g cocoa beans. Therefore the trial A started with 16 bottles at the beginning of fermentation with one bottle taken out every 24 hours except after 72 and 120 hours of fermentation where two bottles were removed.

Day	Temperature (°C)	Replicates		
		Trial A	Trial B	Trial C
1	**30**	1	1	1
2	**30**	1	1	1
3	**30**	2	2	2
4	**35**	1	1	1
5	**40**	2	1	2
6	**45**	1	1	1
7	**45**	1	1	1

to gradually generate an anaerobic headspace for the first 48 h. After two days, the lids were left slightly open, and the headspace was subsequently flushed with sterile compressed air every 24 h.

Bottles were placed inside an incubator with specific temperatures for each day of fermentation according to **Table 4.1**. Except for the first 48 h of fermentation, the beans were turned over using a sterile spatula every 24 h. Samples were taken every 24 h as such: one bottle of 300 g beans was removed from the experiment while the rest of the bottles were allowed to remain in the incubation chain. This was done to avoid disturbance to the fermentation and prevent any aerial or external contamination.

4.2.4. Liquor production with Sensory Profiling

Simultaneous to the above experiment, trials A, B, and C were repeated in seven additional bottles each carrying 300 g of beans. After seven days of fermentation, 2 kg of beans removed from the bottles, spread out over aluminium foil and placed in a draught oven (Binder, Tuttlingen, Germany) at 45°C for 72 h to dry completely. 1 kg of dry bean material from each fermentation trial was shipped to Barry Callebaut, Lebbeke-Wieze, Belgium for liquor and chocolate production. De-shelled cocoa nibs were roasted in a standard oven for 30 min at 120°C and then milled. Liquor was tasted and scored by a trained taste panel of 17 persons.

4.2.5. Material Sampling

Bean samples were taken prior to the start of fermentation and every 24 h. Temperatures of the incubator and the bean mass were measured prior to sampling. For microbial analysis, 20-25 g of beans were re-suspended in 20 mL sterile saline solution (0.9% sodium chloride) in the

sterile bench and shaken vigorously for 30 s. The resulting slurry was immediately applied on to agar medium as explained in section 2.8. For analysis of microbial metabolites, 20-25 g of beans were re-suspended in 20 mL MilliQ water and shaken vigorously for 1 min. Beans were discarded and the resulting slurry stored at -20°C until further processing. The remaining beans inside the bottles were aliquoted to 50-mL centrifuge tubes and stored at -20°C until further processing. Dried beans were aliquoted to 50-mL centrifuge tubes and stored at 4°C until further processing.

4.2.6. Bean sample preparation

Cocoa beans were de-shelled and ground in a Grindomix GM200 (Retsch, Haan, Germany) with speeds in the following order: 3,000 rpm for 20 s, 10,000 rpm for 10 s, 5,000 rpm for 10 s, 10,000 rpm for 5 s, 2,500 rpm for 10 s, and 10,000 rpm for 5 s. These ground cotyledon samples were stored at -20°C until further processing except for the dry bean samples, which were stored at 4°C. Liquor samples were ground at 3,000 rpm for 20 s and stored at 4°C.

Aliquots of 5-6 g of ground sample were defatted using dichloromethane with standard Soxhlet extraction for 18 h using a Büchi B-811 system (Essen, Germany) to produce dry defatted powder (DDP). The DDP was stored at 4°C until further use.

4.2.7. Internal bean pH and moisture content

2.5 g of ground bean material were re-suspended in 22.5 mL MilliQ water in 50-mL centrifuge tubes and shaken head-over-head for 5 min. The suspension was centrifuged at 3,200 *g* for 10 min, and the pellet was discarded. A pH meter (Mettler Toledo, Giessen, Germany) was used to measure the pH of the supernatant. Values for the 72-h time point of trials A, B and C, as well as for the 120-h time points of trials A and C, were from two biological duplicates whereas all other measurements were done with single samples.

For moisture determination, 2 g of ground bean material were placed in an aluminium boat and incubated in the Binder draught oven at 105°C for 24 hours. The resulting powder was cooled down and weighed out to determine the moisture loss from the bean.

4.2.8. Culture-dependent microbial population dynamics analysis

100 µL of the sample for microbial population analysis from section 2.5 were used to prepare serial dilutions of 10^0 – 10^5 in saline solution. 100 µL of each dilution were plated on

YPD-agar supplemented with 100 mg/L chloramphenicol and GYC-agar supplemented with 0.1% cycloheximide. Plates were incubated at 30°C for 3-5 days. Colony-based polymerase chain reaction (PCR) yielding in the amplification of the 16S rRNA gene was performed on unique colonies of each dilution. For this, a small mass of each colony was re-suspended in 40 µL of sterile PCR buffer and vortexed vigorously for 30 s. The suspension was subsequently sonicated for 5 min in a Sonorex ultra sonication bath (Bandelin, Berlin, Germany) and incubated at 95°C for 10 min. Cell debris was spun down using a centrifuge at maximum speed for 1 min and the supernatant was used as DNA extract.

For yeast D1/D2 LSU gene amplification, 1 µL of DNA extract was added to a 49 µL PCR mix containing 5 µL 10X PCR buffer, 1 µL 10 pmol/µL each of forward primer LR6 (CGCCAGTTCTGCTTACC) and reverse primer LR0R (ACCCGCTGAACTTAAGC), 1.5 µL 10 mM dNTP mix and 0.2 µL of 5000 U/mL *Taq* polymerase all re-suspended in sterile MilliQ water (Daniel et al., 2009). For bacterial 16S rRNA gene amplification, 1 µL of DNA extract was added to a 49 µL PCR mix containing 5 µL 10X PCR buffer, 1 µL 10 pmol/µL each of forward primer (AGAGTTTGATCCTGGCTCAG) and reverse primer (TACGGYTACCTTGTTACGACTT), 1.0 µL 10 mM dNTP mix and 0.2 µL of 5000 U/mL *Taq* polymerase all re-suspended in sterile MilliQ water (Alfaro-Espinoza & Ullrich, 2014).

4.2.9. Microbial metabolite analysis

Samples for microbial metabolite analysis as explained in section 2.5 were handled at 4°C with all the equipment cooled to 4°C.

For measurement of ethanol, acetic acid and lactic acid, samples were spun down in an Eppendorf 5810R centrifuge (Hamburg, Germany) at 3,200 *g* for 15 min at 4°C. The pellet was discarded and the supernatant passed through a 0.45-µm pore size filter followed by a pH measurement. The supernatant was analysed using a Waters HPLC system, comprising the following components: Waters 2414 refractive index (RI) detector, Waters 2489 ultraviolet (UV) detector, Waters 1525 binary pump system, Waters 2707 injector system and a Waters column heater module WAT038040 (Waters, MA, USA). 20 µL sample were injected to an Aminex HPX-87H column (Biorad, CA, USA) with a pH range of 1-3 containing 8% cross-linked resin attached to a Micro-guard® cation H cartridge (Biorad). An eluent of 5 mM sulphuric acid was used with a flow rate of 0.6 mL/min. The column temperature was set to 45°C and the internal

heater to 30°C (John et al., 2016). Quantification was performed for citric acid (as a control for extraction), ethanol and lactic acid with RI detection and for acetic acid with UV detection. Analysis was performed using the Breeze™ 2 v6.2 software.

For the measurement of antioxidant capacity, 2 mL of the slurry were centrifuged at 16,000 *g* for 10 min at 4°C. The pellet was discarded and the supernatant was freeze-dried. 500 µL of extraction buffer (MeOH:H_2O:CH_3COOH::70:28:2) was added to the freeze-dried powder, and the suspension was left to sonicate for 10 min and then shaken vigorously for 10 min. The sample was then spun down for 5 min at 16,000 *g*. 10 µL of the supernatant as well as standards of gallic acid were mixed with 200 µL FRAP reagent in a 96-well plate in triplicate. The FRAP reagent was prepared by mixing 4 mL TPTZ solution (10 mM TPTZ in 40 mM HCl), 4 mL 20 mM $FeCl_3.6H_2O$ solution and 40 mL acetate buffer (150 mM acetic acid, 12 mM sodium acetate, pH 3.6). After mixing, the plate was left in the dark for 10 min at room temperature and absorbance was measured at 593 nm using an EZ Read 2000 plate reader (Biochrom, Cambridge, UK).

4.2.10. Protein Extraction and Quantification

Acetone-dry powder was prepared by washing 1 g of ground bean in 10 mL acetone for three 30-min cycles, with a spin down at 3,200 *g* after every wash. The resulting powder was left to dry overnight and was then mixed with 10 mL protein extraction buffer (John et al., 2016). The suspension was heated for 10 min at 80°C and left on head-over-head rotation at room temperature for 1 h. The suspension was then centrifuged at 16,000 *g* for 20 min and the resulting supernatant filtered through a Vivaspin column with a 10-kDa cut off. The filtrate was purified with ice-cold acetone in a ratio of 1:4 with an overnight incubation at -20°C. The subsequent protein precipitate was spun down and re-suspended in 50 mM Tris-HCl (pH 8.0). The purified protein extract was diluted in a ratio of 1:10 in 50 mM Tris-HCl before being quantified using a 2D-Quant protein quantification kit (GE Healthcare, NJ, USA). Results were normalized with the moisture content of the bean samples.

4.2.11. Peptide and Polyphenol Extraction

Extraction was done on 50 mg DDP using 5 mL extraction buffer (MeOH:H_2O:CH_3COOH::70:28:2) with sonication for 10 min followed by stirring for 30 min.

The polyphenol/peptide (PP) extract was then passed through 0.45 µm PTFE membrane filters, spiked with 2 µg/mL hesperetin as an internal standard and was immediately processed (D'Souza et al., 2018).

4.2.12. UHPLC-ESI-MS/MS

PP extracts were processed as described earlier (D'Souza et al., 2017) with an Agilent 1260 HPLC system using a ZORBAX Eclipse Plus C18 column coupled to an Impact HD ultra-high resolution ESI-q-TOF mass spectrometer (Bruker Daltonics, Bremen, Germany). 2 µL of the sample were injected for HPLC and eluted with a mobile phase of a binary gradient between MilliQ water and acetonitrile (B), both spiked with 0.05% formic acid, in a scheme of (t (min), %B): (0, 5); (1, 5); (8, 16.5); (9, 17); (10, 17.5);(11, 17.5); (12, 18.5); (13, 18.5); (23, 95); (25, 95), at a constant flow rate of 0.5 mL/min. A column temperature of 40°C was maintained. The electrospray ionization source had a nebulizer pressure of 1.8 bars, a dry gas flow rate of 9 L/min and a dry gas temperature of 200°C. Before each run, the mass spectrometer was calibrated with 0.1 M sodium formate with calibrations being done in HPC mode (DataAnalysis 4.1, Bruker Daltonics). Data were acquired in both negative- and positive-ion mode with MS^n runs being carried out on a select group of samples.

4.2.13. Polyphenol and Peptide identification

Masses of previously observed polyphenols were queried in negative-ion mode spectra and their corresponding peaks were integrated and normalized to the hesperetin internal standard using DataAnalysis 4.1 (D'Souza et al., 2018).

Mascot generic format (MGF) files extracted from positive-ion mode spectra were analyzed using PeptideShaker. Peptides corresponding to the six most abundant cocoa bean proteins as discovered in previous proteomics research, namely, albumin (CAA39860), vicilin (EOY05738), oleosin 1 (XP_007036421), chitinase class 1 (AAA80656), lipoxygenase (EOY28236) and maturase K (AAQ84277) were searched for based on their fragmentation spectra (D'Souza et al., 2018).

4.3. Results

4.3.1. Temperature regimes help control microbial dynamics

The dynamics of the main microbial metabolites on the bean's exterior play a major role in determining which biochemical and structural changes go on inside the bean. Investigating these changes is therefore important in ascertaining the degree of fermentation. HPLC coupled to refractive index detection was used to quantify the three main microbial metabolites, ethanol, acetic acid and lactic acid, respectively. It was observed that ethanol was produced during all three trials of forced fermentation, with trial B, containing the microbial supplement, achieving

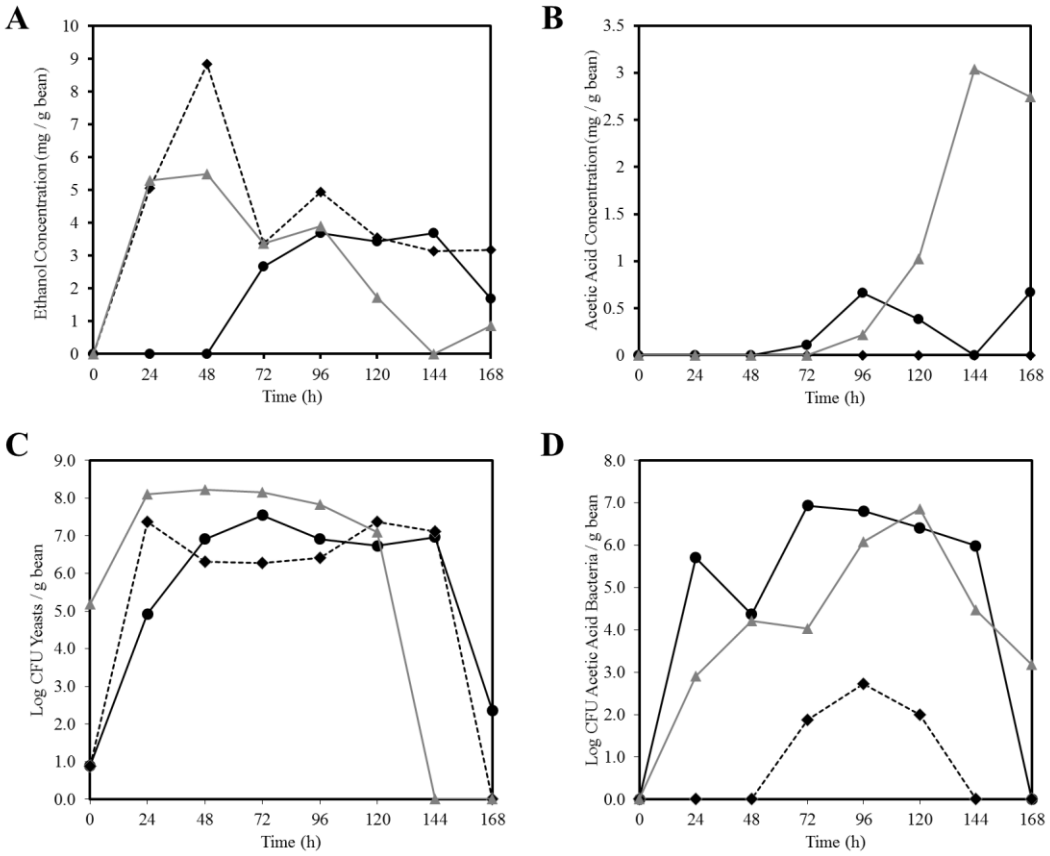


Figure 4.1: Microbial population dynamics and metabolism on the bean pulp during trials of forced fermentation.
Concentrations of ethanol (A) and acetic acid (B) quantified using HPLC-RI on the pulp of the beans and values of log CFU/g bean determined using culture-dependant analyses for yeasts (C) and acetic acid bacteria (D) for trial A (circle with black solid line) as the control, trial B (diamond with black dotted line) which included the use of a microbial supplement and trial C (triangle with grey solid line) which involved the use of a starter culture. Each time point is its own fermentation "heap" that was stopped and sampled after the corresponding time indicated. Microbes were enumerated on plates of YPD, GYC and MRS.

the highest ethanol yield of 9 mg/g bean among the three trials and peaking at 48h. (**Figure 4.1A**).

Acetic acid was present in trial A (control) and trial C (with starter culture) (**Figure 4.1B)**. Concentrations of acetic acid started to rise during the mid-latter stages of the fermentation. Trial C saw acetic acid concentrations reaching values up to 3 mg/g bean, far higher than the control trial. The addition of AAB in trial C prior to fermentation accounts for this. Lactic acid was not detected in any of the trials, which was not surprising since LAB neither were added nor present (data not shown).

Population studies revealed a staggering growth of yeasts over the first 48 hours by a scale of roughly six orders of magnitude (**Figure 4.1C**) with all three trials peaking between 10^7 and 10^8 CFU/g bean. Trial C had a greater population of yeasts overall, primarily due to the inoculation with *S. cerevisiae* prior to fermentation. The most abundant native species of yeasts were found to be *S. cerevisiae*, *Hanseniaspora valbyensis* and *Pichia kudriavzevii*. Trials A and C encountered growth of AAB already at the early stages of the fermentation, reaching maximal values of 10^7 CFU/g bean, whereas growth was delayed by 48 h in the case of trial B and was substantially less throughout the fermentation (**Figure 4.1D**). The observed species of AAB were *Acetobacter tropicalis*, *Gluconobacter oxydans*, *Gluconobacter frateurii* and *Gluconobacter nephelii*. Initially, more diverse microbiomes were found to give way to more specific key players, such as *S. cerevisiae*, *Hanseniaspora valbyensis*, *Pichia. kudriavzevii* and *Acetobacter* sp. and *Gluconobacter* sp. during the course of the fermentation (**Figure 4.2**).

Surprisingly and despite the lack of experimentally added microbial influence, trials A and B, which did not have any microbial inoculation, showed remarkable dynamics and astoundingly similar profiles of growth as seen in trials of spontaneous fermentation (Camu et al., 2007; De Vuyst & Weckx, 2016). It was also evident that high temperatures of 45°C strongly inhibited growth of yeast and bacterial populations towards the end of the fermentation (**Figures 4.1C and 4.1D**). This involved sudden drops in CFU counts as drastic as 10^7 to 10^0 within 24 h.

4.3.2. Production of microbial metabolites correlates with acid influx

Internal pH of the bean is fundamental in determining the fate of the cocoa bean's biochemistry, especially in terms of enzyme activities. Each trial showed unique pH profiles during the course of fermentation. Beans in trials A and C underwent a sharp drop in pH, starting

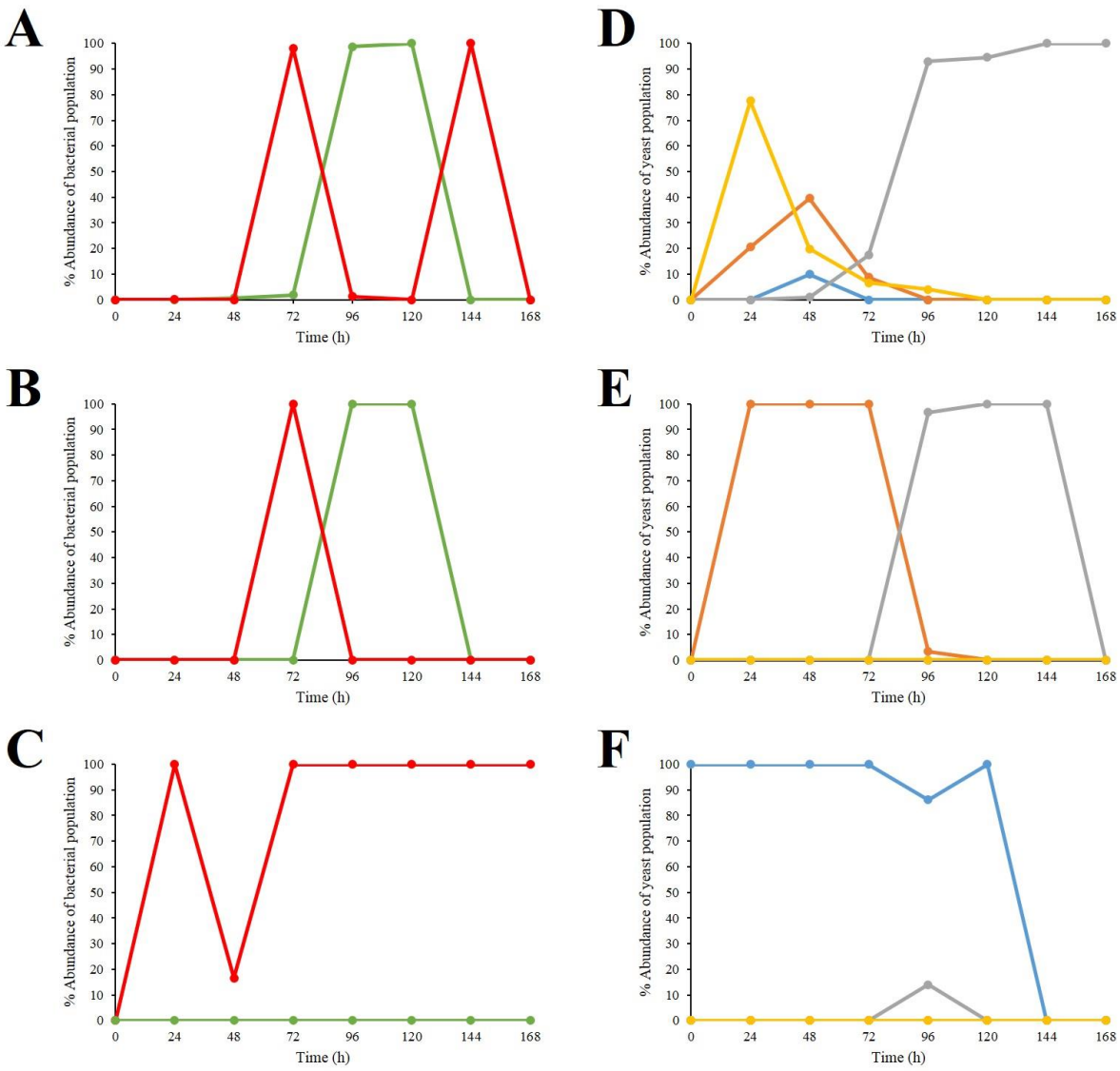


Figure 4.2: Growth of the most abundant species of yeast and bacteria in the trials of Forced Fermentation. Percentage abundance of *Acetobacter* sp. (red) and *Gluconobacter* sp. (green) in bacterial populations of trial A (A), trial B (B) and trial C (C). Percentage abundance of *Saccharomyces cerevisiae* (blue), *Hanseniaspora valbyensis* (orange), *Pichia kudriavzevii* (grey) and *Meyerozyma guilliermondii* (yellow) in yeast populations of trial A (D), trial B (E) and trial C (F).

at around 6.6 and dropping to 4.5 and 5.0 (**Figure 4.3A**). This was not the case for trial B, which only showed a mild pH decrease. According to literature, the pH of a good-quality fermentation generally drops to about pH 4.5 (Thompson et al., 2007), which trials A and C reflected in their profiles. It is evident that the sharpest decline in pH occurred after 96 hours, which is likewise in direct correlation with higher temperatures, the growth of AAB, and the onset of acetic acid

production in the bean's pulp. Conversely, the absence of acetic acid in the yeast extract-supplemented trial B, yielded no substantial decline in pH.

4.3.3. Acid influx affects flavan-3-ol dynamics

Flavan-3-ols are known to be the major polyphenols present in cocoa beans (Albertini et al., 2015; Wollgast & Anklam, 2000). To map their profiles, epicatechin and the procyanidin B-type oligomers were focused on in this study. Trials A and C showed increases in flavan-3-ol content while the opposite was true of trial B (**Figure 4.3B**). Epicatechin constituted the bulk of the measurement while the procyanidin dimers and trimers followed in similar manner (data not shown). The rise in content in trials A and C coincided with the production of acid in the fermentation (**Figure 4.1B)**, such that trial A in particular showed a dramatic increase of about 700 units between 72 and 96 hours of fermentation.

Interestingly, drying of the beans did not cause losses in flavan-3-ols in any of the three trials. Recent research has shown that shorter drying techniques preserve polyphenols (Santhanam Menon, Hii, Law, Shariff, & Djaeni, 2017). Accordingly, since the oven-drying procedure employed a constant temperature and draught at 50°C, the beans would have dried out much faster as opposed to several 12-hour cycles of daylight on a tropical plantation. The faster drying out of the beans therefore could explain the preservation of flavan-3-ols.

4.3.4. pH induced proteolysis favours generation of flavour/aroma precursors

All fermentation trials showed considerable decreases in protein quantity during the fermentation (**Figure 4.3C**). Proteins remained stable during the initial days of fermentation, which could be expected due to the rather mild incubation temperatures and the absence of acid. Towards the latter part of the fermentation, there was still a substantial amount of protein left. In contrast, in a spontaneous fermentation cocoa bean proteins would be almost, if not, completely broken down at the end of the fermentation (Kumari et al., 2016). In consequence, by elongating the fermentation by a span of three additional days, which involved a higher incubation temperature of 55°C, we found that there was indeed a complete breakdown of proteins (data not shown).

The herein employed method for peptide identification yielded one of the largest lists of peptides observed in cocoa bean fermentation with a cohort of 449 identified peptides

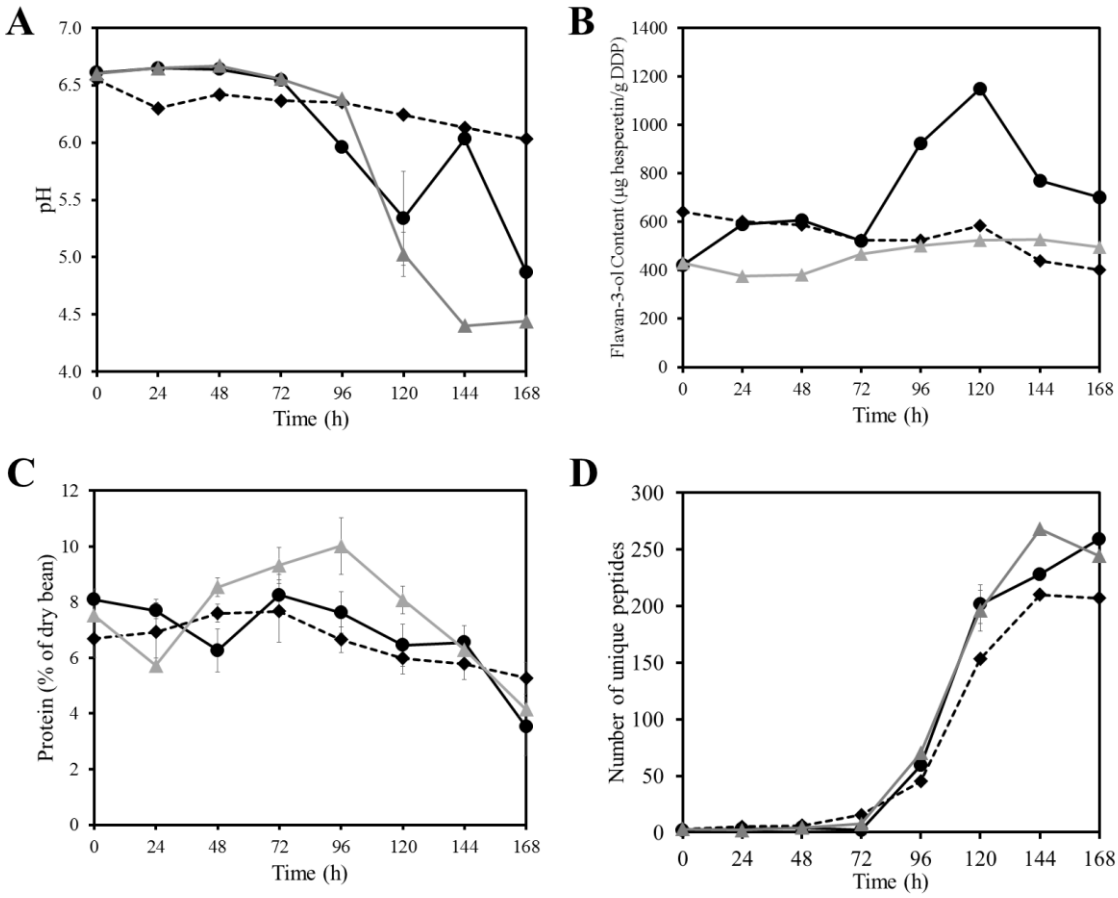


Figure 4.3: Metabolomic data for trials of forced fermentation.
Internal bean pH (A), total bean flavan-3-ol content relative to hesperetin (B), bean protein concentration (C) and diversity of naturally occurring peptides of 4-9mers in the bean (D) for trial A (circle with black solid line) as the control, trial B (diamond with black dotted line) which included the use of a microbial supplement and trial C (triangle with grey solid line) which involved the use of a starter culture. Each time point is its own fermentation heap that was stopped and sampled after the corresponding time indicated.

representing 4 – 23mers (Appendix **Table A1**). The diversity of generated peptides increased with progressive proteolysis, with an especially fast-growing repertoire of 4-9mers for all three trials (**Figure 4.3D**). Trial B had the lowest overall diversity thereby signifying a slower rate of proteolysis as compared to the other two trials. Proteolysis in cocoa beans is known to be driven mainly by aspartic endoprotease and carboxypeptidase activity (Bytof et al., 1995; Guilloteau et al., 2005), .The sum of these activities lead to the generation of flavour precursors. Since neither the production of any acid during the fermentation nor a drop in internal pH occurred in trial B (**Figures 4.1D and 4.3A**), it can thus be deduced that conditions for protease activity were not optimal for the generation of peptides in trial B. This would be especially true for aspartic

protease activity, which has a pH optimum of 3.5. This notion is additionally supported by the observation of much lower relative peptide intensities in trial B (**Figure 4.4**).

Furthermore, trials A and C revealed apparent decreases in peptide abundance and intensity when samples of dried bean and liquor samples were compared (**Figure 4.4**). This prompted the hypothesis that whatever peptides are lost during the roasting process would either directly undergo Maillard reactions to form flavour/aroma components, or be further broken down into smaller peptide fragments, which could possibly be converted into Maillard reaction products. Only the corresponding peptides derived from vicilin were taken into account herein since prior investigations had indicated that vicilin degradation products showed the highest propensity to form flavour/aroma precursors (Voigt et al., 1994, 2018). The resulting selection of peptides was further chiselled by considering only those peptides, which showed a percentage intensity loss of more than 50% between dry bean and liquor samples in all three trials. A

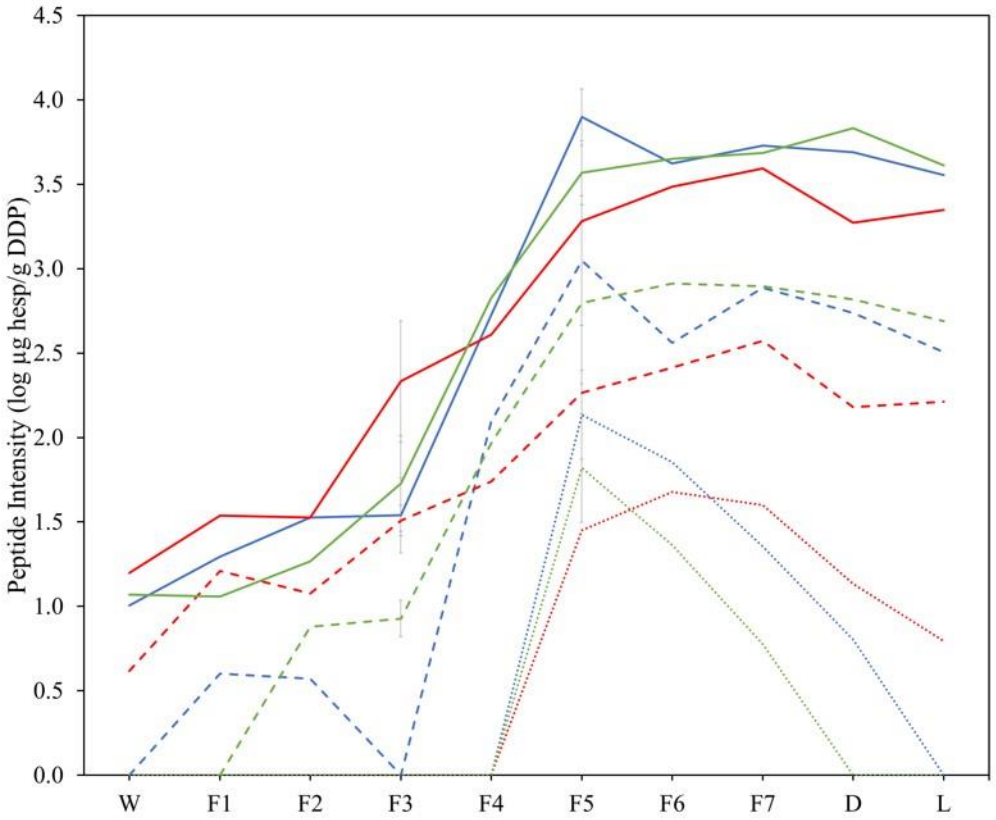


Figure 4.4: Relative intensities of naturally occurring peptides resolved by UHPLC-MS/MS.
Peptides of 4mers to 23mers were identified from positive-ion mode UHPLC-MS/MS spectra of methanol-based extracts of dry-defatted bean powder (DDP) and expressed in terms of their intensity for categories of 4-9mers (solid line), 10-16mers (dashed line) and 17-23mers (dotted line) for trial A (blue), trial B (red) and trial C (green) of forced fermentation. D: dry bean; L: liquor.

handful of peptides (*PVNSPGKY*, *NSPPLK*, *VTNGKGTITF*, *FQNMD*, *DEEGNFKILQ*, *ETVF*, *SEAKEL*, *IFNNPDESYF* and *NNPYY*) were discovered that originated from the full length of the 47-kDa vicilin fragment (**Figure 4.5**). Of those, *PVNSPGKY* was the most intense and showed the greatest decrease in intensity upon roasting (**Table 4.2**). These peptides appeared between 96 and 168 hours of fermentation with peptides *PVNSPGKY* and *NNPYY* appearing later during the fermentation (**Figure 4.6**). The remarkably high hydrophilicity of these nine peptides in combination with the generally ostensive hydrophilic nature of all peptides found in the dry bean supports the notion that hydrophobic free amino acids contribute most to aroma in cocoa (Mohr et al., 1976). This, too, hints toward a prominent role of hydrophilic peptides in cocoa aroma formation.

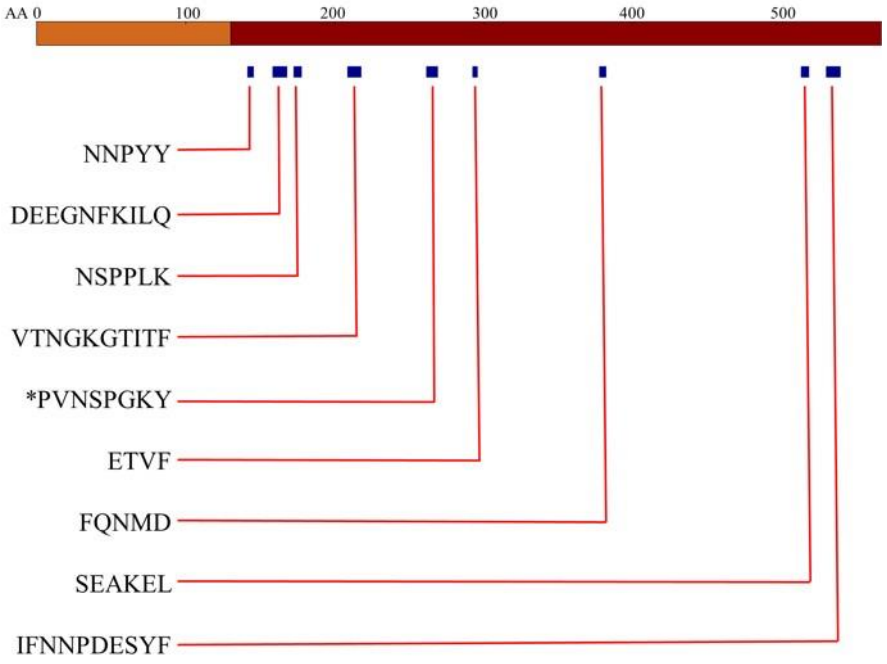


Figure 4.5: Naturally occurring peptides of vicilin found to be present in the dry bean and absent in liquor samples of forced fermentation.
Peptides of vicilin, resolved and identified by UHPLC-MS/MS, are represented by blue bars with their sequence corresponding to their position relative to the red bar above representing the full length amino acid (AA) sequence of vicilin. These peptides are found to have decreased by at least 50% in relative intensity during the processing stages from dry bean to liquor. *shown to have the highest loss in intensity.

Table 4.2: Naturally occurring vicilin-derived peptides which are lost from the dry bean upon roasting.
Nine peptides in total were found to be lost upon roasting in the liquor samples where some peptides show a total loss in relative intensity and some, a partial loss, based on a hesperetin internal standard. Peptides are arranged according to average percentage loss. *shown to have the highest loss in intensity.

Peptide	Percentage loss between dry bean and liquor (%)			
	Trial A	Trial B	Trial C	Average
*PVNSPGKY	100		47	73
NSPPLK	100		43	73
VTNGKGTITF	100		43	71
FQNMD	100		34	67
DEEGNFKILQ	100	28		64
ETVF	62		49	56
SEAKEL	52		57	55
IFNNPDESYF	100	8		54
NNPYY	49		52	51

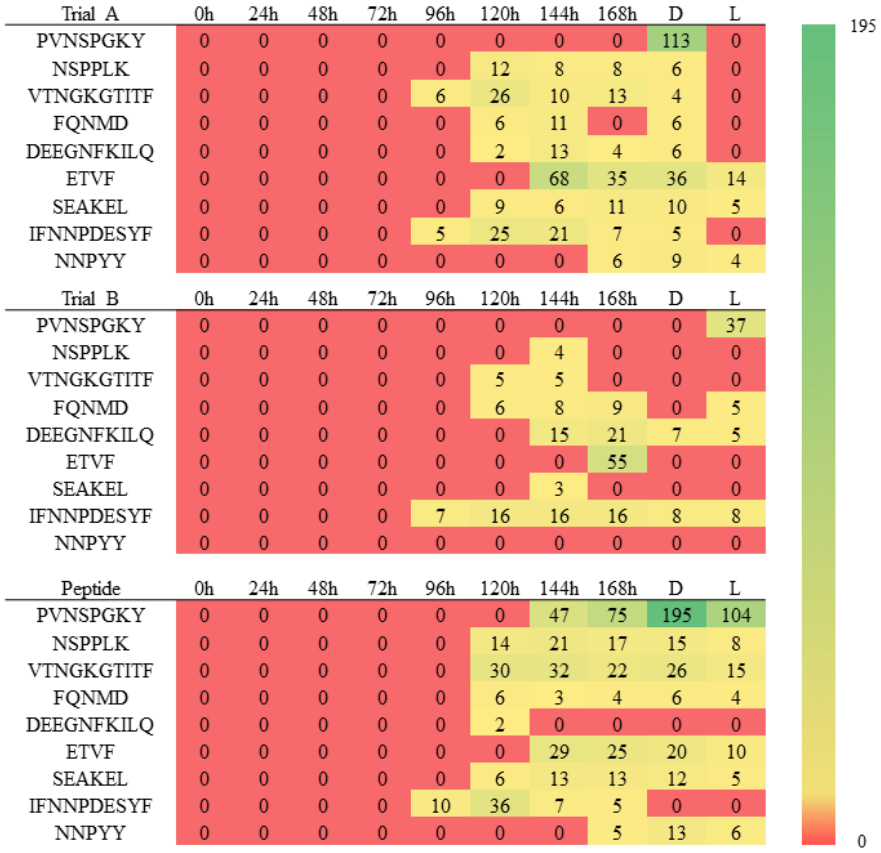


Figure 4.6: Relative intensities of vicilin-derived peptides which are lost from the dry bean upon roasting.
Nine peptides of vicilin, resolved and identified by UHPLC-MS/MS, are represented according to their relative intensities (µg hesperetin/g DDP). The colour bar represents the intensity of each peptide at each processin stage with the highest value being that of *PVNSPGKY* in the dry bean of trial C. D: dry bean; L: liquor. Each time point is its own fermentation heap that was stopped and sampled after the corresponding time indicated.

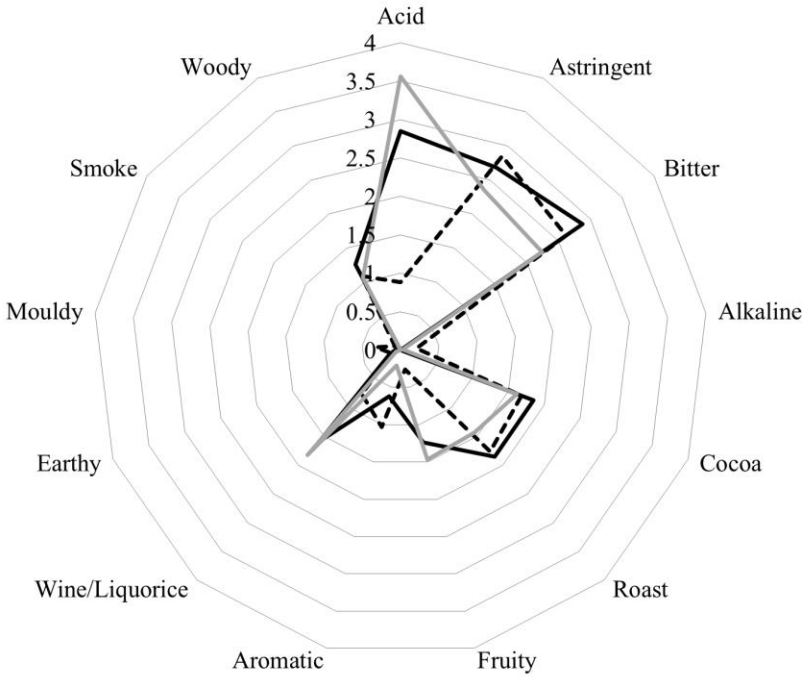


Figure 4.7: Taste profile of liquor samples for all three trials of forced fermentation.
Taste profiles were drawn up for liquor using a trained panel of 17 persons and the values representing averages on a scale of 0 to 4 for trial A (solid black line) as the control, trial B (dashed black line) which included the use of a microbial supplement and trial C (solid grey line) which involved the use of a starter culture.

4.3.5. Sensory profiles are dependent on microbial dynamics

Each of the three trials of forced fermentation gave rise to uniquely tasting liquors signifying the importance of microbial growth dynamics during fermentation (**Figure 4.7**). Flavour notes of acid, astringent, bitter, cocoa, roast, fruity and wine/liquorice were the most prominent. The greatest difference between the three trials was seen in terms of their acid and fruity notes, for which trials A and C were noticeably stronger tasting, by up to 1.5 units. This can be explained by the production of a much higher quantity of acetic acid compared to trial B (**Figure 4.1B**). Although trials B and C tasted similar in terms of their cocoa flavour, trial B had a more astringent flavour. The astringency of trial B could potentially be attributed to the lack of a substantial drop in internal pH (**Figure 4.3A**) in the beans as compared to the other trials, and a much smaller peptide diversity (**Figure 4.3D**) as well as an overall lower peptide intensities (**Figure 4.4**), signifying a somewhat under-fermented batch of beans.

4.4. Discussion

So far, cocoa bean fermentation trials for research purposes were preferred to be carried out in large scale in order to allow ample microbial growth to produce ethanol, acetic acid and lactic acid, and to generate the heap temperatures needed to drive fermentation. Remarkably, the fermentation trials conducted in this study, consisting of bean amounts of only 300 g, have demonstrated that, provided the right conditions, microbial dynamics can be reproduced and can influence internal bean biochemistry very similarly to spontaneous fermentation. Additionally, our results appear to stand contrary to previous research arguing that cocoa pods maintain a sterile interior until opening (Schwan, 1998; Schwan & Wheals, 2004). While this might indeed be the case for pristine pods, most pods are prone to damage or attack by insects or animals, allowing microbes to enter during pod development and through eventually microscopic cracks. Moreover, drying-up of the pod exterior after harvesting also compromises the integrity of the husk allowing microbes to enter. The current trials of forced fermentation would have made it difficult for even a few microbial CFUs originating from the pods to avoid being detected since the initial three-day incubation at 30°C boosted the growth of microbes. This was evidenced by the dramatic increases in the CFU values of yeasts during the first 24 h even when no yeast cells had been added to the trial.

Despite the growth of yeasts and AAB in trials of fermentation, in which no microbes were experimentally added, no LAB could be detected. This may be explained by the rather long initial incubation at 30°C, when yeasts would have outcompeted LAB, which thrive better at temperatures closer to 37°C (Lefeber et al., 2011). Spontaneous fermentation usually undergoes a more dramatic increase in temperature over the first 48 – 72 hours, thus favouring LAB growth. Recently, research by Romanens et al. (2018) included lab-scale fermentations on the site of harvest employing a more sudden temperature increase, reaching 45°C on day two already. Those experiments being carried out on site allowed for an immediate growth of microbes as well as the growth of ample LAB. However, in contrast to the current study, pods were not surface-sterilized and aerial contamination not avoided in their fermentation system. In turn, this could also have contributed to the generally higher CFU counts prior to fermentation in the former study.

The herein applied three-day pre-incubation period at 30°C was assumed to be necessary in order to efficiently acclimatize microbes after transport of pods by flight which would have left

them in the cold, as well as storage at 4°C overnight ahead of starting the fermentation trials. However, a visual inspection of the trials indicated that this pre-incubation period could have been shortened to just two days to generate the desired anoxic conditions during the initial stages and to therefore allow for growth of LAB. LAB have been reported to constitute a major proportion of the microbial population in spontaneous fermentations and are responsible for the development of volatile compounds that lead to specific aroma notes (De Vuyst, Lefeber, Papalexandratou, & Camu, 2010). In contrast, our results indicate that the presence of yeast and AAB alone is sufficient to cause profiles of bean biochemistry similar to those of spontaneous fermentation and result in cocoa-specific flavour.

The onset of acetic acid accumulation marked a critical point during the course of these lab-scale fermentation trials in terms of both, proteins as well as flavan-3-ols. Trial B, which did not undergo any acid production, showed the least proteolytic activity along with a constant decay of flavan-3-ols, while the opposite was true of trials A and C characterized by an initial increase in polyphenol content. These results were consistent with previous findings showing acetic and lactic acids to have protective properties for polyphenols (Eyamo Evina et al., 2016) as well as findings demonstrating increases in polyphenol content during the mid-stages of the fermentation (Albertini et al., 2015). One could speculate that the increase in polyphenol content during fermentation is a result of the bean mounting its ultimate defence response to counter the harshly deteriorating conditions generated by the influx of acid and low pH (Bayliak, Burdyliuk, & Lushchak, 2016).

The majority of cocoa bean research has demonstrated that polyphenols degrade over the course of fermentation such that fermented beans have a much lower content of polyphenols as compared to fresh beans (Jalil & Ismail, 2008). This has been largely attributed to oxidization to insoluble tannins and the leaching out of polyphenols from the bean into the surrounding pulp and subsequent running off in the sweating (Wollgast & Anklam, 2000). In the current study, this polyphenol-leaching effect could be observed in trials B and C as opposed to trial A: the higher ethanol content in the pulp of the former trials might have better dissolved polyphenols in the bean. In support of this, it was found that the pulp in trials B and C contained significantly higher antioxidant capacity indirectly indicative for higher polyphenol content as compared to the pulp samples of trial A (**Figure 4.8**). The drop in bean polyphenol content consequently reduces the overall bitterness in the roasted product, which explains why the liquor of trial A

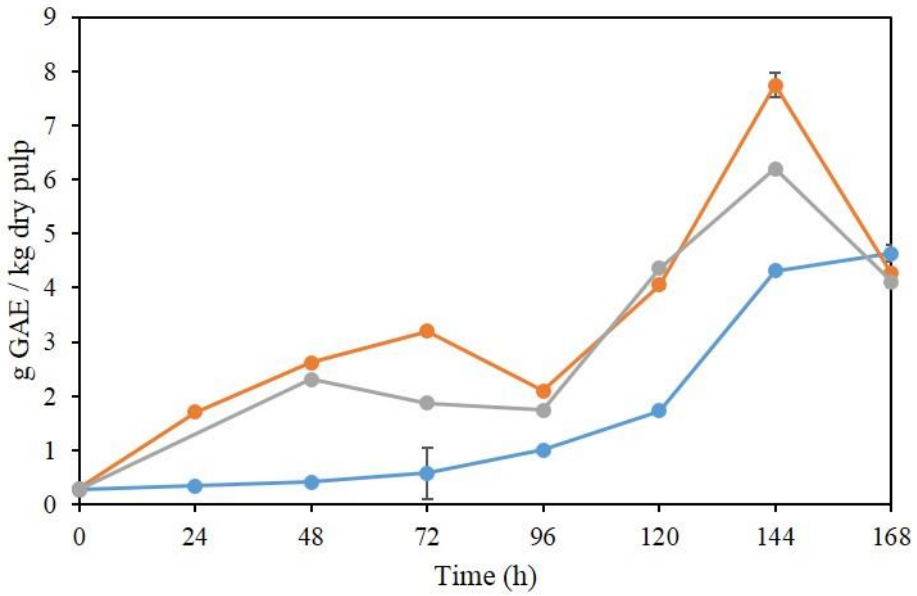


Figure 4.8: Antioxidant capacity measured in methanol extracts of cocoa pulp during trials of forced fermentation.
The FRAP assay was used to measure antioxidant capacity of dry pulp in terms of gallic acid equivalents (GAE) for trial A (blue), trial B (orange) and trial C (orange). n = 3.

tasted the most bitter. However, polyphenol dynamics during fermentation might not be as straightforward as initially thought. The presence of too much acid could eventually lead to a retention of or an increase in polyphenol levels. While this may increase the health benefits of the resulting chocolate, it could spell a compromise in taste in terms of bitterness.

A reduction in the bean's pH during fermentation leads to the generation of more peptides and free amino acids, which would become available to form Maillard reaction products during roasting, thereby contributing to more preferred aroma and flavour notes (Afoakwa et al., 2008). Therefore, pH is a crucial factor, not just because of its influence on acidity in the final product but also because of its role in transforming the bean metabolome thereby affecting the flavour and aroma profiles in several ways.

Furthermore, it became evident that the drop in pH in trials A and C favoured the formation of nine peptides, which were found to be present in the fermented and dried bean but absent in the corresponding liquor. Voigt et al. (2018) have recently reported a series of peptides in *in vitro* studies of proteolysis of cocoa-derived proteins. Interestingly, two of the peptides identified in the current study, namely *PVNSPGKY* and *NNPYY*, had been shown to be also present in the cocoa aroma-rich fractions of pH 4.8 and 5.2 of that former study. In our

experiment, peptide *PVNSPGKY* showed the highest change in intensity of the nine peptides and was occurring in substantial quantity only during the latter stage of fermentation as well as during drying. Several other studies have reported the incidence of this peptide in well-fermented cocoa beans (D'Souza et al., 2018; Kochhar et al., 2004; Voigt et al., 2016). Therefore, it cannot be a coincidence that this particular peptide is only seen in well-fermented beans but not in substandard fermentation trials, such as trial B, while it is lost during roasting. In consequence, peptide *PVNSPGKY* might be a good marker for cocoa-specific flavour in fermented beans. Further research will have to corroborate this hypothesis, especially using a good mix of sound and flawed fermentation trials, along with *in vitro* testing.

A comparison of the three trials of forced fermentation suggested that trials A and C had better resembled a spontaneous fermentation. The use of a yeast extract-based supplement therefore might not always lead to well-fermented beans since the supplement seemed to favour growth of yeasts over that of bacteria. This became obvious since the control trial A and the supplemented trial B started with essentially similar if not identical microbiomes. Consequently, one should use a starter culture to better control population dynamics in a laboratory setting for the purpose of consistency rather than relying on a microbial growth supplement or the native assortment of microbes. Nevertheless, one would have to investigate the use of different starter cultures to include growth of LAB as well as to reduce the overall amount of acid production. With the herein established small-scale lab fermentation technique, the path is set to conduct such studies.

4.5. Conclusions

The current study presents an *in vitro* cocoa fermentation method, which drives a 300-g-scale fermentation to well reproduce a large-scale spontaneous fermentation, both inside and outside the bean. Although trial A remarkably displayed characteristics of good fermentation just by controlling the incubation temperature, the use of a starter culture, as done in trial C might represent the most efficient method to ensure reliability. Further trials based on the starter culture and temperature regime of trial C have been conducted, all of which achieved almost identical microbial dynamics and pH and protein profiles (data not shown). These later trials have also highlighted the profound influence of acid and low pH on the bean biochemistry. We therefore conclude that small-scale fermentation trials in laboratories far away from cocoa plantations are

feasible in order to do more in-depth studies of cocoa bean fermentation without the burden of conducting laborious large-scale fermentations. This would eventually open up new avenues of research for cocoa bean fermentation in terms of metabolomics that have not yet been explicitly explored.

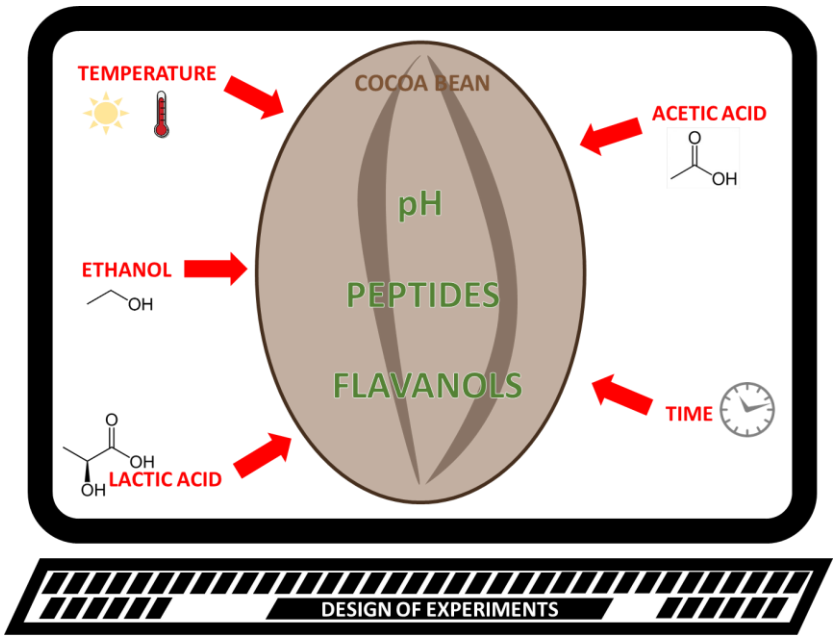
4.6. Acknowledgments

The authors gratefully acknowledge Britta Behrends for her excellent technical support and Marcello Corno for his support in sourcing and sample handling. This work was financially supported by Barry Callebaut.

5. Results – Part III

Modelling Cocoa Bean Fermentation Using Design of Experiments

Warren A. John, Nina L. Böttcher, Britta Behrends, Marcello Corno, Roy N. D'Souza, Nikolai Kuhnert and Matthias S. Ullrich.



Abstract:

The process of cocoa bean fermentation, although it may be vital for the manufacturing of chocolate, is one of the least controlled and predictable processes of fermentation in the food industry. The research presented here deals with using artificial fermentation systems in order to achieve greater control over the external influences that the beans are exposed to, with the aim of modelling changes to pH, peptides, and flavanols within the bean. A design of experiments approach was used, in a first-ever attempt, to study the effects of each factor and their interactions on the mentioned bean components during the incubations. Models with high predictive precision were obtained, showing temperature, incubation time and the concentration of acetic acid to be the key influences over these three responses. Validation experiments proved the model to be robust in predicting changes to these classes of compounds as well as individual compounds within the bean.

5.1. Introduction

The process of microbial fermentation is seen today in many facets, from the diminutive production of methane in cattle rumination to the brawny industrially-applied use of baker's yeast. Whatever the application, fermentation is a vital tool in driving modern day society due to the vast extent to which it is being applied in various industries. Depending on the region, culture and product, fermentation is practiced in a number of ways. Some processes allow fermentation to be more under the influence of nature; while some employ more laborious control.

For example, the production of yoghurt, involving the fermentation of milk by lactic acid bacteria (LAB), encompasses several highly controlled steps, among which are maintaining consistency of the milk, thermalisation, the standardisation of fat content, homogenisation, and the starter cultures and conditions needed for fermentation, to name a few (Sfakianakis & Tzia, 2014). In contrast, the fermentation of cocoa beans, originating from ancient Mesoamerica, is a more laid-back, simplistic process that is practiced differently from one region to another (Saltini et al., 2013). Typically, pods are cut open on the plantations, from which they are harvested, and the beans removed by hand and piled up in heaps. Some farmers cover their heaps in banana leaves while some place the beans in lidded wooden boxes. The heaps are left to stand for a period of three to ten days. It is within these heaps and wooden boxes that fermentation takes place. It is driven naturally and spontaneously, encompassing the successions of yeasts, LAB and acetic acid bacteria (AAB) (Afoakwa et al., 2008). Yeasts begin the process by feeding off the sugar-rich pulp surrounding the beans, and as they consume the available oxygen within the heap, they metabolize anaerobically leading to the formation of ethanol (De Vuyst & Weckx, 2016a). High alcohol content and low oxygen make it difficult for yeasts to thrive and LAB species start to take over and produce lactic acid. Characteristically, the heaps are turned over after two days, which brings oxygen back into the heaps, now boosting the growth of AAB. The AAB start to produce acetic acid from the available ethanol and lactic acid in a considerably exothermic reaction which causes the temperature within the heaps to reach values of 55-60°C (Camu et al., 2007).

The sum of these events occurring on the surface of the beans leads to a metabolomic "metamorphosis" within the bean, essential for forming the basic flavour and aroma pre-cursors for good-quality chocolate. The diffusion of microbial metabolites into the bean coupled with the

changes in temperature result in the breakdown of proteins into peptides and free amino acids that would go on later to form Maillard reaction products (Kongor et al., 2016; Mohr et al., 1976). The formation of specific peptides and amino acids would determine the resulting taste of the chocolate, with diketopiperazines contributing to bitterness (Stark & Hofmann, 2005) and certain other peptides contributing to nutty- or cocoa-specific aroma/flavour (Voigt et al., 2016, 2018). Furthermore, polyphenols are broken down, alleviating bitterness. Our previous findings have shown that the diffusion of acetic acid into the bean is not just responsible for the generation of flavour-specific peptides but also contributes to polyphenol dynamics, influencing taste in terms of bitterness and astringency (John et al., 2019).

Despite the technology of cocoa bean fermentation being rather primitive, the process, by and large, can be viewed as a complex system of interdependent factors, which act externally on the bean, to drive a series of reactions inside the bean, which facilitate the generation of flavour and aroma precursors. The majority of research over the past years has looked into the changing bean metabolome during the course of fermentation as a result of temperature changes and the influx of microbial metabolites (Buamah et al., 1997; Eyamo Evina et al., 2016; Kadow et al., 2015; Leal Jr et al., 2008; Lefeber, Gobert, et al., 2011; Pereira et al., 2012; Romanens et al., 2018; Sandhya et al., 2016). While authors of these findings tried to link specific external factors with changes to specific bean metabolites, there is a still a gap in understanding how these factors affect bean metabolites overall and how these factors interact with each other. Mathematical models based on spontaneous fermentations have also been drawn up in order to study the external environment of the bean and the interplay of microbial successions (Adler et al., 2014; Kresnowati, Gunawan, & Muliyadini, 2015; Moreno-Zambrano, Grimbs, Ullrich, & Hütt, 2018). However, these models address their own limitations of using spontaneous fermentations over which they have almost no control, and the overwhelming amount of variables and variation that need to be taken into account (Moreno-Zambrano et al., 2018).

Biehl and Passern (1982) were the first to utilize submerged artificial fermentation systems, in order to study the effects of acid on bean cotyledons. This research and their subsequent findings opened up new perspectives in studying fermentation under controlled environmental conditions (Biehl, Passern, & Sagemann, 1982; Voigt, Heinrichs, Voigt, & Biehl, 1994). We have successfully employed a similar system in trying to mimic fermentation under completely aseptic conditions, exposing beans to regulated temperature regimes and the main microbial

metabolites generated during fermentation (John et al., 2016). Just as in the case of yoghurt production, these submerged fermentation systems help in achieving fine-tuned control over the conditions that the beans are exposed to, which would otherwise lie in the capricious hands of nature.

Despite achieving control over each factor, simultaneously dealing with multiple factors in experimental design can be a daunting enterprise. Statistical modelling using design of experiments (DoE) has been a favoured approach when it comes to studying multiple factors. DoE, and response surface methodology (RSM), in particular, has been of great industrial benefit in simplifying multi-factorial experimental design, which would otherwise be very cumbersome and cost-ineffective (Montgomery, 2009). Furthermore, DoE has been used extensively in food technology for process optimization, microbiology, various sensory analyses (Granato & de Araújo Calado, 2014), and in cocoa research for the optimization of roasting (Farah et al., 2012).

The following research describes a first-ever approach to experimentally modelling cocoa bean fermentation in a simplistic manner using submerged artificial fermentation systems. DoE was used to assess the individual and combined effects of five factors (temperature, incubation time and concentrations of ethanol, acetic acid and lactic acid) on internal bean pH and various bean metabolites including peptides and polyphenols. The model was validated using a second round of experiments with beans from the same hybrid. We observed notably close model fits and high prediction coefficients for the responses of bean internal pH, diversity as well as quantity of peptides, and total flavan-3-ol (flavanol) content.

5.2. Materials and Methods

5.2.1. Chemicals and reagents

Agar-agar (microbiology grade), calcium carbonate (99%), ethanol (≥ 96%, denatured with 1% MEK), D-glucose monohydrate (≥ 99.5%), isopropanol (Rotisolv®HPLC), lactic acid (90% synthetic grade), MRS-broth (microbiology grade), peptone (from casein, microbiology grade), sodium hydroxide (≥ 99%) and yeast extract (bacteriology grade) were purchased from Carl Roth (Karlsruhe, Germany). Acetic acid (100%, analytical grade), acetonitrile (ACN, LC-MS grade), dichloromethane (stabilized with 20 ppm amylene, analytical grade) and methanol (100%, LC-MS grade) were purchased from Applichem (Darmstadt, Germany). Formic acid (≥ 98% for mass spectrometry) and hesperetin were purchased from Sigma-Aldrich (München, Germany).

5.2.2. Experimental design

Experiments were designed around a mathematical model using the software MODDE Pro v11.0.1 (MKS Umetrics, Umeå, Sweden). The model was designed around five quantitative factors as outlined in **Table 5.1** with orthogonal multiple linear regression (MLR) scaling (Lundstedt et al., 1998). The responses measured were internal bean pH, and the gains in peptide diversity, quantity and flavanol content as compared to the unfermented bean. An RSM model was drawn up with a D-optimal design in order to study the effect of each factor. First- (factor), second- (factor*factor), and third-order (factor*factor*factor) interactions were entered for a comprehensive understanding of factor dynamics (**Table 5.2**). The model with the highest G-efficiency was selected and a list of 47 experiments was generated (Appendix **Table A2**), with the last three experiments being biological replicates in order to account for natural variance. The results for pH were logarithmically transformed with a formula of [$x = 10\text{Log}(1.5y+10)$] and the results for flavanol were linearly transformed with a formula of [$x = 500y$] in order to improve their respective prediction coefficients while all other data remained untransformed.

5.2.3. Artificial fermentation

All equipment used in the processing was sterilized with either 70% ethanol or by autoclaving at 120 °C in order to exclude any microbial contribution to the system. Slightly premature cocoa pods of the Forastero hybrid named “German” were harvested and shipped to Bremen, Germany from Edzendouan in Central Cameroon (GPS coordinates: 4° 17' 33.654" N, 11° 44' 6.292" E). Pods were disinfected by first rinsing with tap water to remove grit and debris, followed by wiping with 70% ethanol and finally with Sterillium® (Bode Chemie GmbH, Hamburg, Germany). To validate sterilization of the system, random swabs from the surfaces of pods were plated on YPD-agar [1% yeast extract, 2% peptone, 2% glucose, 1.5% agar (John et al., 2019)], De Man, Rogosa and Sharpe agar [MRS-agar; 52 g/L MRS broth, 2% agar (Camu et al., 2007)] and glucose-yeast extract-calcium carbonate agar [GYC-agar; 10% glucose, 1% yeast extract, 2% calcium carbonate, 1.5% agar, pH 6.8; (Sharafi et al., 2010)]. Finally, pristine pods were opened up and the beans were de-pulped under high-pressure running water for 15 min. A sample of beans was taken at this point and designated as the “unfermented sample”.

Inside a sterile flow bench, 80-g portions of the de-pulped beans were placed inside a total of 47 sterile 500-mL glass bottles. 80 mL of the respective solutions containing ethanol, acetic

Table 5.1: Factors used in experimental design.
The five main factors known to influence cocoa bean fermentation were plugged into MODDE Pro 11, and designated as quantitative terms.

Factor	Units	Type	Settings
Temperature	**°C**	Quantitative	25 to 60
[Acetic acid]	**g/L**	Quantitative	0 to 25
[Lactic acid]	**g/L**	Quantitative	0 to 25
[Ethanol]	**g/L**	Quantitative	0 to 200
Incubation time	**days**	Multilevel	0.5, 1, 1.5, 2

Table 5.2: Coefficients of factors and their interactions used in experimental design.
The model consisted of first-, second-, and third-order interactions in a d-optimal design. Only the terms that contributed to the values of R2 and Q2 were considered in the model (shaded cells). A colour code is used to represent the factor coefficients for each response (no influence: white; most influential: green). Temp: temperature; HAc: acetic acid; Lac: lactic acid; EtOH: ethanol; Time: incubation time. *denotes significance ($p < 0.05$).

Term	pH	Peptide Diversity	Peptide Intensity	Flavanol Content
Temp	-0.0316*	124.5620*	9.3881*	937.0490*
HAc	-0.0187*	147.8930*	10.9390*	285.5850*
Lac	-0.0122*	22.7425	1.8452*	189.4410
EtOH	-0.0027	-21.2019	-0.6319	82.2264
Time	-0.0047	75.3255*	5.9882*	588.9410*
Temp*Temp	0.0019	-220.4630*	-19.0634*	
HAc*HAc	0.0091*	-74.4359*	-7.1917*	
Lac*Lac	0.0011	-47.6806	-5.2805*	
EtOH*EtOH	0.0018	-32.7188	-3.2403	
Time*Time	0.0078*	-53.3120*		-440.5990*
Temp*HAc		17.3899	1.3311	
Temp*Lac	-0.0114*	-16.7727		
Temp*EtOH	-0.0014	-51.0764*	-4.3025*	206.7610
Temp*Time	0.0022		-0.9297	290.8410*
HAc*Lac	0.0067*	-26.3884	-2.3130	-214.3530
HAc*Time				185.3950
HAc*EtOH				
Lac*EtOH				
Lac*Time				
EtOH*Time	0.0027		1.1659	-256.5940*
Temp*Temp*HAc		-60.1116	-4.9754*	
Temp*Temp*EtOH				
Temp*HAc* Time				
Temp*EtOH*EtOH				
Temp*EtOH*Time	0.0040*		-2.9741*	
Temp*Time*Time	0.0030			
HAc*Time*Time				
EtOH*EtOH*Time	-0.0054			
EtOH*Time*Time				

acid and lactic acid dissolved in MilliQ water in the concentrations dictated by the model, were added to each bottle. The bottles were then sealed shut and placed inside incubators with the corresponding temperature as mentioned in **Table A2**. After the allocated time of incubation, the solution from each bottle was drained off, and the beans rinsed in running deionized water for 5 min before being stored at -20°C.

Small samples of the solutions of experiments N40 and N41 were also plated on YPD-agar, MRS-agar and GYC-agar after incubation in order to check for sterility of the incubations.

5.2.4. Model validation experiments

To validate the prediction accuracy of the model, a set of ten experiments were drawn up for which the model needed to predict the set of factors required to generate the given response profile. Using the optimizer function on MODDE Pro 11, necessary factors were determined for a requested set of responses. The set of factors with the lowest "normalized distance to target" [log(D)] were picked in each case. These factors were then adjusted in terms of temperature such that only five different temperatures would be required. The factors were re-entered into the software and, using the prediction tool, a predicted set of responses was obtained for each of the experiments along with the value of the 95% confidence intervals. The procurement of pods and the subsequent incubation of beans were performed in the same way as mentioned in section 2.3.

5.2.5. Bean sample preparation

Cocoa beans were de-shelled and ground to a fine powder in a Retsch Grindomix GM200 (Haan, Germany). The ground cotyledon samples were stored at -20°C until further processing except for the dry bean samples, which were stored at 4°C. Liquor and chocolate samples were ground at 3,000 rpm for 20 s and stored at 4°C (John et al., 2019).

5-6 g of the ground sample were defatted using dichloromethane with standard Soxhlet extraction for 18 h using a Büchi B-811 system (Essen, Germany) to produce dry defatted powder (DDP)(D'Souza et al., 2018). The DDP was stored at 4°C until further use.

5.2.6. Internal bean pH measurement

2.5 g of ground bean material were shaken head-over-head in 22.5 mL MilliQ water for 5 min to form a suspension. The suspension was centrifuged down at 3,200 *g* for 10 min and the pH of the resulting supernatant was measured using a table-top pH meter (Mettler Toledo GmbH, Giessen, Germany).

5.2.7. Peptide and polyphenol analysis with UHPLC-MS

50 mg of DDP were re-suspended in 5 mL extraction buffer ($MeOH:H_2O:CH_3COOH$/70:28:2) and subjected to ultrasonication for 10 min followed by stirring for 30 min. The polyphenol/peptide (PP) extract was then filtered through 0.45 µm Chromafil® PTFE membrane filters (Macherey-Nagel, Düren, Germany) and spiked with 2 µg/mL hesperetin as an internal standard. PP extracts were processed in the same way as mentioned in D'Souza et al. (2017) with an Agilent 1260 HPLC system using a ZORBAX Eclipse Plus C18 column coupled to an Impact HD ultra-high resolution ESI-Q-q-TOF mass spectrometer (Bruker Daltonics GmbH, Bremen, Germany). 2 µL of the sample were injected for HPLC and eluted with a mobile phase of a binary gradient between MilliQ water and acetonitrile (B), both spiked with 0.05% formic acid, in a scheme of (t (min), %B): (0, 8); (1, 5); (8, 16.5); (9, 17); (10, 17.5);(11, 17.5); (12, 18.5); (13, 18.5); (23, 95); (25, 95), at a constant flow rate of 0.5 mL/min. A column temperature of 40°C was maintained. The electrospray ionization source had a nebulizer pressure of 1.8 bars, a dry gas flow rate of 9 L/min and a dry gas temperature of 200°C. Before each run, the mass spectrometer was calibrated with a solution of sodium formate (0.2% formic acid, 0.1 M NaOH and 50 % isopropanol in MilliQ water) with calibrations being done in HPC mode (DataAnalysis 4.1, Bruker Daltonics GmbH, Bremen, Germany). Data were acquired in both negative- and positive-ion mode with MS^n runs being carried out on a select group of samples (D'Souza et al., 2017). Relative amounts of polyphenols and peptides were obtained by integrating extracted ion chromatograms of matching compounds and were determined in accordance with our previous publications (D'Souza et al., 2018, 2017; Kumari et al., 2018)

5.3. Results and Discussion

Of the many external factors that transform the biochemistry of the cocoa bean during fermentation, the most influential ones are known to be temperature, the concentrations of acetic acid, lactic acid and ethanol, and the duration over which the beans are exposed to these factors (Kadow et al., 2015; Kongor et al., 2016). Thus, these five factors were chosen to model the effect on bean metabolomics in terms of the four main responses of pH, peptide diversity (unique number of different peptides observed), peptide quantity and flavanol content, the latter three being determined by UHPLC-MS. Values for the five factors were chosen using extreme limits reported from experimental data (Camu et al., 2007). Designing the model using limits, including intermediate values, resulted in a set of 47 experiments, generated by the software. The aim was to be able to develop a reliable and robust model despite differences in the quality of the starting material. For this reason, the responses of peptide diversity and quantity as well as total flavanol content were expressed in terms of the change from the original, de-pulped, non-incubated bean.

5.3.1. Model statistics

The accuracy of the model is described by MODDE Pro 11 using four statistical parameters: R2, Q2, model validity and reproducibility. R2 refers to the model fit and a model fit with low significance has values of R2 below 0.5. Q2 refers to the prediction precision which should be above 0.5 for a good model (Consonni, Ballabio, & Todeschini, 2010; Umetrics AB, 2014). A low value for model validity could be due to outliers or transformation problems that contribute to statistically significant model problems. High reproducibility can also result in low model validity (Moldovan et al., 2016). A value for reproducibility above 0.5 denotes biologically and technically reproducible results (Umetrics AB, 2014).

All four responses showed remarkably high model fits and prediction capabilities (**Figure 5.1**). Internal pH showed the highest values followed closely by peptide diversity and peptide quantity. The lowest values of R2 = 0.84 and Q2 = 0.696 were observed for total flavanol gain, which, nevertheless, can still be considered sound values. The values for reproducibility were excellent with relative standard deviation (RSD) values of biological replicates similar to analytical replicate RSD values. Considering model validity, it was observed that values for peptide diversity and quantity as well as that for flavanol content were all above 0.25, whereas that of pH was 0.1. The low model validity in the case of pH could be attributed to the almost

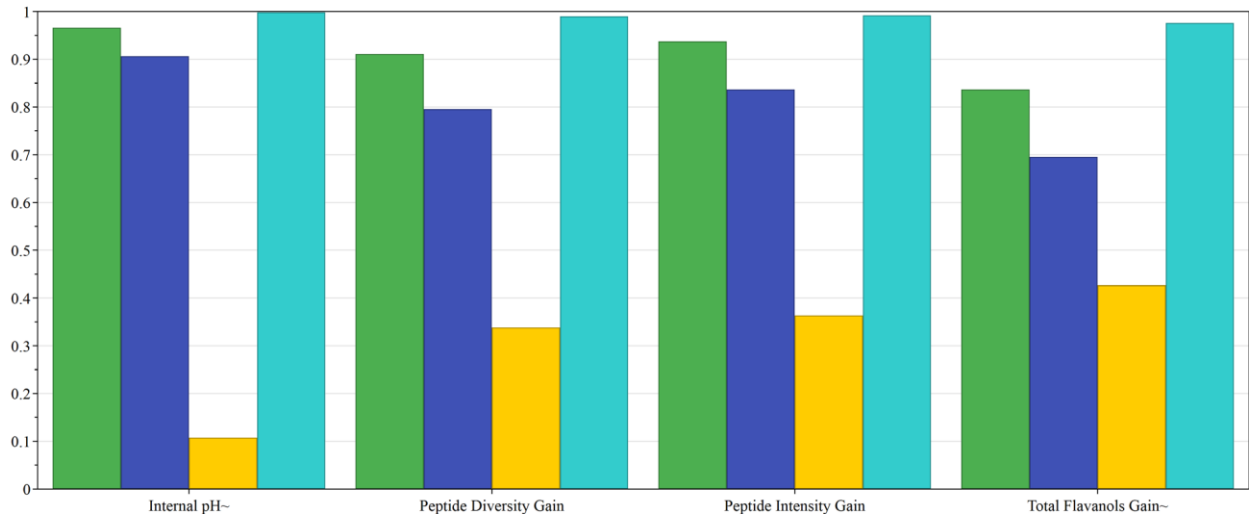


Figure 5.1: Model statistics for internal pH and gains in flavanol content and peptide diversity and quantity. Values of R2 (green), Q2 (dark blue), model validity (yellow) and reproducibility (light blue) are shown for the four responses of internal pH, and gains in peptide diversity, peptide intensity and flavanol content as modelled by MODDE Pro. Five factors and their interactions were used in the model in a d-optimal design.

perfect reproducibility of the three biological replicates and therefore still be considered a good model. The model statistics for these areas of model fit, prediction precision and reproducibility demonstrated the outstanding reliability and accuracy of the entire procedure from bean incubation to analysis via UHPLC-MS.

5.3.2. Factor influence on responses

When the model was constructed, many factors and their interactions were taken into account (**Table 5.2**). From all possible factor interactions, only a handful contributed significantly to the model fit and prediction precision. A significant factor in the model contained error bars that did not cross zero on its coefficient plot (**Figures 5.2A, 5.2G, 5.3A** and **5.3G**). This corresponded to a p-value of less than 0.05 (Umetrics AB, 2014).

Notably, pH had the most factors contributing to it with a total of 18 terms and eight of them being statistically relevant as seen on the factor coefficient plot (**Figure 5.2A**). Gain in total flavanol content had the least number of contributing factors, and out of the total 11 factors, six showed significance (**Figure 5.2G**). Gains for peptide quantity and diversity had a total of 16 and 15 contributing factors, respectively, and of those, seven and 10, respectively, were statistically relevant (**Figure 5.3A** and **5.3G**). Taking into consideration both, factor coefficients and model statistics, in the case of pH, it stands to reason that the herein considered factors were

highly influential within the model. In contrast, in the case of gain in flavanol content, there were missing pieces to the puzzle: certain other factors apart from the five analysed herein. The reactions that underlie the dynamics of polyphenol composition are complex, to say the least. Genetics, environmental stress factors, and even differing matrices within the bean hybrids can lead to polyphenols being formed and processed in different ways (Cheynier, 2005). As our understanding of the factors that govern polyphenols dynamics grows, the possibility of developing better models might increase. Nevertheless, the factors used in this study are the five main factors that come into play during fermentation and one already gets a glimpse of how significantly they are able to influence polyphenol content within the bean.

5.3.2.1. Influence of temperature

For the responses of internal bean pH and total flavanol gain, temperature was able to linearly influence both responses, negatively in the case of pH (**Figure 5.2B**) and positively for flavanol gain (**Figure 5.2H**). Temperature has a direct and linear influence of the diffusion of acetic and lactic acids into the bean which causes a drop in pH (Jinap & Danker, 1993). Consequently, the diffusion of acid into the bean also leads to an increase in flavanol content as shown by our previous findings (John et al., 2019). Moreover, temperature influenced gains in peptide diversity and quantity non-linearly (**Figures 5.3B** and **5.3H**). This could be due to its effect on enzymatic activity. As temperature passes the optimal range for enzymatic activity, denaturation might occur, accompanied by loss of activity, causing a drop in rate of proteolysis and a reduction or modification of peptide formation. One could deduce that the optimal temperature for cocoa bean endoprotease activity is between 46°C and 48°C (**Figures 5.3B** and **5.3H**), which has also been observed in previous findings (Guilloteau et al., 2005).

5.3.2.2. Influence of incubation time

The duration of incubation time had non-linear effects on the gains in total flavanol content and peptide diversity (**Figures 5.2I** and **5.3C**). The effect on presence of peptides could as well be linked to enzymatic activity, which breaks down proteins into peptides initially, but when allowed to continue, breaks down peptides into amino acids and thereby decreasing peptide diversity. Curiously, internal bean pH was also non-linearly correlated to incubation time with an initial decrease followed by an increase in pH. A plausible explanation for this phenomenon

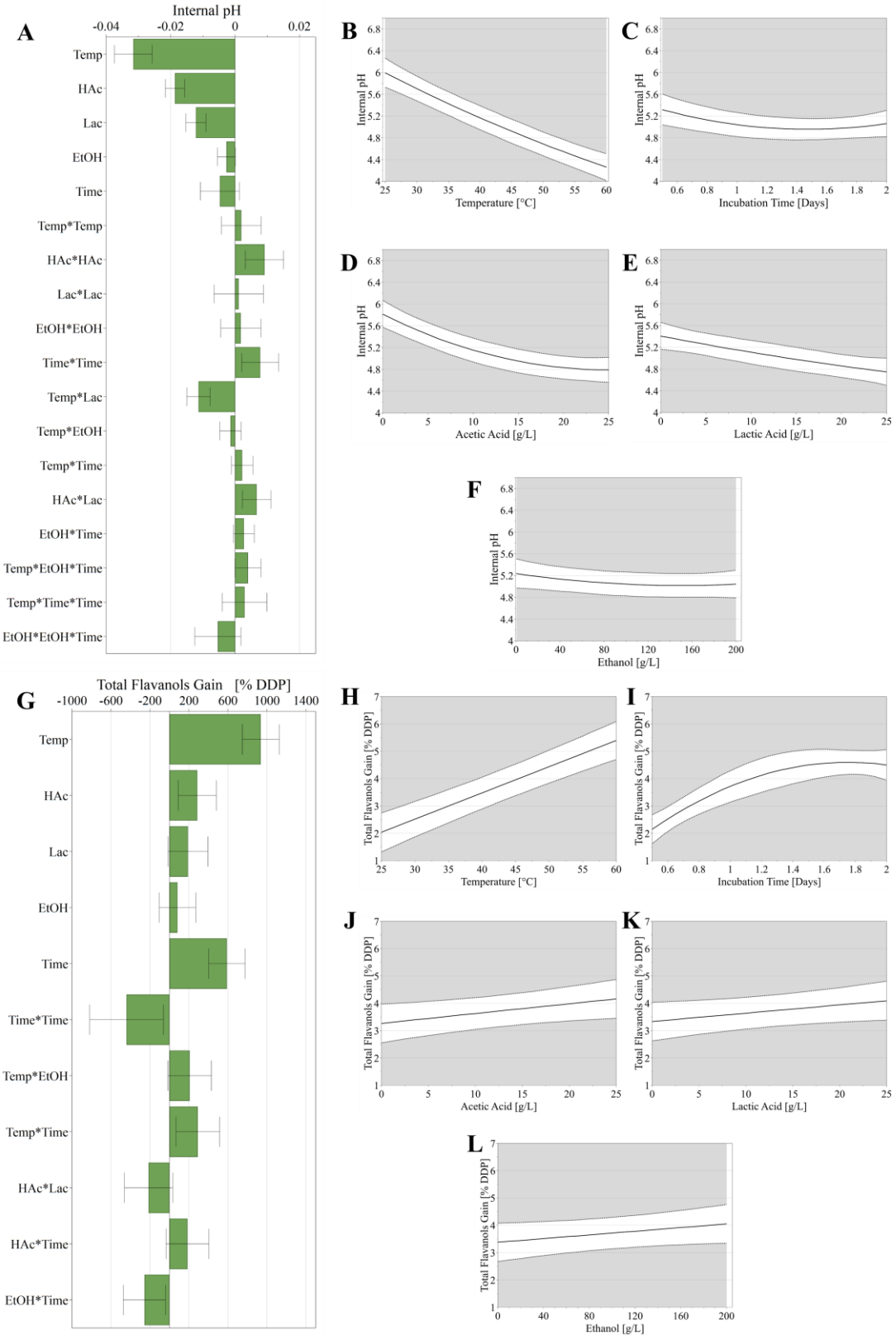


Figure 5.2: Factor effects over pH and gain in flavanol content.
Coefficients of each factor and their interaction terms for internal bean pH (A) and the gain in flavanol content (G). A term containing error bars that do not cross zero is considered statistically significant to the model. The effects of each factor on the two responses are also graphically displayed (B-F, H-L). The confidence intervals for the factor effects are set at 95%.

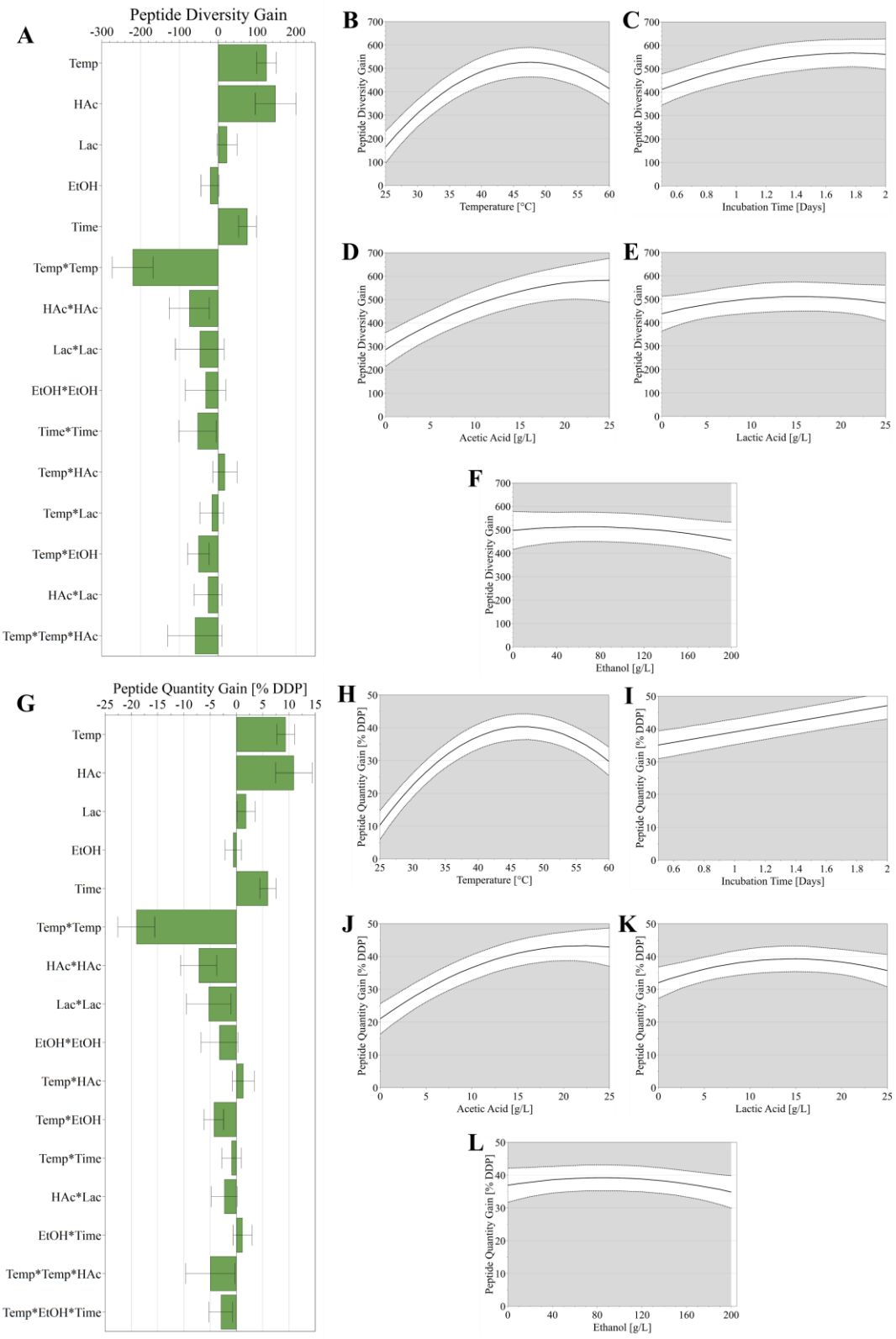


Figure 5.3: Factor effects over gains in peptide diversity and quantity.
Coefficients of each factor and their interaction terms for the gains in peptide diversity (A) and quantity (G). A term containing error bars that do not cross zero is considered statistically significant to the model. The effects of each factor on the two responses are also graphically displayed (B-F, H-L). The confidence intervals for the factor effects are set at 95%.

would be the occurrence of an initial diffusion of citric acid from the residual pulp surrounding the bean that is later countered by the formation of peptides and amino acids. Since amino acids have an average isoelectric point (pI) of 6.03 at 25°C (Damodaran, Parkin, & Fennema, 2008) and each amino tends to exist in its zwitterion state, there would be a consumption of protons, thereby increasing pH. The gradual increase of peptide quantity, associated with the overall increase in free N-terminal amino group concentration and the plateauing of peptide diversity after 1.5 days corroborate this (**Figure 5.3I**).

5.3.2.3. Influence of acetic acid and lactic acid

The concentration of acetic acid had a linear effect on flavanol gain (**Figure 5.2J**) which could be explained as in the case of temperature. Intriguingly, the acetic acid concentration had a significant non-linear influence on internal bean pH (**Figure 5.2D**). Logically, the gains in peptide diversity and quantity were affected in a similar manner (**Figures 5.3D** and **5.3J**). However, this was not observed in the case of lactic acid, which showed a linear relationship to pH (**Figure 5.2E**). The difference in the behaviour of these two acids could potentially be attributed to their buffering capacities. Research on the buffering capacity of wort in beer brewing indicated that an increasing concentration of acetic acid increases the buffering capacity of wort to a greater extent as compared to that of lactic acid (Li, Liu, Kang, & Zheng, 2016). This could also be the case for the cocoa bean matrix, which contains a substantial amount of citric acid. Due to the higher increase in buffering capacity with acetic acid, the relationship changes from linear to non-linear with an increasing concentration of acetic acid. This occurrence is not as pronounced in the case of lactic acid.

The only significant influences of lactic acid were those being direct and linear on pH and non-linear on peptide quantity gain (**Figures 5.2E** and **5.3K**), most likely because low pH leads to increased endoprotease activity (Guilloteau et al., 2005). There was no significant influence of lactic acid on the gain in flavanol content (**Figure 5.2G**) although this interaction showed a linear tendency (**Figure 5.2K**). Despite a significant presence of lactic acid during spontaneous cocoa bean fermentation, it was reported to result in undesired flavour in chocolate and is therefore sought to be kept to a minimum (Holm, Aston, & Douglas, 1993). Further research also corroborates our findings in suggesting that lactic acid bacteria and the subsequent production of lactic acid play a much lesser role in modulating pH, peptides and flavanols, compared to acetic

acid (Ho et al., 2015). Even though, lactic acid itself may seemingly not have a profound influence on cocoa biochemistry and flavour, one cannot just simply argue away the role of lactic acid bacteria. They are known to aid the consumption of citric acid and the reduction of fructose to mannitol (De Vuyst & Weckx, 2016a), the effects of which can cascade into other events contributing to cocoa flavour.

5.3.2.4. Influence of ethanol

Interestingly, ethanol did not directly influence any of the four responses. This had been demonstrated in another experimental study before (De Vuyst & Weckx, 2016a). It has no significant effect on pH (**Figure 5.2A**) while it affects gains in peptide diversity and quantity together with temperature but never by itself. For peptide quantity, ethanol has a significant third factor interaction with temperature and time (**Figures 5.3A** and **5.3G**). For gain in flavanols, ethanol is only able to influence it together with time (**Figure 5.2G**). All these indirect effects of ethanol negatively impact gains in peptide and flavanol content, hinting towards a temperature- and incubation time-dependent leaching effect of ethanol, which has been observed previously (John et al., 2019).

5.3.2.5. Factor Interactions

Overall, using this model, one is able to better understand how the analysed factors come into play when affecting pH, peptides and flavanols within the bean. For example, the contrasting relationship of acetic acid and incubation time, respectively, on flavanol gain are seen clearly in the response surface plot (**Figure 5.4A**). In essence, incubation time is linearly correlated with gain in flavanols while acetic acid has a non-linear correlation. Both, temperature and acetic acid concentration have linear correlations with flavanol gain (**Figure 5.4B**). This lies in stark contrast to the corresponding effects on the gain in peptide quantity. The maximum gain in peptide quantity is achieved at a specific range of the two factors, and the relationship results in a better fit for a quadratic model (**Figure 5.4C**). This is even more pronounced for the effect of temperature and lactic acid, where there is a so-called “sweet spot” of maximal peptide quantity gain (**Figure 5.4D**). The latter two figures also show that maximum peptide diversity gain is more easily achieved in a broader temperature range through the use of acetic acid as compared to the use of lactic acid. Further scrutiny of the factor effects emphasized that

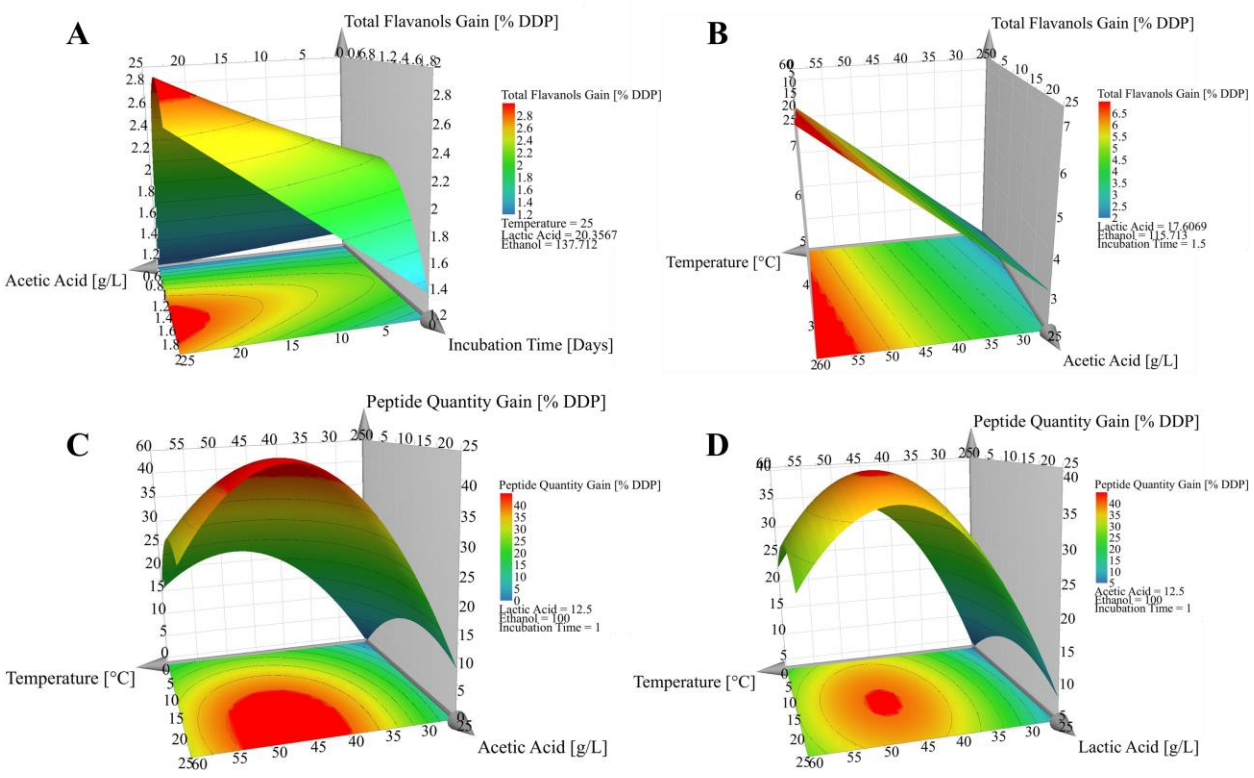


Figure 5.4: Factor interactions upon gains in peptide quantity and flavanol content.
Response of gain in flavanol content with respect to time and [acetic acid] (A), and with respect to temperature and [acetic acid] (B). Response of gain in peptide quantity with respect to temperature and [acetic acid] (C), and with respect to temperature and [lactic acid] (D). The values for the other three factors in each case are kept constant.

temperature, incubation time and concentration of acetic acid were the key players in influencing the model and had significant interaction effects with one another as well as with time and lactic acid concentration (**Figures 5.2** and **5.3, Table 5.2**).

5.3.3. Model validation

A statistically valid model was generated from five quantitative factors and their interaction terms, comprising a set of 47 experiments. Equally important as establishing the model is its validation. Ten validation experiments were drawn up, where a set of desired response profiles was formulated for each (**Table 5.3**). These validation experiments took into consideration extremes of each of the responses as well as mid-level values, to be distributed broadly within the observed range of response values. A set of desired responses was entered into MODDE Pro for which the requisite factors were outlined by the optimizer tool. The factors

Table 5.3: Experiments for model validation.
A set of 10 experiments were formulated where a given set of responses were entered into the software and using the optimizer function, the necessary factors were predicted. The factors were then adjusted such that only five different temperatures could be used. The responses were then predicted for those adjusted factors using the prediction tool in order to get. Max: the response was set to be as high as possible; Min: the response was set to be as low as possible.

Expt.	Requested Response			Predicted Factor					Adjusted Factor					Predicted Response		
	pH	Peptide Quantity Gain (% DDP)	Flavanol Gain (% DDP)	Time (h)	Temperature (°C)	Ethanol (g/L)	Acetic acid (g/L)	Lactic acid (g/L)	Time (h)	Temperature (°C)	Ethanol (g/L)	Acetic acid (g/L)	Lactic acid (g/L)	pH	Peptide Quantity Gain (% DDP)	Flavanol Gain (% DDP)
1	5.0	Max	2.0	12.7	48	29	22.0	13.8	12.7	48	29	22.0	13.8	5.00	40.69	2.28
2	Min	Max	Max	48.0	54.5	54	25.0	20.0	48.0	54	54	25.0	20.0	4.15	45.99	7.06
3	Max	Min	Min	16.0	25	0	0.0	0	16.0	25	0	0	0	7.46	0.00	0.10
4	5.0	Max	2.0	48.0	25	200	22.0	2.8	48.0	25	200	22.0	2.8	5.34	20.27	1.96
5	5.0	Max	7.0	48.0	48	81	25.0	0	48.0	48	81	25.0	0	4.84	46.96	6.31
6	5.5	30	Min	12.0	36	60	24.7	6.9	12.0	35	60	24.7	6.9	5.51	28.68	1.60
7	6.0	15	3.0	19.2	33	187	2.1	23.8	19.2	35	187	2.1	23.8	5.87	13.34	3.16
8	5.5	10	7.0	47.0	60	200	0.7	5.6	47.0	60	200	0.7	5.6	5.50	9.22	5.74
9	5.5	17	3.8	16.3	53	190	4.0	5.0	16.3	54	190	4.0	5.0	5.46	15.92	3.89
10	5.5	17	3.8	16.3	53	190	4.0	5.0	16.3	54	190	4.0	5.0	5.46	15.92	3.89

were altered in terms of their temperature values such that only five different temperatures would be used. Subsequently, the prediction tool was used to accurately predict the responses for the given factors. Only the responses of pH, peptide quantity gain and total flavanol gain were considered since the gains in peptide quantity and diversity were found to be closely related.

The observed results for each response were plotted together with the predicted values within their 95% confidence intervals (**Figure 5.5**). Of all the three responses, pH showed the best prediction accuracy as evidenced by the observed values (spots) being close to or within the predicted range with 95% confidence (bars). Gain in peptide diversity had the next best prediction accuracy where only experiments 2 and 7 showed observed values outside the predicted range. Gain in flavanol content was the least predictable response where the observed values for experiments 3, 6, 7 and 9 lied outside their predicted ranges. Model statistics had already revealed pH to be the most predictable response followed by gains in peptide quantity, peptide diversity and flavanol content in their sequential order (**Figure 5.1**). Despite flavanol content being the least predictable, it still had a considerably good prediction precision as also reflected by the validation experiments. The lack of prediction capability in this instance could stem from other factors that may influence the gain

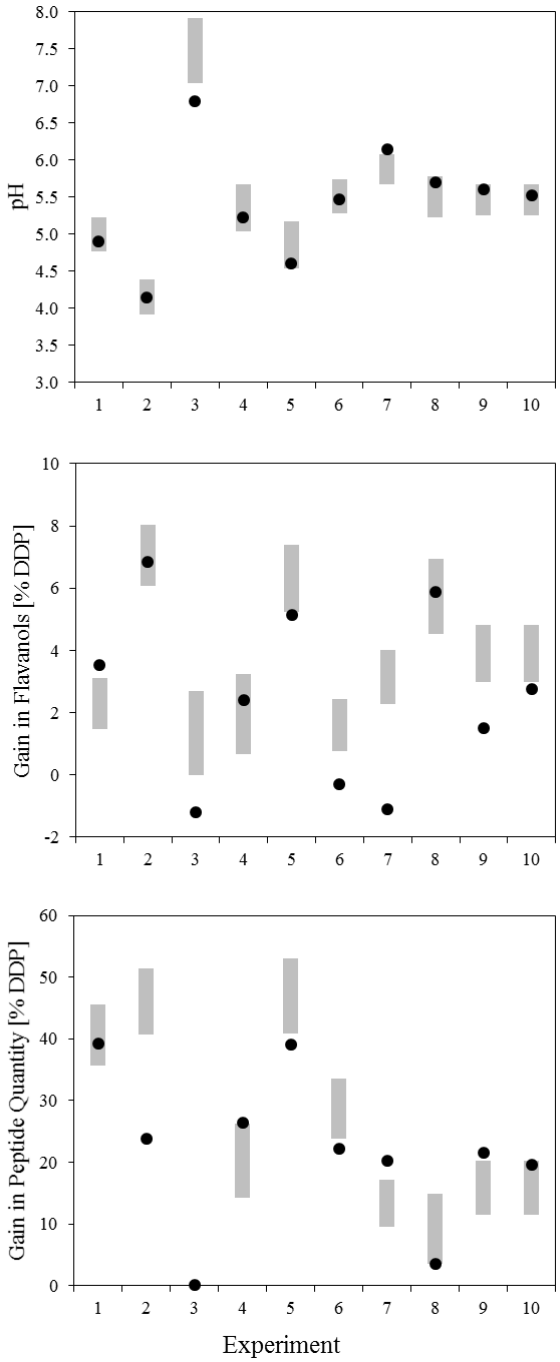


Figure 5.5: Model prediction accuracy.
The model was validated using beans of the same hybrid to see if the observed values (spots) for each experiment lined up with the values predicted by the model within their 95% confidence intervals (bars). This was done for the responses of pH, and gains in peptide intensity and flavanol content. A spot lying within the grey bars would represent a successful prediction.

in flavanols, as discussed in section 5.3.2 with regard to factor coefficients. Furthermore there is also speculation that higher oligomers of catechins with low solubility or high limits of detection in MS experiments may affect and consequently artificially overshadow the observed flavonoid concentration (Wollgast & Anklam, 2000).

The question remains as to what other prominent factors have not been taken into consideration in this model. Research has shown that factors such as oxygen tension, pulp drainage in form of sweating, pectinolysis and the production of other microbial secondary metabolites, to name a few, can affect microbial growth on the surface and subsequently have an impact on the bean's biochemistry (De Vuyst & Weckx, 2016a). A model incorporating every possible factor, however, would be immensely complex, and achieving such accuracy and reliability with biological models can prove to be a formidable task. The simplistic model presented here acts as a starting point to such research and could inspire further research into developing more intricate models of cocoa bean fermentation and subsequent processes of chocolate production. Furthermore, it must be highlighted that the experiments presented here have dealt with only one bean hybrid. One could only speculate as to how much of an influence the genotype of the hybrid has on each of these responses.

5.3.4. Model expansion

Having established that the model works accurately for peptides and flavanols as whole classes, we took it one step further in being able to model individual components of each class of compound. Val-Phe (VF) is a dipeptide representing the precursor sequence of a key diketopiperazine responsible for bitterness in cocoa (Stark & Hofmann, 2005) while the penta-peptide NNPYY has been found to be associated with cocoa flavour (John et al., 2019; Voigt et al., 2018). From the repertoire of peptides detected by the UHPLC-MS method (D'Souza et al., 2018), the amounts of these two peptides were investigated by the herein used model as responses. Epicatechin, which is the major constituent of cocoa polyphenols (Kim & Keeney, 1984), along with clovamide and hydroxyjasmonic sulphate (HJS), known to be important bioactive compounds (Arlorio et al., 2008; Gidda et al., 2003; Patras, Milev, Vrancken, & Kuhnert, 2014) were also studied. Relatively good model fits were observed for all five of these responses, with all values for Q2 and R2 being above 0.5 (**Figure 5.6**). Responses such as loss in HJS, gain in clovamide and gain in VF showed higher values for R2 and Q2 as compared to the

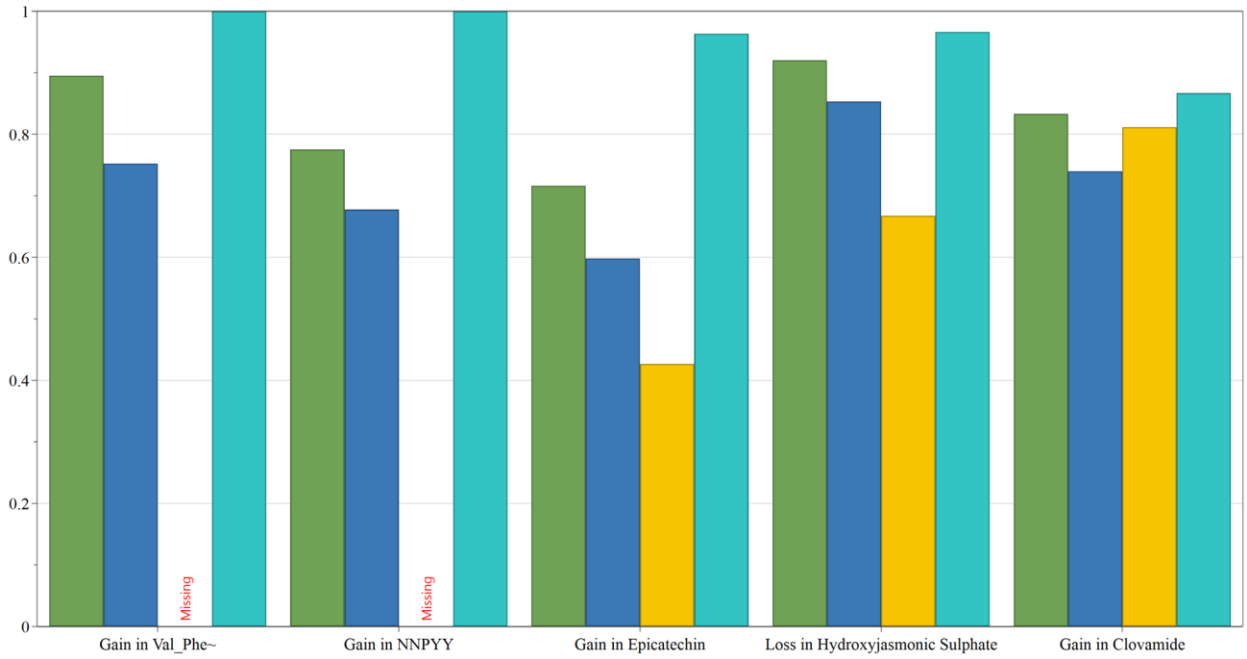


Figure 5.6: Model statistics for internal pH and gains in flavanol content and peptide diversity and quantity. Values of R2 (green), Q2 (dark blue), model validity (yellow) and reproducibility (light blue) are shown for gains of the peptides Val-Phe and NNPYY, loss in hydroxyjasmonic sulphate and the gains in epicatechin and clovamide, as modelled by MODDE Pro. Five factors and their interactions were used in the model in a d-optimal design.

other two responses, epicatechin and NNPYY. The response of gain in epicatechin predictably showed the lowest values of the five responses, since the response of total flavanol content also showed the lowest values of R2 and Q2 initially (**Figure 5.1**). The missing model validities in the responses of both peptides are due to their exceptional absence in all three biological replicates. Moreover, a study of the factor effects also revealed intriguing correlations and factor interactions, reinforcing the point that the effects of these factors are not as simple and straightforward as once thought (**Figure 5.7**).

It is suffice to say that future possibilities with this model and experimental setup are numerous. Our experiments have aptly illustrated that modelling changes to metabolomic components of cocoa beans under controlled conditions is indeed possible. One is able to measure the specific effect of each factor on individual markers, and thereby have some idea of what to tweak in a full-scale spontaneous fermentation to achieve a desired product. In reverse, one is able to predict the effects of a fermentation by measuring specific responses. Additionally, it would be interesting to find out the effects on other aspects of the metabolome such as free amino acids and sugar components, which also contribute extensively to the formation of cocoa

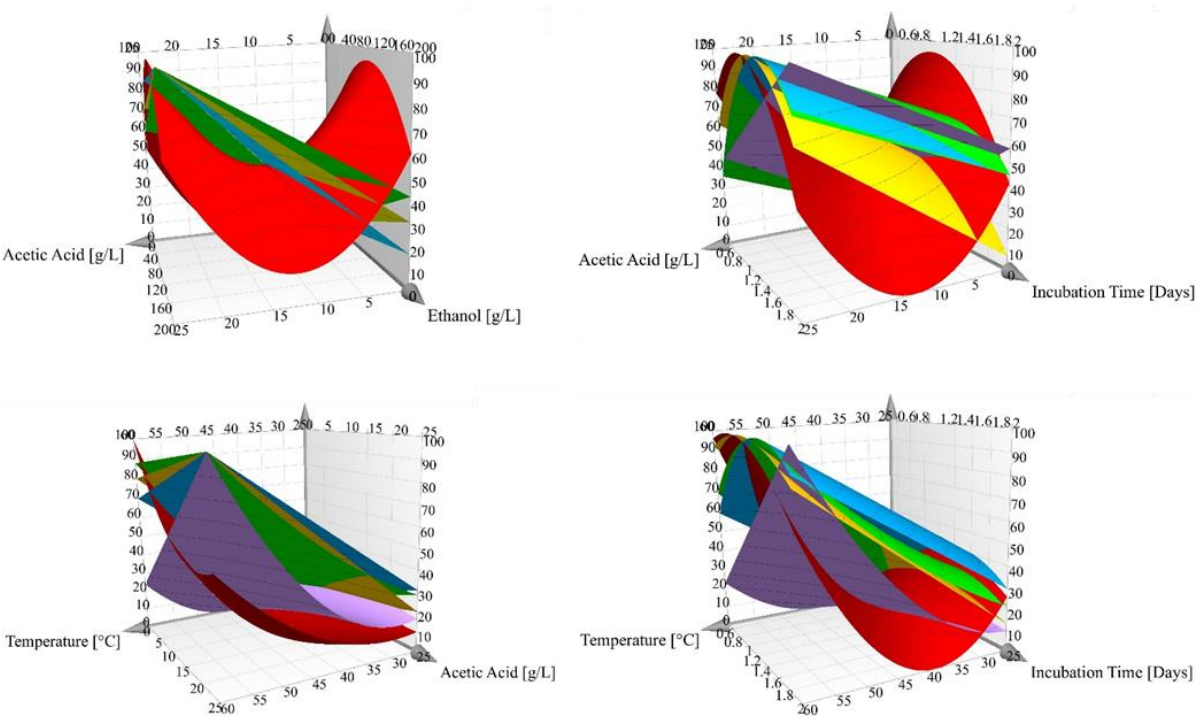


Figure 5.7: Factor interactions upon individual markers as identified and quantified by UHPLC-MS.
The response surface plots reveal the behaviour of responses of loss in **hydroxyjasmonic sulphate**, gains in **clovamide** and **epicatechin**, and the gains in peptides **NNPYY** and **VF**, expressed in percentage scale (y-axis), to the combination of factors being looked at in each diagram.

flavour and aroma. The experimental setup also shows promise for the improvement of poor fermentations using artificial systems. Being able to influence the biochemistry of the bean in a controlled manner could bring in a lot more reliability – and consequently more sustainability – to the industrial chain of chocolate production.

The exploitation of UHPLC-MS, however, would be too cumbersome and expensive a method to use on plantations. However, as research progresses and technology grows, modest and effortless methods could also be developed to measure the profiles of metabolites during the course of spontaneous fermentation in order to make accurate predictions to the resulting bean profile. Building on this, the responses of bean biochemistry could be linked to flavour and aroma of the resulting chocolate. As of now, research into roasting experiments using DoE has only reckoned with process parameters (Farah et al., 2012; Rocha et al., 2017). But, if made to additionally incorporate the starting profiles of flavour and aroma precursors, it could prove to be imperative in revolutionizing the manufacture of chocolate content.

5.4. Conclusions

The presented experiments using DoE have demonstrated that the intricate mechanisms within the cocoa bean can be seen to respond in a predictable manner. The artificial fermentation system was able to successfully guide the external influences that affect the bean in a controlled fashion such that a robust model of the responses in terms of pH, peptides and flavanols could be constructed. Temperature, acetic acid and incubation time had the greatest effects of the five factors and exhibited linear and non-linear effects. Significant interactions between factors were observed for each response, highlighting their collective action in modulating cocoa biochemistry.

5.5. Acknowledgments

The authors gratefully acknowledge Mauricio Moreno-Zambrano and Anne Grimbs for their sound advice.

6. General Discussion and Future Scope

6.1. Factors Affecting Fermentation

Previous studies have already revealed that proteins and polyphenols are affected dramatically during the course of fermentation. Almost all proteins are broken down into peptides and free amino acids by proteolytic activity (Kumari et al., 2016), and polyphenol content (flavanols being the major class) is also reported to decrease during fermentation (Wollgast & Anklam, 2000). The question remains as to what causes these changes during the process. Various reasons have been addressed, such as leaching into the sweating, acidification, high temperatures, death of the embryo, *etc*. The experiments carried out describe the influence of each of these factors over the profiles of polyphenols and proteins, and subsequently, peptide formation.

Chapter 3 illustrated that ethanol, acetic acid and lactic acid diffuse into the bean to a considerable extent. Furthermore, one can indeed influence protein content and proteolysis using just ethanol, acetic acid and lactic acid together with temperature. A control incubation with water revealed that these profiles were not affected to the same extent (data not shown) and the control trial with low concentrations of these components showed that they quantitatively affect proteolysis (**Figure 3.2C**). A more detailed investigation in the chapter 5 made clear the differences between each of these factors. Evidently, temperature and acetic acid had much higher direct influences over proteolysis and polyphenol content, compared to ethanol and lactic acid. Ethanol, curiously, showed indirect influences over bean biochemistry suggesting that while, it may not affect these profiles by itself, it has the propensity to affect other factors such as acetic acid and lactic acid formation, which play a more direct role. This is explained by the conversion of ethanol to lactic and acetic acids by LAB and AAB (De Vuyst & Weckx, 2016a, 2016b).

The work outlined in chapter 5 also investigated the effect of time on the bean. In each of the four responses studied, time had either linear or non-linear effects. This result demonstrated that beans left to stand already undergo certain changes, most likely a result of the starting germination. The reactions triggered by germination would cause proteolysis and other events that would provide nutrients from the storage material in the cotyledons for the embryo to grow (Müntz, 1996; Müntz, Belozersky, Dunaevsky, Schlereth, & Tiedemann, 2001). Consequently,

this has implications for those cocoa-producing parts of the world, which store harvested pods for a period prior to fermentation. Scrutiny of these experiments suggested that time had a positive influence on peptide formation and, hence, a negative influence on protein content, which has also been observed before in Ghanaian pod storage prior to spontaneous fermentation (Afoakwa et al., 2014). In contrast, bean pH was seen to linearly increase during pod storage in the research carried out by Afoakwa et al. (2014), whereas the results of chapter 5 pointed to a quadratic effect of time on pH. Likewise, Afoakwa et al. (2013) also found decreasing polyphenol levels, measured by a Folin-Ciocalteu assay, whereas the herein conducted experiments showed a slight increase in polyphenols during the first two days, consistent with the observations in chapters 3 and 4. The results in this thesis, however, were only obtained over a period of two days whereas those of Afoakwa et al. (2014) were measured over roughly 20 days and therefore cannot be effectively compared due to the varied resolution of the data. Nevertheless, pod storage is typically carried out for only a few days and the evidence presented showed that the duration of pod storage is critical to determining the starting values of proteins and polyphenols before fermentation. This would be important in, for instance, reducing astringency and increasing cocoa aroma by reducing protein content as much as possible, or, preserving polyphenols to enhance the health benefits of the resulting chocolate.

Moreover, there was a considerable amount of interplay between the factors, in the sense that two or more factors at a time were able to influence a response. A good example of this was temperature, ethanol and time together influencing the gain in peptide diversity (**Figure 5.3G**), revealing an interdependence between each factor. This observation emphasized the fact that the influence of each factor in fermentation is not as straightforward as one would assume, which makes cocoa bean fermentation an industrially complex process. Therefore, the establishment of a regulated setup of fermentation, as in the cases of beer or yoghurt production, would be a challenging undertaking.

Notwithstanding, the elimination of microbial activity in the Artificial Fermentation system shows promise in bringing predictability into the process by removing the uncertainty of microbial growth and dynamics. Through the submerged incubation of beans, one could achieve an industrial level of control over fermentation that would result in consistent and desired batches of “fermented” beans to be used in producing specific kinds of chocolate. Even though, it may seem a distant possibility, these experiments have proven that it is a distinct one, and that

predictability can be achieved in the process of fermentation. The research described in chapter 5 is, as of now, the pinnacle in the industrialization of fermentation, but has only dealt in depth with five factors, the responses of pH, peptides and flavanols, and only employed UHPLC-MS as the method of analysis. Further research will still have to be laboriously carried out in terms of fine-tuning each of the factors through an incorporation of sensory analysis, the study of other factors and buffer constituents such as esters and higher alcohols that impact fine-flavour characteristics, and the expansion of analytical methods to study other components of the bean and detect specific markers.

In conclusion, hypothesis statement 1 (section 3.2) is not fully validated and neither fully nullified. The three microbial metabolites were seen to diffuse into the bean and instigate change within it. However, a decreasing order of influence of temperature, acetic acid, ethanol, and finally, lactic acid was observed upon proteins, peptides and polyphenols, as well as significant interaction between the factors.

6.2. Simulating Fermentation

Having now established that one can explicitly influence proteins and polyphenol content using the three microbial metabolites with temperature incubations, the question remained as to whether one could replicate their respective profiles as observed during spontaneous fermentation. The experiments with Artificial Fermentation outlined in chapter 3 exquisitely demonstrated the simulation of proteolysis using these factors. Using literature data to match the concentrations of each microbial metabolite, and replicate the temperature regime that usually accompanies fermentation (Trial 1), complete protein degradation was observed, along with an even more diverse, albeit matching, peptide repertoire of a commercial fermentation. All this was accomplished only within three days of incubation.

The profiles of polyphenols were not investigated in depth in this research. Nevertheless, a Folin-Ciocalteu assay of the samples of Trial 1 revealed there to be a doubling of phenolic content within the bean during those three days of incubation (data not shown). While it is generally supposed that polyphenol content decreases during the course of fermentation, a few reports have pointed to increases at specific periods (Albertini et al., 2015; D'Souza et al., 2017). This phenomenon was also observed in trials of the Forced Fermentation system with the native pod microbiota (Trial A) and with an added starter culture (Trial C) in the chapter 4. All aspects

considered, these occurrences are too frequent to be deemed random errors or the result of sample mishandling. It can thus be deduced that there are other mechanisms at play that lead to elevated levels of polyphenols. Speculation on this matter hints toward plant defence responses as a result of acid influx, to protect an embryo that is yet alive (Bayliak et al., 2016) or the breakdown of long, undetectable, insoluble oligomers during fermentation (Wollgast & Anklam, 2000), among other possibilities.

The results of the Forced Fermentation system, using updated analytical methodology, also showed a substantial breakdown of proteins and the generation of a massive peptide repertoire, especially with the native pod microbiota (Trial A) and the starter culture (Trial C). This was one of the most diverse repertoires of peptides ever observed among published studies. Bean pH, which is an important indicator of fermentation also dropped to a considerable extent, but in the trial with the starter culture, may have dropped too far, most likely the result of an over-production of acetic acid. On the other hand, Trial A was exceptional in replicating the dynamics of yeast and AAB, pH and proteolysis, and, to a good extent, flavanol dynamics. This was especially surprising given that this trial only used the native pod microorganisms to accomplish this feat. This observation proved that pods are not 100% sterile, as studies have previously disputed (Schwan 1998; Schwan and Wheals 2004). Despite the absence of LAB in this instance, this observation could just as well hint to the sufficiency of native pod microbes in fermentation. Furthermore, employing strains of *S. cerevisiae* and *A. pasteurianus*, was also effective in replicating the profiles of pH, proteins, peptides and polyphenols. None of the three trials of Forced Fermentation saw growth of LAB, which also supports the notion that LAB play a minor role in the formation of flavour and aroma precursors (Ho et al., 2015) and encourages research in the pursuit of developing starter cultures using only strains of yeast and AAB. Moreover, the employed temperature regime, was able to efficiently boost the necessary microbial activity, with a succession of yeasts followed by AAB.

Despite the fermentation of beans with native pod microbes leading to better tasting chocolate and more accurate representations of fermentation, further experiments using this setup showed that the reproducibility of microbial dynamics is not that good (data not shown). The use of a starter culture, showed greater promise due to the amount of control and consistency achieved, and these results were corroborated by further trials using the same starter culture. Nevertheless, the taste profiles of the resulting liquors of Trials A and C, despite showing good

tendencies for chocolate production, were still far from the desired profiles. Hence, there is a still a need for refinement of the starter culture to more accurately represent changes in bean biochemistry during spontaneous fermentation to yield good-flavour chocolate.

From the results of chapter 5, one could gather that temperature and acetic acid were the most influential over proteins and polyphenols. Therefore, in order to guide a spontaneous fermentation, one must account for the temperature generated by microbial activity as well as production of acetic acid, which not only affects proteins and polyphenols within the bean but has a direct influence of the acidity of the produced liquor. Temperature is affected by all microbial activity on the bean and AAB are known to be the most efficacious in modulating temperature (Camu et al., 2007). This was also observed in the Forced Fermentation system of chapter 4, where the difference between the bean pile temperature and the incubation temperature was higher when AAB started to grow (data not shown). The formation of acetic acid is also governed by a number of factors. The growth of yeast, their subsequent production of ethanol, and the growth of AAB in terms of their numbers and species specificity, all taken together, would determine the production of acetic acid on the pulp of the bean. Thus, it is essential that one ensures the correct microorganismal growth during fermentation. The selection of strains of yeasts determines the amount of ethanol available for production of acetic acid. *S. cerevisiae* is known to have high ethanol productivity and relatively high ethanol tolerance and is a favoured candidate in the selection of yeast strains. Nevertheless, an over-production of ethanol is disadvantageous to microbial growth (observed in the trial with microbial supplement of the Forced Fermentation system), and may also lead to an over-production of acetic acid (as in the case of the trial with starter culture). In selecting strains of AAB, *Acetobacter* spp. are generally preferred due to their efficient production of acetic acid and a low risk of forming off-flavour compounds, in contrast to *Gluconobacter* spp. All these factors need to be taken into account for strain selection.

In the research presented here, only pH, proteins, peptides and polyphenols were investigated. There remain other components of the metabolome, however, that were not touched upon. These include sugars and free amino acids, in particular, which are known to be vital contributors to flavour and aroma. The lack of instrumentation and labour were hindrances to studying these aspects of the bean. However, further investigation will still need to be done in order to assess how these components are affected by these factors, such that a concise model of

fermentation can be developed. The realization of such a study could have a variety of applications. One example would be to preserve natural bean sugars, which are known to decrease significantly during fermentation (Rohan, 1964), while still forming flavour and aroma precursors, such that the produced chocolate would not need any added sugar.

In conclusion, the commercial-scale profiles of proteins and polyphenols can be reproduced in lab-scale fermentation, which validates hypothesis statement 2. An effective temperature regime, according to all three trials of forced fermentation, was established. The Forced Fermentation system shows greater promise in the immediate establishment of a more controlled method of fermentation, exploiting starter cultures and temperature "boosts" to drive the fermentation. This would also be considered a more natural process and, once perfected in terms of strain selection, can prove to be an effective tool in regulating fermentation even in the face of adverse environmental conditions.

6.3. Formation of Flavour and Aroma

When it comes to the desirable flavour and aroma characteristics of fine-flavour chocolate, there is a vast number of contributing compounds. The majority of these components are formed during the process of roasting, but their precursors mainly stem from the processes of fermentation and drying. Among these precursors are peptides, amino acids and sugars, in particular, which are the fundamental components needed for Maillard reactions that take place during drying and, primarily, roasting.

Experiments conducted by Voigt and co-workers over the years have methodically elucidated and formation of cocoa- and nutty-specific aroma from the breakdown of cocoa proteins and that proteolysis of vicilin is the major contributor (Voigt et al., 1994, 2016, 2018; Voigt & Lieberei, 2015). The fermentation process plays on the activation of enzymes that would usually be used in the generation of subunits for protein synthesis by a germinating embryo, and leaves the bean with a combination of hydrophilic peptides and hydrophobic amino acids, along with a dead embryo. The peptide repertoires obtained in the trials of fermentation outlined in this thesis also have denoted a preference for hydrophilicity. Further scrutiny of the peptide termini in chapter 3 pointed to endo- and exoprotease activity of aspartic proteinase and carboxypeptidase, respectively. The same experiments also revealed a pH dependency of these enzymes but not to the limits described in literature, which were much narrower than herein

observed. These results were further corroborated by an analysis of the N- and C-termini of the peptides of the Forced Fermentation system in chapter 4 (data not shown). A significant endoprotease activity at pH 5 could either suggest that there are other proteases at play or that the cocoa bean matrix provides a different working environment compared to the solely *in vitro* experiments previously carried out (Guilloteau et al., 2005; Laloi et al., 2002; Voigt et al., 1994). Nevertheless, these experiments highlighted the importance of bean pH and also verified the claim that a bean pH value of around 5.0 at the end of fermentation is optimal. Accordingly, bean pH was established as a good marker for fermentation, in terms of proteolysis.

A great advantage of the experiments in the second manuscript was the incorporation of the data from roasted beans and liquor samples. Through an analysis of peptides along the processing chain, it became clear that there were a handful of peptides speculated to be involved in Maillard reactions, due to their unambiguous disappearance upon roasting. Two peptides (*NNPYY* and *PVNSPGKY*) stood out in particular due to their occurrences in literature as being thought to be peptides responsible for cocoa and nutty flavour (Voigt et al., 2018). Despite there being no substantial evidence as yet to back these claims, the results of chapter 4 supported this line of thought. In order to verify these claims, synthetic versions of these two peptides were roasted with glucose and fructose in a matrix of deodorized cocoa butter for around 10 min at 200°C. Preliminary results hint at *PVNSPGKY* contributing to nutty flavour, however, the results for *NNPYY* were inconclusive (data not shown). These peptides will have to be further investigated and in greater detail in order to also discover what Maillard reaction products they form.

A downside to the analytical techniques employed in these manuscripts for the study of the cocoa peptidome was that only tetra-peptides and higher oligomeric peptides were studied. Even though the method of UHPLC-MS was able to detect di- and tri-peptides, they were not focused on in this research due to the sheer size of those repertoires. Moreover, the low sequence specificity of trimeric and smaller oligopeptides made it difficult to ascertain their respective precursor proteins. Logically, a study of these di- and tri-peptides should also be carried out due to them having a greater propensity to form flavour components.

An expansion of these analytical methods to a more diverse sample set of spontaneous fermentation trials and their subsequent roasting trials would also help expose other peptides that may contribute to cocoa flavour. Furthermore, such an analysis would also help in ascertaining

whether there are peptides unique to certain criteria, which could also serve as markers in tracing a batch of cocoa beans to a specific region or hybrid, for example. Such an approach has already revealed that there are peptidomic fingerprints between regions (Kumari et al., 2018) but has not yet revealed the key peptides responsible.

Polyphenols are also key contributors to taste since they impart bitter flavour. Ideally, a fermentation trial should end with a much lower level of polyphenols than initially, to alleviate bitterness in the resulting chocolate. Flavanols constitute the major class of phenolic compounds in the bean and the experiments of chapter 4 follow the convention of having lower flavanol content at the end of the microbial fermentation trials than at the beginning. The drop in polyphenols is attributed to thermal degradation, oxidation by polyphenol oxidase and leaching into the sweating (Aprotosoaie et al., 2016; Voigt & Lieberei, 2015).

However, both the Artificial and Forced Fermentation systems showed increases in polyphenol content at some point along the incubation/fermentation. Speculation on this occurrence has been discussed in section 6.2. This increase would also affect the taste of the cocoa. Results of chapter 5 also revealed that temperature has a profound influence over the gain in flavanols and that this is mainly achieved during the first 24 – 36 hours of incubation. While the results of chapter 4 pointed to a correlation between the increase in flavanols and the formation of acetic acid, the DoE experiments in chapter 5 showed this to be not as strong a correlation as initially thought. Albeit, decreasing pH, through the diffusion of lactic and acetic acids into the bean was seen to have a linear correlation with flavanol content. Having observed that temperature shows a strong positive correlation with flavanol content, it stands to reason that thermal degradation of flavanols during fermentation is almost negligible. Instead, these results point toward an enzymatic formation of flavanols. Such an observation has never before been studied in literature and these results are proof that the story of polyphenols during fermentation is more complex than once thought and needs further investigation.

The abundance of polyphenols in cocoa and in particular during the course of fermentation suggests that polyphenols also may be involved in other biochemical reactions. Evidence of polyphenols depleting sugar and amino acid concentrations and preventing the generation of cocoa flavour suggests that they may be binding to these precursors and affecting Maillard reactions that take place later on (Misnawi et al., 2004; Noor-Soffalina et al., 2009). An excess amount of polyphenols in fermented and dried beans will consequently reduce cocoa and aroma

flavour in the chocolate.

Therefore, to eliminate the unpleasant bitter notes of polyphenols in chocolate and improve upon cocoa flavour, one must account for the spike in temperature, the influx of acid, and the duration of the fermentation over which this occurs. According to these results, spontaneous fermentation that sees the growth of AAB after three to four days of fermentation would be ideal in maintaining a generally low flavanol content.

In conclusion, certain peptides show promise in being markers of fermentation. *PVNSPGKY* and *NNPYY* are examples of this as they are generated towards the end of fermentation and disappear upon roasting. Polyphenols, which not only contribute to bitterness but also affect the formation of cocoa flavour during roasting, need to be considered and kept to narrow limits in order to improve the taste of chocolate. In this regard, parameters such as temperature and acid production need to be carefully monitored.

6.4. Implementation of Model Fermentation Systems

In light of the reliability and robustness of the Artificial Fermentation system, a study investigating the incubation of beans with acetic and citric acid for their effects on anthocyanin content was conducted. Fresh and de-pulped beans were incubated in varying temperatures and concentrations of acetic acid and citric acid for specific periods of time using a DoE approach. The acids exhibited a preservation effect on the anthocyanins, which has been previously described in literature (Khoo et al., 2017), but also resulted in increased levels of anthocyanins in unnaturally high acid concentrations (data not shown). The application of submerged incubations, therefore, could be implemented in the production of chocolate which requires high anthocyanin content. The simplicity of the setup would allow for post-harvest processes of submerged incubation and immediate drying on the plantations themselves, in order to maintain high levels of anthocyanins prior to shipment.

An example of an application of the Forced Fermentation system would be in testing out the behaviour of certain strains during fermentation. An investigation into the potential of using *Bacillus subtilis* (natto) for the purpose of naturally raising bean pH at the end of very acidic fermentation trials showed that the strains being used could not grow on acidic cocoa pulp (data not shown). However, there are other *Bacillus* spp. that are seen to grow on the pulp towards the end of fermentation, and genetically engineering these species could result in bacteria that grow

under such acidic conditions and be able to naturally alkalize the bean.

The observations made using these systems in the laboratory could also be used to advise farmers on how to improve consistency between trials of fermentation. From the results in chapter 5, temperature is an important parameter in terms of pH change, proteolysis and flavanol dynamics. Thus, it is essential that beans are exposed to the correct temperature changes during fermentation. This can be easily achieved by avoiding fermenting beans during damp and cold weather spells, or conducting fermentation within a more temperature-regulated environment. This is to ensure that the temperature at the start of fermentation would be high enough to promote ample microorganismal growth, which would naturally drive the rest of the temperature regime.

Furthermore, farmers could also be advised to prevent the turning of perceptibly very alcoholic fermentations too frequently. AAB convert ethanol to acetic acid in the presence of oxygen. This means that as bean piles containing naturally high ethanol content in the pulp are turned more often, more acetic acid is produced which can result in very acidic beans. Further measures, such as alkalization and long conching times, would need to be taken upon such beans in order to get rid of the excess acidity. This phenomenon can already be avoided at the stage of fermentation and therefore incur a more profitable production process.

In conclusion, both these systems of fermentation were immensely beneficial in controlling the process of fermentation and investigations into the biochemistry of the cocoa bean. Larger scale direct implementations of these systems will have to reckon with certain factors:

1. Equipment: both systems require controlled temperature conditions and the artificial system, in particular, requires the use of chemicals. Data logging will also have to be done which requires special analytical devices.
2. Labour: the artificial system requires that beans be de-pulped prior to incubation which would require a considerable amount of time and manpower.
3. Space: specialized equipment will need to be placed somewhere indoors and will take up space, unlike a spontaneous fermentation where beans are piled outside.

Feasibility studies will therefore have to be conducted in order to find out whether the systems are suitable for their intended purpose and if the investments are worth the returns.

7. Conclusions

The first two objectives outlined in section 2.4 have been successfully realized. Both model systems have shown great control over influencing the changes in the biochemistry of the bean using microbes and artificial incubations. These experiments were conducted thousands of miles away from the plantations where the pods were harvested from. Despite minor setbacks, the results are proof that an efficient logistics chain had been established from plantation to processing, allowing for a detailed study of the metabolome of the cocoa bean.

The third objective was also accomplished through a series of good-quality and poor fermentation trials that were observed in these experiments. Fortunately, there were more successful fermentation trials than otherwise, through which even slight discrepancies between these fermentation trials could be analysed. It was concluded that poor quality fermentation was the result of adverse temperature conditions and incorrect microbial dynamics which affected bean pH and the profiles of proteins, peptides and polyphenols within the bean. Certain peptides were seen to be candidates for markers of cocoa aroma potential, and high polyphenol content was seen be disadvantageous to taste in terms of bitterness and potentially reducing cocoa aroma. The aspect of bean pH was observed and confirmed to be one of the key markers of fermentation.

Methods of refining fermentation, as outlined in the fourth objective, were still underway. The Forced Fermentation system using a starter culture of only yeast and AAB was able to guide the fermentation successfully and produce decent chocolate. The Artificial Fermentation system using DoE revealed what factors needed to be taken into consideration but could not as yet be connected to sensory profiles which would have been the next step. Nevertheless, through the knowledge generated by the research in this thesis, the “black box” stigmatisation of cocoa bean fermentation has been rescinded, clearing the path for innovation and industrialization.

8. References

Adler, P., Frey, L. J., Berger, A., Bolten, C. J., Hansen, C. E., & Wittmann, C. (2014). The key to acetate: Metabolic fluxes of acetic acid bacteria under cocoa pulp fermentation-simulating conditions. *Applied and Environmental Microbiology*, *80*(15), 4702–4716. https://doi.org/10.1128/AEM.01048-14

Afoakwa, E. O., Budu, A. S., Mensah-Brown, H., Takrama, J. F., & Akomanyi, E. (2014). Changes in Biochemical and Physico-chemical Qualities during Drying of Pulp Preconditioned and Fermented Cocoa (*Theobroma cacao*) Beans. *Journal of Nutritional Health & Food Science*, *2*(3), 1–8. https://doi.org/10.15226/jnhfs.2014.00121

Afoakwa, E. O., Kongor, J. E., Takrama, J. F., Budu, A. S., & Mensah-Brown, H. (2013). Effects of Pulp Preconditioning on Total Polyphenols, O-diphenols and Anthocyanin Concentrations during Fermentation and Drying of Cocoa (*Theobroma cacao*) Beans. *Journal of Food Science and Engineering*, *3*, 235–245.

Afoakwa, E. O., Paterson, A., Fowler, M. S., & Ryan, A. (2008). Flavor formation and character in cocoa and chocolate: A critical review. *Critical Reviews in Food Science and Nutrition*, *48*(9), 840–857. https://doi.org/10.1080/10408390701719272

Afoakwa, E. O., Quao, J., Takrama, J., Budu, A. S., & Saalia, F. K. (2012). Influence of pulp-preconditioning and fermentation on fermentative quality and appearance of ghanaian cocoa (*Theobroma cacao*) beans. *International Food Research Journal*, *19*(1), 127–133. https://doi.org/10.1016/j.envsci.2018.01.004

Afoakwa, E. O., Quao, J., Takrama, J., Budu, A. S., & Saalia, F. K. (2013). Chemical composition and physical quality characteristics of Ghanaian cocoa beans as affected by pulp pre-conditioning and fermentation. *Journal of Food Science and Technology*, *50*(6), 1097–1105. https://doi.org/10.1007/s13197-011-0446-5

Albertini, B., Schoubben, A., Guarnaccia, D., Pinelli, F., Vecchia, M. Della, Ricci, M., … Blasi, P. (2015). Effect of Fermentation and Drying on Cocoa Polyphenols. *Journal of Agricultural and Food Chemistry*, *63*, 9948–9953. https://doi.org/10.1021/acs.jafc.5b01062

Alfaro-Espinoza, G., & Ullrich, M. S. (2014). Marinobacterium mangrovicola sp. nov., a marine nitrogen-fixing bacterium isolated from mangrove roots of *Rhizophora mangle*. *International Journal of Systematic and Evolutionary Microbiology*, *64*, 3988–3993. https://doi.org/10.1099/ijs.0.067462-0

Amin, I., Jinap, S., Jamilah, B., Harikrisna, K., & Biehl, B. (2002). Oligopeptide patterns produced from *Theobroma cacao* L of various genetic origins. *Journal of the Science of Food and Agriculture*, *82*, 733–737. https://doi.org/10.1002/jsfa.1098

Andersson, M., Koch, G., & Lieberei, R. (2006). Structure and function of the seed coat of *Theobroma cacao* L . and its possible impact on flavour precursor development during fermentation. *Journal of Applied Botany and Food Quality*, *80*, 48–62.

Aprotosoaie, A. C., Luca, S. V., & Miron, A. (2016). Flavor Chemistry of Cocoa and Cocoa Products-An Overview. *Comprehensive Reviews in Food Science and Food Safety*, *15*(1), 73–91. https://doi.org/10.1111/1541-4337.12180

Arlorio, M., Locatelli, M., Travaglia, F., Coïsson, J. D., Grosso, E. Del, Minassi, A., … Martelli, A. (2008). Roasting impact on the contents of clovamide (N-caffeoyl-L-DOPA) and the antioxidant activity of cocoa

beans (*Theobroma cacao* L.). *Food Chemistry*, *106*(3), 967–975. https://doi.org/10.1016/j.foodchem.2007.07.009

Ascrizzi, R., Flamini, G., Tessieri, C., & Pistelli, L. (2017). From the raw seed to chocolate: Volatile profile of Blanco de Criollo in different phases of the processing chain. *Microchemical Journal*, *133*, 474–479. https://doi.org/10.1016/j.microc.2017.04.024

Auld, D. S., & Vallee, B. L. (1970). Kinetics of Carboxypeptidase A. II Inhibitors of the Hydrolysis of Oligopeptides, *9*(3), 602–609.

Bartley, B. G. . (2005). *The Genetic Diversity of Cacao and its Utilization*. Oxfordshire: CABI Publishing.

Bayliak, M. M., Burdyliuk, N. I., & Lushchak, V. I. (2016). Effects of pH on antioxidant and prooxidant properties of common medicinal herbs. *Open Life Sciences*, *11*(1), 298–307. https://doi.org/10.1515/biol-2016-0040

Bernaert, H., Blondeel, I., Allegaert, L., & Lohmueller, T. (2012). Industrial Treatment of Cocoa in Chocolate Production: Health Implications. In A. Conti, R. Paoletti, A. Poli, & F. Visioli (Eds.), *Chocolate and Health* (pp. 17–31). Milano: Springer.

Biehl, B., Brunner, E., Passern, D., Quesnel, V. C., & Adomako, D. (1985). Acidification, proteolysis and flavour potential in fermenting cocoa beans. *Journal of the Science of Food and Agriculture*, *36*(7), 583–598. https://doi.org/10.1002/jsfa.2740360710

Biehl, B., & Passern, D. (1982). Proteolysis during Fermentation-like Incubation of Cocoa Seeds. *Journal of the Science of Food and Agriculture*, *33*, 1280–1290.

Biehl, B., Passern, D., & Sagemann, W. (1982). Effect of Acetic Acid on Subcellular Structures of Cocoa Bean Cotyledons. *Journal of the Science of Food and Agriculture*, *33*, 1101–1109.

Bonvehí, J. S. (2005). Investigation of aromatic compounds in roasted cocoa powder. *European Food Research and Technology*, *221*, 19–29. https://doi.org/10.1007/s00217-005-1147-y

Bonvehí, J. S., & Coll, F. V. (1997). Parameters affecting the quality of processed cocoa powder: acidity fraction. *Zeitschrift Für Lebensmittel-Untersuchung Und -Forschung*, *204*, 287–292. https://doi.org/DOI 10.1007/s002170050079

Bonvehí, J. S., & Coll, V. (2002). Factors Affecting the Formation of Alkylpyrazines during Roasting Treatment in Natural and Alkalinized Cocoa Powder. *Journal of Agricultural and Food Chemistry*, 3743–3750.

Borchers, A. T., Keen, C. L., Hannum, S. M., & Gershwin, M. E. (2000). Cocoa and Chocolate: Composition, Bioavailability, and Health Implications. *Journal of Medicinal Food*, *3*(2), 77–105. https://doi.org/10.1089/109662000416285

Buamah, R., Dzogbefia, V. P., & Oldham, J. H. (1997). Pure yeast culture fermentation of cocoa (*Theobroma cacao* L): effect on yield of sweatings and cocoa bean quality. *World Journal of Microbiology & Biotechnology*, *13*, 457–462. https://doi.org/10.1023/a:1018536519325

Buyukpamukcu, E., Goodall, D. M., Hansen, C. E., Keely, B. J., Kochhar, S., & Wille, H. (2001). Characterization of peptides formed during fermentation of cocoa bean. *Journal of Agricultural and Food Chemistry*, *49*(12), 5822–5827. https://doi.org/10.1021/jf0104127

Bytof, G., Biehl, B., Heinrichs, H., & Voigt, J. (1995). Specificity and stability of the carboxypeptidase activity in

ripe, ungerminated seeds of *Theobroma cacao* L. *Food Chemistry*, *54*(1), 15–21. https://doi.org/10.1016/0308-8146(95)92657-6

Caligiani, A., Marseglia, A., Prandi, B., Palla, G., & Sforza, S. (2016). Influence of fermentation level and geographical origin on cocoa bean oligopeptide pattern. *Food Chemistry*, *211*, 431–439. https://doi.org/10.1016/j.foodchem.2016.05.072

Camu, N., De Winter, T., Verbrugghe, K., Cleenwerck, I., Vandamme, P., Takrama, J. S., … De Vuyst, L. (2007). Dynamics and biodiversity of populations of lactic acid bacteria and acetic acid bacteria involved in spontaneous heap fermentation of cocoa beans in Ghana. *Applied and Environmental Microbiology*, *73*(6), 1809–1824. https://doi.org/10.1128/AEM.02189-06

Castro-Alayo, E. M., Idrogo-Vásquez, G., Siche, R., & Cardenas-Toro, F. P. (2019). Formation of aromatic compounds precursors during fermentation of Criollo and Forastero cocoa. *Heliyon*, *5*(1), e01157. https://doi.org/10.1016/j.heliyon.2019.e01157

Cheynier, V. (2005). Polyphenols in foods are more complex than often thought. The American Journal of Clinical Nutrition. *The American Journal of Clinical Nutrition*, *81*, 223s-229s. https://doi.org/81/1/223S [pii]

Clarence-Smith, W. G. (2000). The commodity chain. In *Cocoa and Chocolate, 1765-1914* (pp. 1–6). London: Routledge.

Consonni, V., Ballabio, D., & Todeschini, R. (2010). Evaluation of model predictive ability by external validation techniques. *Journal of Chemometrics*, *24*(3–4), 194–201. https://doi.org/10.1002/cem.1290

Crown, P. L., & Hurst, W. J. (2009). Evidence of cacao use in the Prehispanic American Southwest. *Proceedings of the National Academy of Sciences*, *106*(7), 2110–2113. https://doi.org/10.1073/pnas.0812817106

D'Hondt, K., Bosch, D., Van Damme, J., Goethals, M., Vandekerckhove, J., & Krebbers, E. (1993). An aspartic proteinase present in seeds cleaves Arabidopsis 2 S albumin precursors in vitro. *Journal of Biological Chemistry*, *268*(28), 20884–20891.

D'Souza, R. N., Grimbs, A., Grimbs, S., Behrends, B., Corno, M., Ullrich, M. S., & Kuhnert, N. (2018). Degradation of cocoa proteins into oligopeptides during spontaneous fermentation of cocoa beans. *Food Research International*, *109*(May), 506–516. https://doi.org/10.1016/j.foodres.2018.04.068

D'Souza, R. N., Grimbs, S., Behrends, B., Bernaert, H., Ullrich, M. S., & Kuhnert, N. (2017). Origin-based polyphenolic fingerprinting of *Theobroma cacao* in unfermented and fermented beans. *Food Research International*, *99*(1), 550–559. https://doi.org/10.1016/j.foodres.2017.06.007

Damodaran, S., Parkin, K. L., & Fennema, O. R. (2008). *Fennema's Food Chemistry* (4th Ed.). Florida: CRC Press Taylor and Francis Group.

Daniel, H., Vrancken, G., Takrama, J. F., Camu, N., & Vos, P. De. (2009). Yeast diversity of Ghanaian cocoa bean heap fermentations. *FEMS Yeast Research*, *9*, 774–783. https://doi.org/10.1111/j.1567-1364.2009.00520.x

De Vuyst, L., & Weckx, S. (2016a). The cocoa bean fermentation process: from ecosystem analysis to starter culture development. *Journal of Applied Microbiology*, *121*(1), 5–17. https://doi.org/10.1111/jam.13045

De Vuyst, L., & Weckx, S. (2016b). The Functional Role of Lactic Acid Bacteria in Cocoa Bean Fermentation. In F. Mozzi, R. R. Raya, & G. M. Vignolo (Eds.), *Biotechnology of Lactic Acid Bacteria Novel Applications* (Ed.,

2nd, pp. 248–278). John Wiley and Sons, Ltd. https://doi.org/10.1002/9780813820866.ch17

Dreiss, M., & Greenhill, S. (2008). *Chocolate : pathway to the gods*. Arizona: University of Arizona Press.

Endraiyani, V., Ludescher, R. D., Di, R., & Karwe, M. V. (2017). Total Phenolics and Antioxidant Capacity of Cocoa Pulp: Processing and Storage Study. *Journal of Food Processing and Preservation*, *41*, e13029. https://doi.org/10.1111/jfpp.13029

Eteng, M. U., Eyong, E. U., Akpanyung, E. O., Agiang, M. A., & Aremu, C. Y. (1997). Recent advances in caffeine and theobromine toxicities: A review. *Plant Foods for Human Nutrition*, *51*, 231–243.

Everitt, E., Persson, M. J., & Wohlfart, C. (1988). pH-Dependent exposure of endoproteolytic cleavage sites of the adenovirus 2 hexon protein. *FEMS Microbiology Letters 49*, *49*, 229–233.

Eyamo Evina, V. J., De Taeye, C., Niemenak, N., Youmbi, E., & Collin, S. (2016). Influence of acetic and lactic acids on cocoa flavan-3-ol degradation through fermentation-like incubations. *LWT - Food Science and Technology*, *68*, 514–522. https://doi.org/10.1016/j.lwt.2015.12.047

Farah, D. M. H., Zaibunnisa, A. H., & Misnawi. (2012). Optimization of cocoa beans roasting process using response surface methodology based on concentration of pyrazine and acrylamide. *International Food Research Journal*, *19*(4), 1355–1359.

Fowler, M. S. (2009). Cocoa Beans: From Tree to Factory. In S. T. Beckett (Ed.), *Industrial Chocolate Manufacture and Use* (4th Ed., pp. 10–47). York: Blackwell Publishing.

Gidda, S. K., Miersch, O., Levitin, A., Schmidt, J., Wasternack, C., & Varin, L. (2003). Biochemical and molecular characterization of a hydroxyjasmonate sulfotransferase from *Arabidopsis thaliana*. *Journal of Biological Chemistry*, *278*(20), 17895–17900. https://doi.org/10.1074/jbc.M211943200

Granato, D., & de Araújo Calado, V. M. (2014). The use and importance of design of experiments (DOE) in process modelling in food science and technology. In *Mathematical and Statistical Methods in Food Science and Technology* (pp. 1–18). John Wiley and Sons, Ltd. https://doi.org/10.1002/9781118434635.ch01

Guehi, T. S., Konan, Y. M., Koffi-Nevry, R., N'dri, D. Y., & Manizan, N. P. (2007). Enumeration and identification of main fungal isolates and evaluation of fermentation's degree of Ivorian raw cocoa beans. *Australian Journal of Basic and Applied Sciences*, *1*(4), 479–486.

Guilloteau, M., Laloi, M., Michaux, S., Bucheli, P., & McCarthy, J. (2005). Identification and characterisation of the major aspartic proteinase activity in *Theobroma cacao* seeds. *Journal of the Science of Food and Agriculture*, *85*, 549–562. https://doi.org/10.1002/jsfa.1777

Hansen, C. E., Del Olmo, M., & Burri, C. (1998). Enzyme activities in cocoa beans during fermentation. *Journal of the Science of Food and Agriculture*, *77*(2), 273–281. https://doi.org/10.1002/(SICI)1097-0010(199806)77:2<273::AID-JSFA40>3.0.CO;2-M

Hashim, P., Selamat, J., Muhammad, K., & Ali, A. (1999). Effect of drying time, bean depth and temperature on free amino acid, peptide-N, sugar and pyrazine concentrations of Malaysian cocoa beans. *Journal of the Science of Food and Agriculture*, *79*, 987–994. https://doi.org/10.1002/(SICI)1097-0010(19990515)79:7<987::AID-JSFA314>3.0.CO;2-H

Ho, V. T. T., Zhao, J., & Fleet, G. (2015). The effect of lactic acid bacteria on cocoa bean fermentation.

International Journal of Food Microbiology, *205*, 54–67. https://doi.org/10.1016/j.ijfoodmicro.2015.03.031

Holm, C. S., Aston, J. W., & Douglas, K. (1993). The effects of the organic acids in cocoa on the flavour of chocolate. *Journal of the Science of Food and Agriculture*, *61*(1), 65–71. https://doi.org/10.1002/jsfa.2740610111

Ismail, A., Lamin, K., Gangadharan, D., Ahmad, A. C., Bakar, S., & Yazik, N. M. (2009). *Cocoa Planting Manual*. Malaysia.

Jacobsen, N. E., & Bartlett, P. A. (1981). A Phosphonamidate Dipeptide Analogue as an Inhibitor of Carboxypeptidase A. *J . Am. Chem. SOC*, *103*, 654–657. https://doi.org/10.1021/ja00393a026

Jalil, A. M. M., & Ismail, A. (2008). Polyphenols in cocoa and cocoa products: Is there a link between antioxidant properties and health? *Molecules*, *13*(9), 2190–2219. https://doi.org/10.3390/molecules13092190

Janek, K., Niewienda, A., Wöstemeyer, J., & Voigt, J. (2016). The cleavage specificity of the aspartic protease of cocoa beans involved in the generation of the cocoa-specific aroma precursors. *Food Chemistry*, *211*, 320–328. https://doi.org/10.1016/j.foodchem.2016.05.033

Jinap, S., & Danker, R. M. (1993). Effect of clone and temperature onacetic acid diffusion and death of cocoa beans. *ASEAN Food Journal*, *8*(4), 139–144.

John, W. A., Böttcher, N. L., Aßkamp, M., Bergounhou, A., Kumari, N., Ho, P.-W., … Ullrich, M. S. (2019). Forcing Fermentation: Profiling Proteins, Peptides and Polyphenols in Lab-scale Cocoa Bean Fermentation. *Food Chemistry*, *278*, 786–794. https://doi.org/10.1016/j.foodchem.2018.11.108

John, W. A., Kumari, N., Böttcher, N. L., Koffi, K. J., Grimbs, S., Vrancken, G., … Ullrich, M. S. (2016). Aseptic artificial fermentation of cocoa beans can be fashioned to replicate the peptide profile of commercial cocoa bean fermentations. *Food Research International*, *89*, 764–772. https://doi.org/10.1016/j.foodres.2016.10.011

Kadow, D., Niemenak, N., Rohn, S., & Lieberei, R. (2015). Fermentation-like incubation of cocoa seeds (*Theobroma cacao* L.) - Reconstruction and guidance of the fermentation process. *LWT - Food Science and Technology*, *62*, 357–361. https://doi.org/10.1016/j.lwt.2015.01.015

Kamphuis, H. J. (2017). Production of cocoa mass, cocoa butter and cocoa powder. In S. T. Beckett (Ed.), *Beckett's Industrial Chocolate Manufacture and Use* (4th Ed, pp. 50–71). York: Blackwell Publishing. https://doi.org/10.1002/9781118923597.ch3

Khoo, H. E., Azlan, A., Tang, S. T., & Lim, S. M. (2017). Anthocyanidins and anthocyanins: colored pigments as food, pharmaceutical ingredients, and the potential health benefits. *Food & Nutrition Research*, *61*(1), 1361779. https://doi.org/10.1080/16546628.2017.1361779

Kim & Keeney. (1984). (-)-Epicatechin Content in Fermented and Unfermented Cocoa Beans. *Journal of Food Science*, *49*, 1090–1092.

Kirchhoff, P. M., Biehl, B., & Crone, G. (1989). Peculiarity of the accumulation of free amino acids during cocoa fermentation. *Food Chemistry*, *31*, 295–311. https://doi.org/10.1016/0308-8146(89)90071-X

Kochhar, S., Hansen, C. E., Juillerat, M. A., Wille, H., Buyukpamukcu, E., Keely, B., & Goodall, D. M. (2004). *US 20040202761A1*. US.

Kongor, J. E., Hinneh, M., de Walle, D. Van, Afoakwa, E. O., Boeckx, P., & Dewettinck, K. (2016). Factors

influencing quality variation in cocoa (*Theobroma cacao*) bean flavour profile - A review. *Food Research International*, *82*, 44–52. https://doi.org/10.1016/j.foodres.2016.01.012

Kongor, J. E., Takrama, J. F., Budu, A. S., Mensah-brown, H., & Afoakwa, E. O. (2013). Effects of Fermentation and Drying on the Fermentation Index and Cut Test of Pulp Pre-conditioned Ghanaian Cocoa (*Theobroma cacao*) Beans. *Journal of Food Science and Engineering*, *3*, 625–634.

Kratzer, U., Frank, R., Kalbacher, H., Biehl, B., Wöstemeyer, J., & Voigt, J. (2009). Subunit Structure of the vicilin-like globular storage protein of cocoa seeds and the origin of cocoa- and chocolate- specific aroma precursors. *Food Chemistry*, *113*, 903–913. https://doi.org/10.1016/j.foodchem.2008.08.017

Kresnowati, M., Gunawan, A., & Muliyadini, W. (2015). Kinetics model development of cocoa bean fermentation. *AIP Conference Proceedings*, *1699*, 030004.

Kumari, N. (2018). *Unravelling the proteome and peptidome biochemistry in Theobroma cacao*. Jacobs University Bremen.

Kumari, N., Grimbs, A., D'Souza, R. N., Verma, S. K., Corno, M., Kuhnert, N., & Ullrich, M. S. (2018). Origin and varietal based proteomic and peptidomic fingerprinting of *Theobroma cacao* in non-fermented and fermented cocoa beans. *Food Research International*, *111*(April), 137–147. https://doi.org/10.1016/j.foodres.2018.05.010

Kumari, N., Jean, K., Grimbs, S., Souza, R. N. D., Kuhnert, N., Vrancken, G., & Ullrich, M. S. (2016). Biochemical fate of vicilin storage protein during fermentation and drying of cocoa beans. *Food Research International*, *90*, 53–65. https://doi.org/10.1016/j.foodres.2016.10.033

Laloi, M., Mccarthy, J., Morandi, O., Gysler, C., & Bucheli, P. (2002). Molecular and biochemical characterisation of two aspartic proteinases TcAP1 and TcAP2 from *Theobroma cacao* seeds. *Planta*, *215*, 754–762.

Lambrechts, M. G., & Pretorius, I. S. (2000). Yeast and its Importance to Wine Aroma - A Review. *South African Journal of Enology and Viticulture*, *21*(Special Issue), 97–129. https://doi.org/10.1080/00207230701791448

Laurent, V., Risterucci, A. M., & Lanaud, C. (1994). Genetic diversity in cocoa revealed by cDNA probes. *Theoretical and Applied Genetics*, *88*(2), 193–198. https://doi.org/10.1007/BF00225897

Leal Jr, G. A., Luiz, H. G., Efraim, P., Tavares, F. C. D. A., & Figueira, A. (2008). Fermentation of cacao (*Theobroma cacao* L.) seeds with a hybrid *Kluyveromyces marxianus* strain improved product quality attributes. *FEMS Yeast Research*, *8*, 788–798. https://doi.org/10.1111/j.1567-1364.2008.00405.x

Lee, A. H., Neilson, A. P., O'Keefe, S. F., Ogejo, J. A., Huang, H., Ponder, M., … Stewart, A. C. (2019). A laboratory-scale model cocoa fermentation using dried, unfermented beans and artificial pulp can simulate the microbial and chemical changes of on-farm cocoa fermentation. *European Food Research and Technology*, *245*(0), 511–519. https://doi.org/10.1007/s00217-018-3171-8

Lefeber, T., Gobert, W., Vrancken, G., Camu, N., & De Vuyst, L. (2011). Dynamics and species diversity of communities of lactic acid bacteria and acetic acid bacteria during spontaneous cocoa bean fermentation in vessels. *Food Microbiology*, *28*(3), 457–464. https://doi.org/10.1016/j.fm.2010.10.010

Lefeber, T., Janssens, M., Camu, N., & De Vuyst, L. (2010). Kinetic analysis of strains of lactic acid bacteria and acetic acid bacteria in cocoa pulp simulation media toward development of a starter culture for cocoa bean

fermentation. *Applied and Environmental Microbiology*, *76*(23), 7708–7716. https://doi.org/10.1128/AEM.01206-10

Lefeber, T., Janssens, M., Moens, F., Gobert, W., & De Vuyst, L. (2011). Interesting starter culture strains for controlled cocoa bean fermentation revealed by simulated cocoa pulp fermentations of cocoa-specific lactic acid bacteria. *Applied and Environmental Microbiology*, *77*(18), 6694–6698. https://doi.org/10.1128/AEM.00594-11

Lefeber, T., Papalexandratou, Z., Gobert, W., Camu, N., & De Vuyst, L. (2012). On-farm implementation of a starter culture for improved cocoa bean fermentation and its influence on the flavour of chocolates produced thereof. *Food Microbiology*, *30*, 379–392. https://doi.org/10.1016/j.fm.2011.12.021

Lerceteau, E., Robert, T., Pétiard, V., & Crouzillat, D. (1997). Evaluation of the extent of genetic variability among *Theobroma cacao* accessions using RAPD and RFLP markers. *Theoretical and Applied Genetics*, *95*, 10–19. https://doi.org/10.1007/s001220050527

Lerceteau, E., Rogers, J., Pétiard, V., & Crouzillat, D. (1999). Evolution of cacao bean proteins during fermentation : a study by two-dimensional electrophoresis. *Journal of the Science of Food and Agriculture*, *79*, 619–625.

Li, H., Liu, F., Kang, L., & Zheng, M. (2016). Study on the buffering capacity of wort. *Journal of the Institute of Brewing*, *122*(1), 138–142. https://doi.org/10.1002/jib.286

Liendo, R., Padilla, F. C., & Quintana, A. (1997). Characterization of cocoa butter extracted from Criollo cultivars of *Theobroma cacao* L. *Food Research International*, *30*(9), 727–731. https://doi.org/10.1016/S0963-9969(98)00025-8

Lipp, M., Simoneau, C., Ulberth, F., Anklam, E., Crews, C., Brereton, P., … Wiedmaier, C. (2001). Composition of genuine cocoa butter and cocoa butter equivalents. *Journal of Food Composition and Analysis*, *14*(4), 399–408. https://doi.org/10.1006/jfca.2000.0984

Lopes, U. V., & Pires, J. L. (2015). Botany and Production of Cocoa. In R. F. Schwan & G. Fleet (Eds.), *Cocoa and Coffee Fermentations* (pp. 43–70). Florida: Taylor & Francis.

Lund, M. N., & Ray, C. A. (2017). Control of Maillard Reactions in Foods: Strategies and Chemical Mechanisms. *Journal of Agricultural and Food Chemistry*, *65*, 4537–4552. https://doi.org/10.1021/acs.jafc.7b00882

Lundstedt, T., Seifert, E., Abramo, L., Thelin, B., Nyström, A., Pettersen, J., & Bergman, R. (1998). Experimental Design and Optimization. *Chemometrics and Intelligent Laboratory Systems*, *43*, 3–40. https://doi.org/10.1007/978-3-540-49148-4-3

Marseglia, A., Sforza, S., Faccini, A., Bencivenni, M., Palla, G., & Caligiani, A. (2014). Extraction, identification and semi-quantification of oligopeptides in cocoa beans. *Food Research International*, *63*, 382–389. https://doi.org/10.1016/j.foodres.2014.03.046

Meersman, E., Steensels, J., Struyf, N., Paulus, T., Saels, V., Mathawan, M., … Verstrepen, K. J. (2016). Tuning Chocolate Flavor through Development of Thermotolerant *Saccharomyces cerevisiae* Starter Cultures with Increased Acetate Ester Production. *Applied and Environmental Microbiology*, *82*(2), 732–746.

Meyer, B., Biehl, B., Said, M. Bin, & Samarakoddy, R. J. (1989). Postharvest pod storage: A method for pulp preconditioning to impair strong nib acidification during cocoa fermentation in Malaysia. *Journal of the*

Science of Food and Agriculture, *48*, 285–304. https://doi.org/10.1002/jsfa.2740480305

Misnawi, A., Jinap, S., Jamilah, B., & Nazamid, S. (2004). Sensory properties of cocoa liquor as affected by polyphenol concentration and duration of roasting. *Food Quality and Preference*, *15*, 403–409. https://doi.org/10.1016/S0950-3293(03)00097-1

Moens, F., Lefeber, T., & De Vuyst, L. (2014). Oxidation of Metabolites Highlights the Microbial Interactions and Role of *Acetobacter pasteurianus* during Cocoa Bean Fermentation. *Applied and Environmental Microbiology*, *80*(6), 1848–1857. https://doi.org/10.1128/AEM.03344-13

Mohr, V. W., Landschreiber, E., & Severin, T. (1976). Zur Spezifität des Kakaoaromas. *European Journal of Lipid Science and Technology*, *78*(2), 88–95.

Moldovan, R. C., Bodoki, E., Kacsó, T., Servais, A. C., Crommen, J., Oprean, R., & Fillet, M. (2016). A micellar electrokinetic chromatography–mass spectrometry approach using in-capillary diastereomeric derivatization for fully automatized chiral analysis of amino acids. *Journal of Chromatography A*, *1467*, 400–408. https://doi.org/10.1016/j.chroma.2016.08.035

Montgomery, D. C. (2009). *Design and analysis of experiments* (5th Ed.). New York: John Wiley and Sons, Inc.

Moreno-Zambrano, M., Grimbs, S., Ullrich, M. S., & Hütt, M. T. (2018). A mathematical model of cocoa bean fermentation. *Royal Society Open Science*, *5*(10), 180964. https://doi.org/10.1098/rsos.180964

Motamayor, J. C., Lachenaud, P., da Silva e Mota, J. W., Loor, R., Kuhn, D. N., Brown, J. S., & Schnell, R. J. (2008). Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L). *PLoS ONE*, *3*(10), e3311. https://doi.org/10.1371/journal.pone.0003311

Motamayor, J. C., Risterucci, A. M., Heath, M., & Lanaud, C. (2003). Cacao domestication II: progenitor germplasm of the Trinitario cacao cultivar. *Heredity*, *91*(3), 322–330. https://doi.org/10.1038/sj.hdy.6800298

Mottram, D. S. (2007). The Maillard Reaction: Source of Flavour in Thermally Processed Foods. In R. G. Berger (Ed.), *Flavours and Fragrances* (pp. 269–284). Berlin: Springer.

Müntz, K. (1996). Proteases and proteolytic cleavage of storage proteins in developing and germinating dicotyledonous seeds. *Journal of Experimental Botany*, *47*(298), 605–622. https://doi.org/10.1093/jxb/47.5.605

Müntz, K., Belozersky, M. A., Dunaevsky, Y. E., Schlereth, A., & Tiedemann, J. (2001). Stored proteinases and the initiation of storage protein mobilization in seeds during germination and seedling growth. *Journal of Experimental Botany*, *52*(362), 1741–1752. https://doi.org/10.1093/jexbot/52.362.1741

Nascimento, M. da S. do, Pena, P. O., Brum, D. M., Imazaki, F. T., Tucci, M. L. S., & Efraim, P. (2013). Behavior of Salmonella during fermentation, drying and storage of cocoa beans. *International Journal of Food Microbiology*, *167*, 363–368. https://doi.org/10.1016/j.ijfoodmicro.2013.10.003

Noor-Soffalina, S. S., Jinap, S., Nazamid, S., & Nazimah, S. A. H. (2009). Effect of polyphenol and pH on cocoa Maillard-related flavour precursors in a lipidic model system. *International Journal of Food Science and Technology*, *44*, 168–180. https://doi.org/10.1111/j.1365-2621.2008.01711.x

Obiri, B. D., Bright, G. A., McDonald, M. A., Anglaaere, L. C. N., & Cobbina, J. (2007). Financial analysis of shaded cocoa in Ghana. *Agroforestry Systems*, *71*, 139–149. https://doi.org/10.1007/s10457-007-9058-5

Papalexandratou, Z., & De Vuyst, L. (2011). Assessment of the yeast species composition of cocoa bean fermentations in different cocoa-producing regions using denaturing gradient gel electrophoresis. *FEMS Yeast Research*, *11*(7), 564–574. https://doi.org/10.1111/j.1567-1364.2011.00747.x

Patras, M. A., Milev, B. P., Vrancken, G., & Kuhnert, N. (2014). Identification of novel cocoa flavonoids from raw fermented cocoa beans by HPLC-MSn. *Food Research International*, *63*, 353–359. https://doi.org/10.1016/j.foodres.2014.05.031

Pereira, G. V. de M., Miguel, M. G. da C. P., Ramos, Cí. L., & Schwan, R. F. (2012). Microbiological and physicochemical characterization of small-scale cocoa fermentations and screening of yeast and bacterial strains to develop a defined starter culture. *Applied and Environmental Microbiology*, *78*(15), 5395–5405. https://doi.org/10.1128/AEM.01144-12

Pereira, G. V. de M., Soccol, V. T., & Soccol, C. R. (2016). Current state of research on cocoa and coffee fermentations. *Current Opinion in Food Science*, *7*, 50–57. https://doi.org/10.1016/j.cofs.2015.11.001

Qin, X. W., Lai, J. X., Tan, L. H., Hao, C. Y., Li, F. P., He, S. Z., & Song, Y. H. (2017). Characterization of volatile compounds in Criollo, Forastero, and Trinitario cocoa seeds (*Theobroma cacao* L.) in China. *International Journal of Food Properties*, *20*(10), 2261–2275. https://doi.org/10.1080/10942912.2016.1236270

Rashidah, S., Jinap, S., Nazamid, S., & Jamilah, B. (2007). Characterisation of the ability of globulins from legume seeds to produce cocoa specific aroma. *International Food Research Journal*, *14*(2), 103–114.

Reineccius, G. A., Andersen, D. A., Kavanagh, T. E., & Keeney, P. G. (1972). Identification and Quantification of the Free Sugars in Cocoa Beans. *Journal of Agricultural and Food Chemistry*, *20*(2), 199–202. https://doi.org/10.1021/jf60180a033

Ritter, U. G. (1999). *Flüchtige Aromastoffe der Kakao (Theobroma cacao L.) und der Cupuacu pulpe (Theobroma grandiflorum Spreng).* Universität Berlin.

Rocha, I. S., Santana, L. R. R. de, Soares, S. E., & Bispo, E. da S. (2017). Effect of the roasting temperature and time of cocoa beans on the sensory characteristics and acceptability of chocolate. *Food Science and Technology*, *37*(4), 522–530. https://doi.org/10.1590/1678-457x.16416

Rodriguez-Campos, J., Escalona-Buendía, H. B., Contreras-Ramos, S. M., Orozco-Avila, I., Jaramillo-Flores, E., & Lugo-Cervantes, E. (2012). Effect of fermentation time and drying temperature on volatile compounds in cocoa. *Food Chemistry*, *132*, 277–288. https://doi.org/10.1016/j.foodchem.2011.10.078

Rodriguez-Campos, J., Escalona-Buendía, H. B., Orozco-Avila, I., Lugo-Cervantes, E., & Jaramillo-Flores, M. E. (2011). Dynamics of volatile and non-volatile compounds in cocoa (*Theobroma cacao* L.) during fermentation and drying processes using principal components analysis. *Food Research International*, *44*, 250–258. https://doi.org/10.1016/j.foodres.2010.10.028

Rohan, T. A. (1964). The Precursors of Chocolate Aroma: A Comparative Study of Fermented and Unfermented Cocoa Beans. *Journal of Food Science*, *29*(4), 456–459. https://doi.org/10.1111/j.1365-2621.1964.tb01760.x

Romanens, E., Näf, R., Lobmaier, T., Pedan, V., Leischtfeld, S. F., Meile, L., & Schwenninger, S. M. (2018). A lab-scale model system for cocoa bean fermentation. *Applied Microbiology and Biotechnology*, *102*, 3349–3362. https://doi.org/10.1007/s00253-018-8835-6

Rombouts, J. (1953). Critical review of the yeast species previously described from cocoa. *Tropical Agriculture*, *30*, 34–41.

Saltini, R., Akkerman, R., & Frosch, S. (2013). Optimizing chocolate production through traceability: A review of the influence of farming practices on cocoa bean quality. *Food Control*, *29*, 167–187. https://doi.org/10.1016/j.foodcont.2012.05.054

Sandhya, M. V. S., Yallappa, B. S., Varadaraj, M. C., Puranaik, J., Rao, L. J., Janardhan, P., & Murthy, P. S. (2016). Inoculum of the starter consortia and interactive metabolic process in enhancing quality of cocoa bean (*Theobroma cacao*) fermentation. *LWT - Food Science and Technology*, *65*, 731–738. https://doi.org/10.1016/j.lwt.2015.09.002

Santhanam Menon, A., Hii, C. L., Law, C. L., Shariff, S., & Djaeni, M. (2017). Effects of drying on the production of polyphenol-rich cocoa beans. *Drying Technology*, *35*(15), 1799–1806. https://doi.org/10.1080/07373937.2016.1276072

Schwan, R. F. (1998). Cocoa fermentations conducted with a defined microbial cocktail inoculum. *Applied and Environmental Microbiology*, *64*(4), 1477–1483. https://doi.org/10.1017/CBO9781107415324.004

Schwan, R. F., & Wheals, A. E. (2004). The microbiology of cocoa fermentation and its role in chocolate quality. *Critical Reviews in Food Science and Nutrition*, *44*(4), 205–221. https://doi.org/10.1080/10408690490464104

Sfakianakis, P., & Tzia, C. (2014). Conventional and Innovative Processing of Milk for Yogurt Manufacture; Development of Texture and Flavor: A Review. *Foods*, *3*(1), 176–193. https://doi.org/10.3390/foods3010176

Sharafi, S. M., Rasooli, I., & Beheshti-Maal, K. (2010). Isolation, characterization and optimization of indigenous acetic acid bacteria and evaluation of their preservation methods. *Iranian Journal of Microbiology*, *2*(1), 38–45.

Sirbu, D. (2018). *Comprehensive Analysis of Cocoa Lipidomics - unravelling the unknown chemistry of cocoa butter*.

Smit, H. J., & Blackburn, R. J. (2005). Reinforcing effects of caffeine and theobromine as found in chocolate. *Psychopharmacology*, *181*, 101–106. https://doi.org/10.1007/s00213-005-2209-3

Stark, T., & Hofmann, T. (2005). Structures, sensory activity, and dose/response functions of 2,5-diketopiperazines in roasted cocoa nibs (*Theobroma cacao*). *Journal of Agricultural and Food Chemistry*, *53*(18), 7222–7231. https://doi.org/10.1021/jf051313m

Steensels, J., & Verstrepen, K. J. (2014). Taming Wild Yeast : Potential of Conventional and Nonconventional Yeasts in Industrial Fermentations. *Annual Reviews in Microbiology*, *68*, 61–80. https://doi.org/10.1146/annurev-micro-091213-113025

Steitz, T. A., Ludwig, M. L., Quiocho, F. A., & Lipscomb, W. N. (1967). The Structure of Carboxypeptidase A. *The Journal of Biological Chemistry*, *242*(20), 4662–4668.

Thompson, S. S., Miller, K. B., & Lopez, A. S. (2007). Cocoa and Coffee. In M. P. Doyle & L. R. Beuchat (Eds.), *In Food Microbiology: Fundamentals and Frontiers* (3rd Ed., pp. 837–850). Washington D.C.: ASM Press.

Toxopeus, H. (1985). Botany, types and populations. In G. A. . Wood & R. . Lass (Eds.), *Cocoa* (4th Ed., pp. 11–37). York: Blackwell Science.

Tran, P. D., Van de Walle, D., De Clercq, N., De Winne, A., Kadow, D., Lieberei, R., … Van Durme, J. (2015).

Assessing cocoa aroma quality by multiple analytical approaches. *Food Research International*, *77*, 657–669. https://doi.org/10.1016/j.foodres.2015.09.019

Umetrics AB. (2014). *User Guide to MODDE*. Malmö.

Vail, G. (2009). Cacao Use in Yucatán Among the Pre-Hispanic Maya. In *Chocolate: History, Culture, and Heritage* (pp. 3–15). Hoboken, New Jersey: John Wiley and Sons, Inc.

Van Lancker, F., Adams, A., & De Kimpe, N. (2012). Impact of the N-terminal amino acid on the formation of pyrazines from peptides in maillard model systems. *Journal of Agricultural and Food Chemistry*, *60*(18), 4697–4708. https://doi.org/10.1021/jf301315b

Van Lancker, F., Adams, A. N., & De Kimpe, N. (2010). Formation of pyrazines in maillard model systems of lysine-containing dipeptides. *Journal of Agricultural and Food Chemistry*, *58*(4), 2470–2478. https://doi.org/10.1021/jf903898t

Voigt, J., & Biehl, B. (1993). The major seed proteins of *Theobroma cacao* L. *Food Chemistry*, *47*, 145–151.

Voigt, J., Heinrichs, H., Voigt, G., & Biehl, B. (1994). Cocoa-specific aroma precursors are generated by proteolytic digestion of the vicilin-like globulin of cocoa seeds. *Food Chemistry*, *50*(2), 177–184. https://doi.org/10.1016/0308-8146(94)90117-1

Voigt, J., Janek, K., Textoris-Taube, K., Niewienda, A., & Wöstemeyer, J. (2016). Partial purification and characterisation of the peptide precursors of the cocoa-specific aroma components. *Food Chemistry*, *192*, 706–713. https://doi.org/10.1016/j.foodchem.2015.07.068

Voigt, J., & Lieberei, R. (2015). Biochemistry of Cocoa Fermentation. In R. F. Schwan & G. Fleet (Eds.), *Cocoa and Coffee Fermentations* (pp. 193–226). Florida: Taylor & Francis.

Voigt, J., Textoris-Taube, K., & Wöstemeyer, J. (2018). pH-Dependency of the proteolytic formation of cocoa- and nutty-specific aroma precursors. *Food Chemistry*, *255*, 209–215. https://doi.org/10.1016/j.foodchem.2018.02.045

Wadsworth, J. (1652). Chocolate or an Indian Drink.

Werber, D., Dreesman, J., Feil, F., van Treeck, U., Fell, G., Ethelberg, S., … Ammon, A. (2005). International outbreak of Salmonella Oranienburg due to German chocolate. *BMC Infectious Diseases*, *5*(7), 1–10. https://doi.org/10.1186/1471-2334-5-7

Wollgast, J., & Anklam, E. (2000). Review on polyphenols in *Theobroma cacao*: Changes in composition during the manufacture of chocolate and methodology for identification and quantification. *Food Research International*, *33*(6), 423–447. https://doi.org/10.1016/S0963-9969(00)00068-5

Wood, G. A. . (1985a). Environment. In G. A. . Wood & R. . Lass (Eds.), *Cocoa* (4th Ed., pp. 38–79). Blackwell Science.

Wood, G. A. . (1985b). History and Development. In G. A. . Wood & R. . Lass (Eds.), *Cocoa* (4th Ed., pp. 1–10). York: Blackwell Science.

Wood, G. A. ., & Lass, R. . (2001). *Cocoa* (4th ed.). Cornwall, UK: Blackwell Science.

Zhang, D., & Motilal, L. (2016). Origin, Dispersal, and Current Global Distribution of Cacao Genetic Diversity. In B. A. Bailey & L. W. Meinhardt (Eds.), *Cacao Diseases: A History of Old Enemies and New Encounters* (pp.

3–31). Springer International Publishing.

Ziegleder, G. (1990). Linalool contents as characteristic of some flavor grade cocoas. *Zeitschrift Für Lebensmittel-Untersuchung Und -Forschung*, *191*, 306–309. https://doi.org/10.1007/BF01202432

Ziegleder, G. (1991). Composition of flavor extracts of raw and roasted cocoas. *Zeitschrift Für Lebensmittel-Untersuchung Und -Forschung*, *192*, 521–525.

Ziegleder, G. (2009). Flavour Development in Cocoa and Chocolate. In S. T. Becket (Ed.), *Industrial Chocolate Manufacture and Use* (4th Ed., pp. 169–191). York: Blackwell Publishing.

Appendix

Table A1: List of naturally occurring peptides resolved by UHPLC-MS/MS.
Peptides of 4mers to 23mers were identified from positive-ion mode UHPLC-MS/MS spectra of methanol-based extracts of dry-defatted bean powder. Proteins were identified from albumin (21 kDa seed protein), vicilin, oleosin 1, chitinase class 1, lipoxygenase and maturase K. RT: retention time.

Protein	Identified Peptide	Calculated m/z	RT (min)
Albumin	IEKA	230.6419	0.9
Albumin	VPIR	242.6657	1.9
Albumin	TVWR	281.1608	4.7
Albumin	DNEW	282.1084	6.1
Albumin	WRLD	295.1583	8.9
Albumin	GVLG	345.2132	4.7
Albumin	AGVL	359.2289	5.9
Albumin	KAGV	374.2398	0.7
Albumin	TDGV	391.1823	1.1
Albumin	SSIS	393.198	1
Albumin	TGVQ	404.214	0.9
Albumin	SPVL	415.2551	6.8
Albumin	TDVN	448.2038	0.8
Albumin	PEIV	457.2657	5
Albumin	IVVQ	458.2973	4.4
Albumin	GVQY	466.2296	4.5
Albumin	VNIE	474.2558	4.4
Albumin	PVIF	475.2915	15.3
Albumin	STVW	492.2453	9.1
Albumin	DELQ	504.23	1.7
Albumin	NEWA	519.2198	3.7
Albumin	AWMF	554.2432	17
Albumin	EWAW	591.2562	16.6
Albumin	WAWM	593.2541	17
Albumin	KAGVL	244.1656	2
Albumin	TDGVK	260.1423	0.6
Albumin	KASKT	267.6659	0.5
Albumin	SAGKW	274.645	1.7
Albumin	VPIRD	300.1792	2.9
Albumin	FVPIR	316.1999	7.1
Albumin	STVWR	324.6768	5.2
Albumin	TVWRL	337.7028	15
Albumin	VWRLD	344.6925	11
Albumin	AGVLG	416.2504	5.2
Albumin	GTPVI	486.2922	6.7
Albumin	ANSPV	487.2511	2.3
Albumin	ALSDN	519.2409	1.1
Albumin	NSPVL	529.298	7
Albumin	SPVLD	530.2821	5.1
Albumin	PEIVV	556.3341	10.9
Albumin	PVIFS	562.3235	11.8
Albumin	TGVQY	567.2773	5.3
Albumin	TPVIF	576.3392	16
Albumin	DVNIE	589.2828	5.3
Albumin	DNEWA	634.2467	6.2
Albumin	SDNEW	650.2416	6.3
Albumin	NEWAW	705.2991	16.5
Albumin	KAGVLG	272.6763	1.7
Albumin	NSAGKW	331.6665	3.8
Albumin	TDVNIE	345.6689	4.2
Albumin	TVWRLD	395.2163	12.7
Albumin	DNEWAW	410.6667	16.5
Albumin	AGGGGL	431.2249	2.5
Albumin	EPGPNT	614.278	1.4
Albumin	VSTDVN	634.3042	1.9
Albumin	NSPVLD	644.325	5.4
Albumin	TDGDEL	649.2675	4.6
Albumin	DELQTG	662.2992	2.7
Albumin	LSDNEW	763.3257	9.8
Albumin	ANSPVLD	358.1847	4.3
Albumin	VTTDGVK	360.2003	1.3
Albumin	NGTPVIF	374.2054	16.1
Albumin	DNSAGKW	389.1799	5
Albumin	ALSDNEW	417.6851	8.5
Albumin	TDVNIEF	419.2031	13.5
Albumin	TVWRLDN	452.2378	9.1
Albumin	DNEWAWM	476.1869	17
Albumin	GAGGGGL	488.2463	2.9
Albumin	AGGGGLA	502.262	2.3
Albumin	GGGGLAL	544.3089	7
Albumin	GEPGPNT	671.2995	1.6
Albumin	DTDGDEL	764.2945	5
Albumin	TDGDELQ	777.3261	2
Albumin	DSKDDVVR	311.8261	1.6
Albumin	LSSISGAG	346.1847	1.5
Albumin	GEPGPNTL	392.6954	7.2
Albumin	YDNSAGKW	470.7116	6.4
Albumin	GAGGGGLA	559.2835	2.7
Albumin	DNEWAWMF	1098.435	17.7
Albumin	NYDNSAGKW	527.7331	7.2
Albumin	ALSDNEWAW	546.2433	16.8
Albumin	DNEWAWMFK	613.7686	17.1
Albumin	ISGAGGGGL	688.3624	6.1
Albumin	TDGDELQTG	935.3952	4.4
Albumin	NADSKDDVVR	373.5194	1.5
Albumin	DSKDDVVRVS	373.8595	5.1
Albumin	DTDGDELQTG	525.7147	4.7
Albumin	DNYDNSAGKW	585.2465	7.9
Albumin	ALSDNEWAWM	611.7635	17.2
Albumin	ISGAGGGGLA	759.3995	5.7
Albumin	SISGAGGGGL	775.3945	7.2
Albumin	ADSKDDVVRVS	397.5386	2.6
Albumin	SNADSKDDVVR	402.5301	1.5
Albumin	SSISGAGGGGL	431.7169	7.5
Albumin	ANSPVLDTDGD	552.2462	5.9
Albumin	LDNYDNSAGKW	641.7886	9.8
Albumin	SISGAGGGGLA	846.4316	6.7
Albumin	NADSKDDVVRVS	435.5529	5
Albumin	DTDGDELQTGVQ	639.2782	6.8
Albumin	TDGDELQTGVQY	663.2964	11
Albumin	SNADSKDDVVRVS	464.5636	5
Albumin	VTTDGVKGEPGPN	635.8173	4.7
Albumin	TDGVKGEPGPNTL	642.8251	7.9
Albumin	DTDGDELQTGVQY	720.8099	11.4
Albumin	SSISGAGGGGLAL	1046.548	15.4
Albumin	SNADSKDDVVRVST	498.2461	4.8
Albumin	VWRLDNYDNSAGKW	575.2777	16.1
Albumin	WVTTDGVKGEPGPN	728.857	9.1
Albumin	ANSPVLDTDGDELQ	737.3388	11.9
Albumin	SNADSKDDVVRVSTD	536.5884	4.6
Albumin	TVWRLDNYDNSAGKW	608.9603	16.2
Albumin	VTTDGVKGEPGPNTL	742.8832	9.1
Albumin	ANSPVLDTDGDELQT	787.8627	12.6
Albumin	NADSKDDVVRVSTDVN	578.6149	7.5
Albumin	ANSPVLDTDGDELQTG	816.3734	12.1
Albumin	ANSPVLDTDGDELQTGVQY	674.6481	16
Chitinase Class 1	PDGP	385.1718	18.7
Chitinase Class 1	FIAA	421.2445	8.7
Chitinase Class 1	DLLN	474.2558	5.1
Chitinase Class 1	FSLF	513.2708	16
Chitinase Class 1	VPGFG	476.2504	8
Chitinase Class 1	NPDLL	571.3086	4
Lipoxygenase	FELK	268.6576	5.7
Lipoxygenase	AGGI	317.1819	2.3
Lipoxygenase	GVSP	359.1925	0.9
Lipoxygenase	LGAF	407.2289	10.1
Lipoxygenase	LLPS	429.2708	5.6
Lipoxygenase	AVNF	450.2347	9
Lipoxygenase	EPLI	471.2813	10.1
Lipoxygenase	EGYF	515.2136	8
Lipoxygenase	EGKW	519.2562	5.7
Lipoxygenase	FEDL	523.2399	9.9
Lipoxygenase	LPSVK	272.1787	4.6
Lipoxygenase	NFGMY	316.1309	5.3
Lipoxygenase	FGVPG	476.2504	9.5
Lipoxygenase	PNDLV	557.293	5.8
Maturase K	AGFA	365.1819	2.4
Maturase K	IGSL	389.2395	4.5
Maturase K	LSSV	405.2344	1.5
Maturase K	GVFL	435.2602	12.3
Maturase K	VLIP	441.3071	10.1
Maturase K	DSDI	449.1878	1
Maturase K	DFLG	451.2187	8.8
Maturase K	YSLD	497.2242	4.8
Maturase K	LDFL	507.2813	11.8
Maturase K	LSHY	519.2562	6.7
Maturase K	LEEF	537.2555	12.3
Maturase K	FLFL	539.3228	17.2
Maturase K	NDSNQ	577.2212	6.6
Maturase K	FQNNL	635.3148	9.2
Maturase K	QYYFY	783.3348	9.8
Maturase K	PFLEDK	748.3876	16.3

Protein	Identified Peptide	Calculated m/z	RT (min)
Oleosin 1	KDYAME	756.3233	7.2
Peroxidase	FDNK	262.1292	0.7
Peroxidase	ISSL	419.25	2.2
Peroxidase	NLSL	446.2609	3
Peroxidase	ISEAD	534.2406	1.8
Peroxidase	TAFEI	580.2977	8
Peroxidase	NFDPT	593.2566	3.7
Vicilin	AKPE	222.6263	0.6
Vicilin	KLVD	237.6498	1.1
Vicilin	ILPH	240.1525	3.6
Vicilin	NGRF	247.1295	1.6
Vicilin	VPHY	258.1343	3.9
Vicilin	SQQR	259.6377	0.6
Vicilin	QRFA	261.1452	2.8
Vicilin	FKLN	261.1577	5.9
Vicilin	QTRF	276.1504	2.4
Vicilin	KYEL	276.655	4.5
Vicilin	QRIF	282.1686	6.8
Vicilin	DYRL	283.6503	7.1
Vicilin	VAPA	357.2132	1.1
Vicilin	QGAI	388.2191	2.2
Vicilin	GTIT	391.2187	1
Vicilin	FLAG	407.2289	2.6
Vicilin	SPPL	413.2395	5.2
Vicilin	LTIA	417.2708	3.3
Vicilin	GIND	418.1932	1.2
Vicilin	FGVP	419.2289	10.3
Vicilin	LSFG	423.2238	8
Vicilin	HAVT	427.23	0.7
Vicilin	QSPV	430.2296	1.9
Vicilin	LAIN	430.266	9.3
Vicilin	MDVA	435.1908	2.6
Vicilin	GDVF	437.2031	7.7
Vicilin	LPVN	442.266	4.3
Vicilin	FGLN	450.2347	7
Vicilin	TFGE	453.198	4
Vicilin	ETFG	453.198	4.3
Vicilin	YGAF	457.2082	8.1
Vicilin	LVDN	460.2402	1.4
Vicilin	VLET	461.2606	3.7
Vicilin	SPVY	465.2344	4.8
Vicilin	EGNF	466.1932	3.4
Vicilin	STVY	469.2293	4
Vicilin	DQPL	472.2402	2.8
Vicilin	IFVP	475.2915	11.8
Vicilin	GEFQ	480.2089	2
Vicilin	TVFN	480.2453	8.2
Vicilin	SYYG	489.198	4
Vicilin	EVLE	489.2555	4.4
Vicilin	FIVL	491.3228	16.8
Vicilin	EAIY	495.2449	6.1
Vicilin	ETVF	495.2449	8.5
Vicilin	DFSQ	496.2038	2
Vicilin	FGEF	499.2187	11.8
Vicilin	LEEI	503.2712	6.1
Vicilin	EEIL	503.2712	9
Vicilin	QNMD	507.1868	1
Vicilin	IFNN	507.2562	4.7
Vicilin	NIFN	507.2562	8.6
Vicilin	DNIF	508.2402	10.6
Vicilin	FVVF	511.2915	16.8
Vicilin	DESY	513.1827	1.9
Vicilin	VTFF	513.2708	14.1
Vicilin	YEVL	523.2762	9.1
Vicilin	INDY	524.2351	1.8
Vicilin	EETF	525.2191	5.3
Vicilin	FMSF	531.2272	16.1
Vicilin	FQNM	539.2282	4.4
Vicilin	FSYE	545.2242	6.6
Vicilin	ESYF	545.2242	9.7
Vicilin	ELFF	555.2813	16.8
Vicilin	NDYR	567.2522	4.5
Vicilin	SPPLK	271.1709	1.4
Vicilin	SPGKY	276.1448	1.1
Vicilin	KAKPE	286.6738	0.5
Vicilin	KLVDN	294.6712	1.5
Vicilin	KDQPL	300.6712	2.5
Vicilin	IFVPH	306.6788	7.9
Vicilin	QNGRF	311.1588	1.9
Vicilin	FILPH	313.6867	11.5
Vicilin	VPHYN	315.1557	2.4
Vicilin	NFKIL	317.6998	10.9
Vicilin	VRQMD	324.6603	1.1
Vicilin	FKLNQ	325.187	4.8
Vicilin	NQRIF	339.1901	7.1
Vicilin	INDYR	340.6717	1.8
Vicilin	KYELF	350.1892	11.8
Vicilin	VAVSA	446.2609	3.8
Vicilin	GTVVS	462.2558	1.8
Vicilin	AVAFG	464.2504	7.2
Vicilin	SPGDV	474.2195	1.9

Protein	Identified Peptide	Calculated m/z	RT (min)
Vicilin	VAFGL	506.2973	16.2
Vicilin	SQSPV	517.2617	2.3
Vicilin	GSTVY	526.2508	4.2
Vicilin	NSPPL	527.2824	5.6
Vicilin	QGAIF	535.2875	10.9
Vicilin	ENSPP	543.2409	1.1
Vicilin	NMDVA	549.2337	3.2
Vicilin	DVFVA	550.2871	8.8
Vicilin	VTDGY	554.2457	3.3
Vicilin	AEAIY	566.2821	6.7
Vicilin	DGYGY	574.2144	5.5
Vicilin	FEANP	577.2617	6.5
Vicilin	GINDY	581.2566	5.4
Vicilin	ETFGE	582.2406	4.9
Vicilin	DQPLN	586.2831	2.1
Vicilin	EVLET	590.3032	5.6
Vicilin	NPNTF	592.2726	3.3
Vicilin	QSPVY	593.293	5.2
Vicilin	EEGNF	595.2358	4.4
Vicilin	TFGEF	600.2664	14.4
Vicilin	VDNIF	607.3086	13.3
Vicilin	ETVFN	609.2879	6.4
Vicilin	QQGMF	610.2654	8.7
Vicilin	SYNVQ	610.2831	3.4
Vicilin	SYEVL	610.3083	11
Vicilin	AFSYE	616.2613	6.4
Vicilin	DNIFN	622.2831	8.3
Vicilin	EDFSQ	625.2464	3.2
Vicilin	DFSQF	643.2722	12.2
Vicilin	FQNMD	654.2552	4.9
Vicilin	EEETF	654.2617	5.8
Vicilin	DESYF	660.2511	10.3
Vicilin	NNPYY	670.2831	6.4
Vicilin	AGHAVT	278.1479	0.8
Vicilin	RGTVVS	309.6821	1.2
Vicilin	SPGDVF	311.1476	7.5
Vicilin	FGVPSK	317.6816	5.7
Vicilin	NSPPLK	328.1923	1.6
Vicilin	GQQGMF	334.147	6.5
Vicilin	SEAKEL	338.6792	2.2
Vicilin	SPGKYE	340.6661	1.1
Vicilin	AIFVPH	342.1974	9.5
Vicilin	SKDQPL	344.1872	3.1
Vicilin	DEEGNF	355.635	3.4
Vicilin	KDQPLN	357.6927	1.3
Vicilin	VPHYNS	358.6717	2.7
Vicilin	EEGNFK	362.169	0.7
Vicilin	NQNGRF	368.1803	1.5
Vicilin	GINDYR	369.1825	4.5
Vicilin	IFVPHY	388.2105	13.5
Vicilin	NNQRIF	396.2116	7.2
Vicilin	DVAVSA	561.2879	4.6
Vicilin	AVAFGL	577.3344	16.4
Vicilin	NAVAFG	578.2933	9.1
Vicilin	PLSPGD	585.2879	3.4
Vicilin	SFGVPS	593.293	9.2
Vicilin	VTDGYG	611.2671	3
Vicilin	DGYGYA	645.2515	5.7
Vicilin	DVFVAP	647.3399	10.9
Vicilin	NQGAIF	649.3304	10.8
Vicilin	ENSPPL	656.325	6.1
Vicilin	ANPNTF	663.3097	3.7
Vicilin	FEANPN	691.3046	3.5
Vicilin	SYYGAF	707.3035	13.9
Vicilin	EETFGE	711.2832	5.5
Vicilin	IFNNPD	719.3359	6.5
Vicilin	VDNIFN	721.3515	11
Vicilin	NPDESY	724.2784	3.7
Vicilin	ETFGEF	729.309	15.5
Vicilin	ESYNVQ	739.3257	4.3
Vicilin	SYEVLE	739.3509	9.8
Vicilin	EDFSQF	772.3148	12.8
Vicilin	NGKGTIT	345.6927	1
Vicilin	AENSPPL	364.1847	6.5
Vicilin	GAIFVPH	370.7081	8.5
Vicilin	FGVPSKL	374.2236	12.7
Vicilin	VPSKLVD	379.2264	5.6
Vicilin	SEAKELS	382.1953	1.4
Vicilin	FKLNQGA	389.2163	6.4
Vicilin	QQVKAPL	392.2398	6.4
Vicilin	ENSPPLK	392.7136	2.8
Vicilin	SKDQPLN	401.2087	1.6
Vicilin	SNQNGRF	411.6963	2.1
Vicilin	GQQGMFR	412.1976	5.4
Vicilin	DEEGNFK	419.6825	3
Vicilin	KLVDNIF	424.7475	14.7
Vicilin	GINDYRL	425.7245	11.2
Vicilin	RDEEGNF	433.6856	4.2
Vicilin	KESYNVQ	434.214	2.6
Vicilin	QDNQEKL	437.7169	2.6
Vicilin	IFVPHYN	445.232	11.2

Protein	Identified Peptide	Calculated m/z	RT (min)
Vicilin	NPYYFPK	464.7318	10.4
Vicilin	APLSPGD	656.325	4.6
Vicilin	DVFVAPA	718.377	11.4
Vicilin	ASKDQPL	758.4043	3.4
Vicilin	DGYGYAQ	773.3101	5.1
Vicilin	VTDGYGY	774.3305	6.7
Vicilin	FEANPNT	792.3523	4.6
Vicilin	EANPNTF	792.3523	5.9
Vicilin	NIFNNPD	833.3788	10.3
Vicilin	EEETFGE	840.3258	6
Vicilin	SYEVLET	840.3985	10.7
Vicilin	IFNNPDE	848.3785	7
Vicilin	NPDESYF	871.3468	11.1
Vicilin	NGKGTITF	419.2269	9
Vicilin	AENSPPLK	428.2322	3.3
Vicilin	PVNSPGKY	431.2269	4.1
Vicilin	VPSKLVDN	436.2478	5.2
Vicilin	ASKDQPLN	436.7272	1.9
Vicilin	SKDQPLNA	436.7272	3.7
Vicilin	DSEAKELS	439.7087	2.4
Vicilin	FASKDQPL	453.24	7
Vicilin	GINDYRLA	461.2431	10.3
Vicilin	SPGKYELF	470.7424	15.5
Vicilin	DEEGNFKI	476.2245	11.4
Vicilin	NAQNNQRI	479.2467	1.7
Vicilin	KLVDNIFN	481.7689	12.4
Vicilin	QDNQEKLT	488.2407	2.1
Vicilin	IFVPHYNS	488.748	11.1
Vicilin	FNNPDESY	493.1985	5.4
Vicilin	YSNQNGRF	493.2279	4.5
Vicilin	AQNNQRIF	495.7594	7.5
Vicilin	NNPYYFPK	521.7533	10.8
Vicilin	APLSPGDV	755.3934	9.3
Vicilin	VPAGSTVY	793.409	6.6
Vicilin	VTDGYGYA	845.3676	6.8
Vicilin	IFNNPDES	935.4105	6.8
Vicilin	IFVPHYNSK	368.8661	8.6
Vicilin	RNNPYYFPK	400.205	8.4
Vicilin	VAPAGHAVT	411.7271	2.5
Vicilin	VKAPLSPGD	442.2478	6.2
Vicilin	VTNGKGTIT	445.7507	1.6
Vicilin	APLSPGDVF	451.7345	16.3
Vicilin	GVPSKLVDN	464.7585	6.3
Vicilin	ASKDQPLNA	472.2458	4
Vicilin	FGVPSKLVD	481.2713	14.3
Vicilin	NQGAIFVPH	491.7589	10.4
Vicilin	PVNSPGKYE	495.7482	4.4
Vicilin	FASKDQPLN	510.2615	4.1
Vicilin	VDNIFNNPD	524.2407	12.1
Vicilin	DEEGNFKIL	532.7666	16
Vicilin	IFNNPDESY	549.7406	11
Vicilin	NAQNNQRIF	552.7809	7.6
Vicilin	DFSQFQNMD	566.2242	14.8
Vicilin	VTDGYGYAQ	973.4262	6.2
Vicilin	IFVPHYNSKA	392.5451	9.3
Vicilin	VAPAGHAVTF	485.2613	9.4
Vicilin	VTNGKGTITF	519.2849	9.7
Vicilin	FGVPSKLVDN	538.2928	13.3
Vicilin	FASKDQPLNA	545.78	6.5
Vicilin	PVNSPGKYEL	552.2902	9.6
Vicilin	SQSPVYSNQN	562.2544	4.1
Vicilin	QGAIFVPHYN	573.2905	14.2
Vicilin	NQGAIFVPHY	573.2905	15.7
Vicilin	DEEGNFKILQ	596.7959	14.5
Vicilin	NIFNNPDESY	606.762	15.6
Vicilin	IFNNPDESYF	623.2748	16.4
Vicilin	EDFSQFQNMD	630.7455	12.2
Vicilin	IFVPHYNSKAT	426.2277	9.2
Vicilin	VAPAGHAVTFF	558.7955	16.2
Vicilin	VKAPLSPGDVF	565.3162	16.2
Vicilin	SFGVPSKLVDN	581.8088	14.4
Vicilin	FVTNGKGTITF	592.8191	15.6
Vicilin	PVNSPGKYELF	625.8244	16
Vicilin	EANPNTFILPH	626.8197	15.6
Vicilin	VVSQDNQEKLT	630.8251	2.5
Vicilin	GLNAQNNQRIF	637.8336	10.9
Vicilin	DNIFNNPDESY	664.2755	15.4
Vicilin	NIFNNPDESYF	680.2962	16.8
Vicilin	AENSPPLKGIND	627.8199	7.7
Vicilin	QVKAPLSPGDVF	629.3455	16.2
Vicilin	LSFGVPSKLVDN	638.3508	16.1
Vicilin	ESYNVQRGTVVS	669.836	7
Vicilin	QSPVYSNQNGRF	698.8338	8
Vicilin	PVNSPGKYELFF	699.3586	16.8
Vicilin	FEANPNTFILPH	700.3539	16.3
Vicilin	VDNIFNNPDESY	713.8097	16
Vicilin	DNIFNNPDESYF	737.8097	16.4
Vicilin	ETVFNTQREKLEE	541.6055	7.6
Vicilin	QQVKAPLSPGDVF	693.3748	16.1
Vicilin	AENSPPLKGINDY	709.3515	11.4
Vicilin	SQSPVYSNQNGRF	742.3499	8.2

Protein	Identified Peptide	Calculated m/z	RT (min)
Vicilin	AENSPPLKGINDYR	525.2705	7.4
Vicilin	ASKDQPLNAVAFGL	715.8855	16.6
Vicilin	FQQVKAPLSPGDVF	766.909	16.4
Vicilin	LSQSPVYSNQNGRF	798.8919	9.8
Vicilin	KLVDNIFNNPDESY	834.3992	16.2
Vicilin	QGAIFVPHYNSKATF	560.6229	14.2
Vicilin	AENSPPLKGINDYRL	562.9652	15.7
Vicilin	FASKDQPLNAVAFGL	789.4197	16.8
Vicilin	KLVDNIFNNPDESYF	907.9334	16.7
Vicilin	NQGAIFVPHYNSKATF	598.6372	15.9
Vicilin	SEAKELSFGVPSKLVDN	607.3194	14.8
Vicilin	GVPSKLVDNIFNNPDESY	669.9935	16.5
Vicilin	FGVPSKLVDNIFNNPDESY	719.0163	16.9
Vicilin	APLSPGDVFVAPAGHAVTF	926.9832	16.8
Vicilin	FGVPSKLVDNIFNNPDESYF	768.0391	17.2
Vicilin	QQVKAPLSPGDVFVAPAGHAVT	730.3953	16.1
Vicilin	LSFGVPSKLVDNIFNNPDESYF	834.7444	17.5
Vicilin	QQVKAPLSPGDVFVAPAGHAVTF	779.4181	16.6
Vicilin	ELSFGVPSKLVDNIFNNPDESYF	877.7586	17.5

Table A2: List of experiments generated by D-optimal design on MODDE Pro 11.
A list of 47 experiments was generated incorporating five quantitative factors and their interaction terms in first- second- and third-order. Experiments N45, N46 and N47 are biological replicates.

Experiment Name	Temperature [°C]	Acetic Acid [g/L]	Lactic Acid [g/L]	Ethanol [g/L]	Incubation Time [days]
N1	60	0	25	200	0.5
N2	30	25	0	200	0.5
N3	25	5	5	20	0.5
N4	60	0	5	0	0.5
N5	25	25	20	20	0.5
N6	25	25	5	140	0.5
N7	25	15	25	200	0.5
N8	60	25	0	66.7	0.5
N9	60	25	6.7	200	0.5
N10	60	5	0	133.3	0.5
N11	60	13.3	25	0	0.5
N12	40	0	25	0	0.5
N13	30	0	25	133.3	0.5
N14	40	0	5	200	0.5
N15	40	25	0	0	0.5
N16	50	25	20	200	0.5
N17	40	12.4	12.4	97.4	0.5
N18	60	25	20	0	1
N19	25	5	18.3	200	1
N20	60	0	25	66.7	1
N21	60	18.3	0	200	1
N22	30	11.7	0	0	1
N23	30	0	25	0	1.5
N24	25	25	5	20	1.5
N25	60	5	0	0	1.5
N26	25	25	20	200	1.5
N27	60	25	20	133.3	1.5
N28	50	0	25	200	1.5
N29	30	5	0	133.3	1.5
N30	60	25	0	0	2
N31	60	0	25	0	2
N32	30	0	25	200	2
N33	60	0	5	200	2
N34	30	0	5	0	2
N35	60	25	20	200	2
N36	25	25	20	80	2
N37	25	10	25	20	2
N38	25	18.3	5	200	2
N39	60	13.3	25	200	2
N40	40	25	0	200	2
N41	30	25	13.3	0	2
N42	50	20	25	0	2
N43	60	12.5	12.5	100	2
N44	45	0	15	100	2
N45	40	12.4	12.4	97.4	1.5
N46	40	12.4	12.4	97.4	1.5
N47	40	12.4	12.4	97.4	1.5

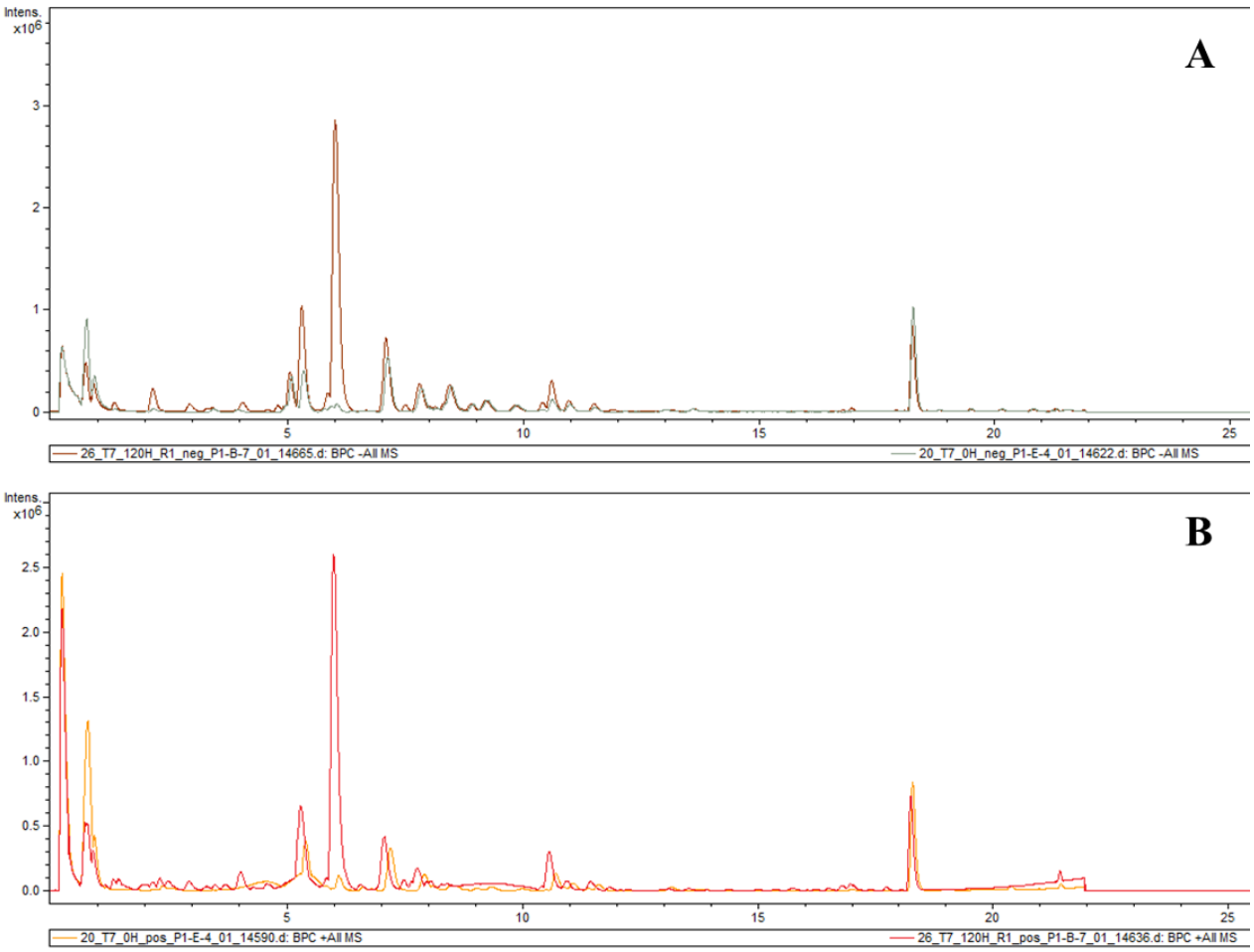


Figure A1: Base peak chromatogram of methanol extracts of cocoa in Forced Fermentation.
Negative mode (A) and positive mode (B) spectra are shown for methanol extracts of Trial A of the forced fermentation system. 0h (brown line) and 120h (green line) are shown for negative mode along with 0h (yellow line) and 120h (red line) are shown for positive mode demonstrating the dramatic increase in detectable compounds during the course of fermentation.

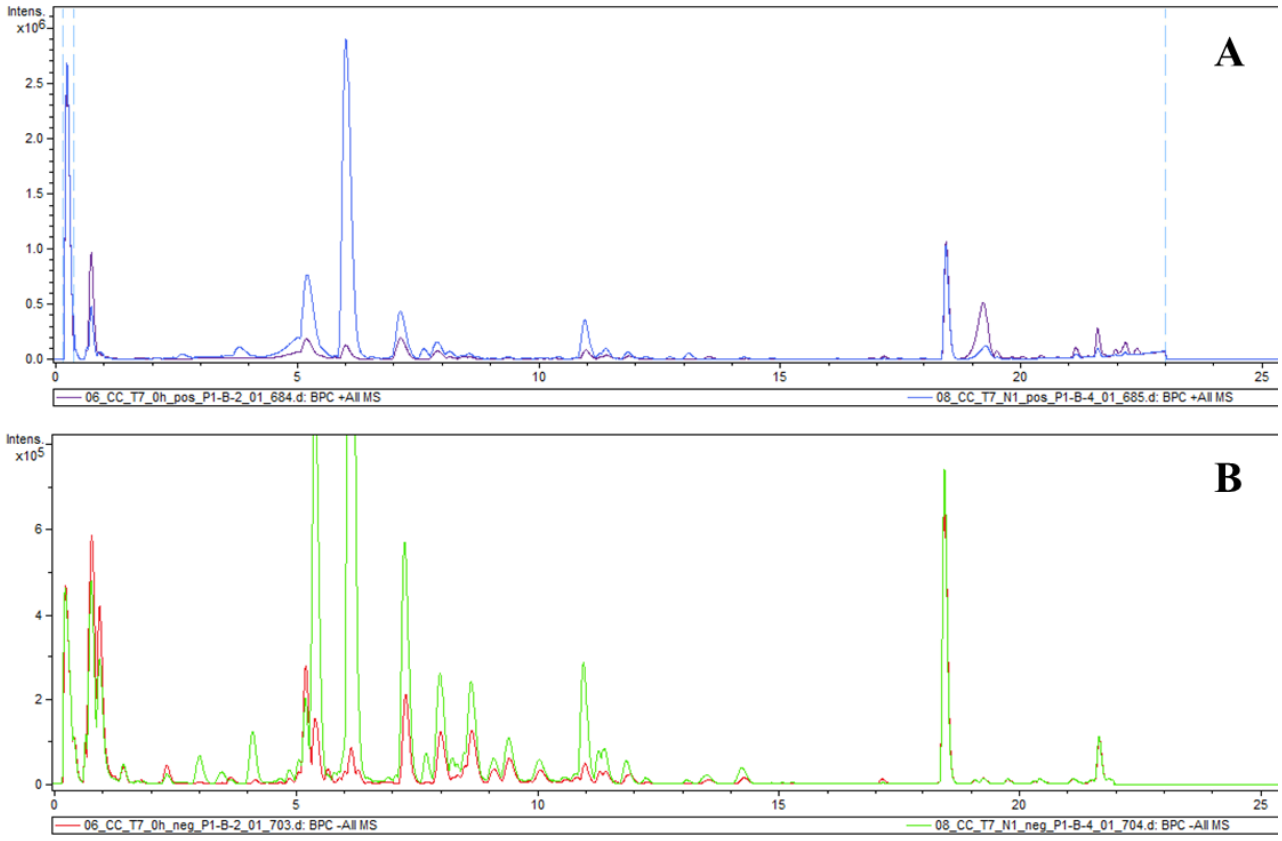


Figure A2: Base peak chromatogram of methanol extracts of cocoa in Artificial Fermentation.
Positive mode (A) and negative mode (B) spectra are shown for methanol extracts of Artificial Fermentation system outlined in chapter 5. 0h (purple line) and expt. N1 (blue line) are shown for positive mode along with 0h (red line) and expt. N1 (green line) are shown for negative mode demonstrating the dramatic increase in detectable compounds during the course of bean incubation in the specified conditions in Table A2.