UNIVERSIDAD AUTÓNOMA CHAPINGO

DEPARTAMENTO DE INGENIERÍA

AGROINDUSTRIAL

POSGRADO EN CIENCIA Y TECNOLOGÍA AGROALIMENTARIA

OBTENCIÓN DE MOLÉCULAS PLATAFORMA POR METODOLOGÍAS VERDES, A PARTIR DE RESIDUOS DE CÁSCARA DE CACAO

TESIS

Que como requisito parcial para obtener el grado de:

DOCTORA EN CIENCIAS AGROALIMENTARIAS

Presenta:

EDNA ELENA SUÁREZ PATLÁN

Bajo la supervisión de: **HOLBER ZULETA PRADA, DR.**

Chapingo, Estado de México, mayo de 2024.

UNIVERSIDAD AUTÓNOMA CHAPINGO

DEPARTAMENTO DE INGENIERÍA

AGROINDUSTRIAL

POSGRADO EN CIENCIA Y TECNOLOGÍA AGROALIMENTARIA

PRODUCTION OF PLATFORM MOLECULES FROM COCOA POD HUSK BY GREEN METHODS

THESIS

That is a partial requirement to obtain the degree of:

DOCTORA EN CIENCIAS AGROALIMENTARIAS

Submit:

EDNA ELENA SUÁREZ PATLÁN

Supervised by: **HOLBER ZULETA PRADA, DR.**

Chapingo, Estado de México, may 2024.

OBTENCIÓN DE MOLÉCULAS PLATAFORMA POR METODOLOGÍAS VERDES, A PARTIR DE RESIDUOS DE CÁSCARA DE CACAO

Dissertation written by EDNA ELENA SUÁREZ PATLÁN under the direction of the Committee of Advisors approved and accepted as a partial requirement for the degree of:

DOCTORA EN CIENCIAS AGROALIMENTARIAS

Director:

Holber Zuleta Prada Doctor

Codirector:

Teodoro Espinosa Solares Doctor

Advisor:

José Enrique Herbert Pucheta Doctor

External reader:

Emanuel Hernández Núñez Doctor

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

APPENDIX LIST

The standards of pentoses, hexoses, and disaccharide sucrose are separated and identified in Hotelling ellipses (which define a confidence level of 95%). Supervised methods such as partial least squares discriminant analysis (PLS-DA) were applied to increase the discriminant capacity of the NIR-MIR matrix under study. This discriminant method applies PLS regressions from which possible class members can be predicted, and the discriminations are evaluated with permutation tests and cross-validations (Alamar et al., 2020). Figure 13B shows the PLS-DA showing a total variance of 60.2 %, with 53 % for CP1 and 7.2 % for CP2. These analyses allowed greater separation within the group of pretreatments and polymers. However, there is no significant discrimination in this group. The standards for pentoses, hexoses, and sucrose remain bounded by Hotelling ellipses (at a 95% confidence level) except for arabinose, which appears not to discriminate effectively from the pretreatment group. Sparse partial least squares discriminant analysis (sPLS-DA) is a variant of PLS-DA that performs variable selection and classification during model calibration and discards uninformative variables. The sPLS algorithm efficiently estimates many variables (Jimenez Carvelo et al., 2021). Figure 13C shows the sPLS-DA discriminant analysis with an overall variance of 21.3 %, considering 12.5 % for PC1 and 8.8 % for PC2. This analysis shows a remarkable increase in the separation of pretreatments and polymers within the group, delimiting each pretreatment and polymer by species. Likewise, a closer approach of the galactose, arabinose, ribose, xylose, and mannose standards to the group of pretreatments and polymers is observed. Finally, the discriminant analysis to orthogonal projections to latent structures (OPLS-DA) is a supervised model that provides dimension reduction and identification of spectral features due to its ability to distinguish between relevant and irrelevant data variations within a group, which also allows

ABBREVIATION LIST

CPH	Cocoa Pod Husk
CBS	Cocoa Bean Shell
PM	Platform Molecules
TAPPI	Technical Association of the Pulp and Paper Industry
ASTM	American Society for Testing and Materials
ABTS	(2,2'-azinobis [3-etilbenzotiazolina-6-ácido sulfónico]-sal de diamonio)
mg GAE/g	mg Gallic Acid Equivalents/g
mg CE/g	mg Catechin Equivalents/g
µmol TE/g	µmol Trolox Equivalent/g
PEH	Pretreatment Extraction Hydrothermal
MA-HTE	Microwave-Assisted Hydrothermal Extraction
MA-HTP	Microwave-Assisted Hydrothermal Pretreatment
NIR	Near-Infrared
MIR	Mid-Infrared
FT-MIR	Fourier Transform Mid-Infrared
FTIR-ATR	Fourier Transform Infrared - Attenuated Total Reflection
NMR	Nuclear Magnetic Resonance
^{13}C NMR	Carbon-13 Nuclear Magnetic Resonance Spectroscopy
1D-NMR	One Dimensional - Nuclear Magnetic Resonance
2D-NMR	Two-dimensional - Nuclear Magnetic Resonance
^{1}H NMR	Proton Nuclear Magnetic Resonance
1D-^{1}H-DP	One Dimensional - Proton - Direct Polarization Scheme
1D-^{1}H-water suppression	Noesy Presaturation Scheme

MSA	Multivariate Statistical Analysis
PCA	Principal Component Analysis
PLS-DA	Partial Least Squares - Discriminant Analysis
sPLS-DA	Sparse Partial Least Square - Discriminant Analysis
OPLS-DA	Orthogonal Projections to Latent Structures - Discriminant Analysis
HSQC	Heteronuclear Single Quantum Correlation
TOCSY	Totally correlated Spectroscopy
J-Res	Homonuclear J-Resolved Spectroscopy
HMDB	Human Metabolome Database
DES	Deep Eutectic Solvents
NSC	Non-Structural Carbohydrates

ACKNOWLEDGMENTS

To the Consejo Nacional de Humanidades, Ciencias y Tecnología (CONAHCYT) for Postgraduate Scholarship No. 763652.

To the Universidad Autónoma Chapingo and the Posgrado en Ciencia y Tecnología Agroalimentaria.

The Laboratory of Natural Products of the Departamento de Preparatoria Agrícola of the Universidad Autónoma Chapingo.

To the Geoquímica Marina and LANNBIO laboratories, CINVESTAV-IPN, Unidad Mérida.

To Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP).

To C. Carlos Hernández Echeverría for providing the raw material for the study.

Dr. Holber Zuleta Prada, Dr. José Enrique Herbert Pucheta, and Dr. Teodoro Espinosa Solares are commended for their research guidance.

To Dr. Emanuel Hernández Núñez for his guidance during the research stay at CINVESTAV-IPN, Unidad Mérida.

We thank Dr. Benito Reyes Trejo and Dra for their collaboration. Diana Guerra Ramírez.

Valeria, Josué, Carlos, and Consuelo, thank you for your invaluable collaboration in the laboratory.

To my teachers, lab mates, classmates, and friends inside and outside the graduate program for your friendship and invaluable support.

DEDICATORY

To my beloved family

BIOGRAPHY

Personal Information

Name:	Edna Elena Suárez Patlán
Date of Birth:	May 18th, 1977
Birthplace:	Morelia, Michoacán, México
CURP:	SUPE7705184SA
ORCID	0000-0002-7883-625X

Employment Data

Researcher to the Instituto Nacional de Investigaciones Forestales, Agrícolas y pecuarias (INIFAP), Centro de Investigación Regional Golfo Centro, Campo Experimental San Martinito, Puebla. (2012 to the present).

Technical-Scientific Collaborator at the Center for Research on Properties and Uses of Wood, Departamento de Madera, Celulosa y Papel, Centro Universitario de Ciencias Exactas e Ingenierías, Universidad de Guadalajara. Zapopan, Jalisco-México. (2009-2012).

Research Assistant. Forestry Area of Crisoba Industrial, S.A. de C.V. Morelia, Michoacán-México. (2004-2005).

Academic Progress

Master: Universidad de Guadalajara, Departamento de Madera, Celulosa y Papel. Professional license: 11837628

Bachelor's Degree: Universidad Michoacana de San Nicolás de Hidalgo, Facultad de Ingeniería en Tecnología de la Madera. Professional license: 4792809

RESUMEN GENERAL

OBTENCIÓN DE MOLÉCULAS PLATAFORMA POR METODOLOGÍAS VERDES, A PARTIR DE RESIDUOS DE CÁSCARA DE CACAO

En el mundo se producen anualmente alrededor de 48 millones de toneladas de residuos de la cosecha del cacao, estos se pueden aprovechar para la obtención de moléculas plataforma (PM), estas sirven como base para la obtención de productos de valor agregado. El objetivo de este trabajo fue obtener y elucidar PM mediante pretratamiento hidrotérmico asistido por microondas (MA-HTP) y la combinación de herramientas estadísticas y espectroscópicas, como alternativa para el aprovechamiento de cáscara de mazorca de cacao (CPH). Se determinaron componentes lignocelulósicos, capacidad antioxidante y perfil mineral de CPH. Se evaluaron los MA-HTP mediante metabolómica dirigida y no dirigida MIR-NIR y RMN, se determinaron las PM a partir de RMN de una y dos dimensiones. Se optimizó el MA-HTP por análisis de superficie de respuesta Box Behnken (RSA-BBD), se elucidó y cuantificó por ^{1}H-RMN las PM obtenidas en los extractos de CPH. Como resultado se obtuvo un contenido de celulosa, hemicelulosa y lignina de 30.5 %, 19 % y 23.9 % respectivamente, 7.7 % de cenizas con un perfil de 13 minerales; Fe, Ba y Zn, con 164.4 µg/g, 63.5 µg/g y 63.4 µg/g fueron los más abundantes. Se obtuvo un contenido fenólico: 24.5 mg EAG/g, capacidad antioxidante ABTS: 304.8 µmol ET/g y FRAP: 145.8 µmol ET/g. Mediante metabolómica MIR-NIR y NMR y OPLS-DA se determinó que el MA-HPT de 200°C se agrupó con los estándares de mono- y disacáridos. De la elucidación por RMN se identificaron: Alosa, Celobiosa, D-fructosa, D-galactosa, Fructosa 6-fosfato, N-acetilmanosamina, Gluconolactona, Ácido glucónico, Glicerol, Glicerol 3-fosfato, Citidina, D-maltosa, D-xilosa y Trehalosa. Mediante RSA-BBD, se predijeron puntos óptimos de temperatura, potencia y tiempo, para la extracción de glucosa: 135.4°C-180.6 W y 5.8 min; sacarosa: 154.3°C-256.3 W y 20.2 min y fructosa de CPH: 129.5°C-173.8 W y 5.27 min. Para glucosa: 142.2°C-204.4 W y 10.5 min; sacarosa: 148.8°C-215.6 W y 14.3 min y fructosa de HMC-CPH: 151.6°C-231.6 W y 13 min.

Palabras clave: Pretratamiento hidrotérmico asistido por microondas (MA-HTP), infrarrojo cercano y medio (NIR-MIR), resonancia magnética nuclear (NMR), análisis de superficie de respuesta (RSA), diseño Box Behnken (BBD).

Tesis: Doctorado en Ciencias Agroalimentarias, Universidad Autónoma Chapingo.
Autor: M.C. Edna Elena Suárez Patlán.
Director de Tesis: Dr. Holber Zuleta Prada

GENERAL ABSTRACT

PRODUCTION OF PLATFORM MOLECULES FROM COCOA POD HUSK BY GREEN METHODS

Approximately 48 million tons of cocoa crop residues are produced annually, which can be used to obtain platform molecules (PM) as a base to produce value-added products. This work aimed to obtain and elucidate PM by microwave-assisted hydrothermal pretreatment (MA-HTP) and combining statistical and spectroscopic tools as an alternative to cocoa pod husks (CPH). Lignocellulosic components, antioxidant capacity, and mineral profile of CPH were determined. MA-HTPs were evaluated by targeted and non-targeted metabolomics MIR-NIR and NMR, and PM was determined by one- and two-dimensional NMR. MA-HTP was optimized by Box Behnken Response Surface Analysis (RSA-BBD), and PM obtained in CPH extracts was elucidated and quantified by ^{1}H-RMN. As a result, cellulose, hemicellulose, and lignin contents of 30.5%, 19%, and 23.9% were obtained, respectively, and 7.7% ash with a profile of 13 minerals were obtained; Fe, Ba, and Zn with 164.4 µg/g, 63.5 µg/g, and 63.4 µg/g were the most abundant. Phenolic content: 24.5 mg EAG/g, antioxidant capacity ABTS: 304.8 µmol ET/g and FRAP: 145.8 µmol ET/g. MIR-NIR NMR and OPLS-DA metabolomics showed that the 200°C MA-HPT clustered with the mono - and disaccharide standards. NMR elucidation identified Allose, Cellobiose, D-fructose, D-galactose, Fructose 6-phosphate, N-acetylmanosamine, Gluconolactone, Gluconic acid, Glycerol, Glycerol 3-phosphate, Cytidine, D-maltose, D-xylose, and Trehalose. RSA-BBD was used to predict the optimum temperature, power, and time for the extraction of glucose: 135.4°C-180.6 W and 5.8 min; sucrose: 154.3°C-256.3 W and 20. 2 min and fructose from CPH: 129.5°C-173.8 W and 5.27 min. For glucose: 142.2°C-204.4 W and 10.5 min; for sucrose, 148.8°C-215.6 W and 14.3 min and fructose from HMC-CPH: 151.6°C-231.6 W and 13 min.

Keywords: Microwave-assisted hydrothermal pretreatment (MA-HTP), near and mid-infrared (NIR-MIR), nuclear magnetic resonance (NMR), response surface analysis (RSA), Box Behnken design (BBD).

Doctoral Thesis: Doctorado en Ciencias Agroalimentarias, Universidad Autónoma Chapingo.
Author: Edna Elena Suárez Patlán, M. C.
Advisor: Holber Zuleta Prada, Dr.

GENERAL INTRODUCTION

Our planet's limited fossil resources have become increasingly unsustainable, and their production and use have caused major environmental impacts with significant contributions to climate change (Coma et al., 2017; Dahunsi et al., 2019). On the other hand, agro-industrial activities generate large amounts of waste; these wastes are lignocellulosic materials that, if not properly managed, can cause environmental problems, as their disposal and conversion are costly due to high energy costs and lack of appropriate technologies (Fong Lopez et al., 2019; Kumari & Singh, 2018; Zheng et al., 2017). However, lignocellulosic biomass is a potential material for chemical, cosmetic, food, pharmaceutical, and other industries. One example is the hydrolysis and fermentation of lignocellulosic materials to replace fossil fuels. Its feasibility lies in the fact that it is a globally available material, has a low cost, and does not compete with human food (Batista Meneses et al., 2022; Cornejo et al., 2019). Globally, the generation of lignocellulosic waste is approximately 220 million tons annually (Kumari & Singh, 2018). Two thousand seventeen more than 37 million tons of agro-industrial waste were recorded worldwide (Batista Meneses et al., 2022). Of these, approximately 48 million tons are residues from the cocoa industry (Quiceno Suarez et al., 2024a). This lignocellulosic material comprises two main carbohydrate polymers (cellulose and hemicellulose) and an aromatic polymer (lignin). It is also rich in fiber, protein, and fatty acids (oleic, linoleic, elaidic, and stearic), its volatile compounds are composed of pyrazines, esters, alcohols, and aldehydes, it is rich in polyphenols, epicatechin, theobromine, and caffeine (Araya Farias et al., 2019; Galbe & Wallberg, 2019; Quiceno Suarez et al., 2024a; Vasquez et al., 2019).

However, lignocellulosic material is highly interconnected, so the release of one of its components requires the removal of the others, which complicates its use (Gatt et al., 2018). Microorganisms or chemical processes can convert this renewable resource into high-value chemicals. A limiting factor for converting and utilizing its compounds is the pretreatments, many of which are non-selective and use high temperatures and polluting catalysts (Song et al., 2020). Therefore, it is imperative to search for environmentally friendly pretreatments to obtain these molecules from a feedstock that does not compete with food and is readily

available, such as cocoa pod husk. Although this material is a sustainable energy resource, its cell wall structure is complex, with a high degree of polymerization, cellulose crystallinity, lignification, and heterogeneity in composition, even within the same species (Ahmed et al., 2019). For this reason, it must be pretreated before conversion into chemicals and biofuels to improve its digestibility and reduce its recalcitrance (Araya Farias et al., 2019; Climent et al., 2014).

Pretreatment of lignocellulosic material is the step before depolymerization and extraction of sugars and other compounds to obtain chemical and fuel products. Some conventional treatments for lignocellulosic materials include sudden steam explosion, organosolv process, hydrothermal treatment, and acid and alkaline treatments. However, other pretreatments exist, such as ionic liquids, deep eutectic solvents, microwave or ultrasound-assisted hydrothermal pretreatments, etc. Each pretreatment depends on the nature of the lignocellulosic material, the proposed application, and the objective (Galbe & Wallberg, 2019; Kruyeniskia et al., 2019).

At the end of the 20th century, the concept of platform molecules emerged, where a platform molecule (bio-derivative) can be used as the basis for obtaining high-value chemicals, such as glucose monosaccharides, which are hydrolyzed and fermented for ethanol production, hemicellulose hexoses, which are isolated from starch, and pentoses, which are ideally used for the production of furfural, most commonly levulinic acid (LA), furfural, and 5-hydroxymethylfurfural (HMF) 1. The bio-derivatives from biomass are the following 4-diacyl acids (succinic, fumaric, and maleic), furan-2,5-dicarboxylic acid, 3-hydroxypropionic acid, aspartic acid, glucaric acid, glutamic acid, itaconic acid, levulinic acid, 3-hydroxybutyrolactone, glycerol, sorbitol, xylitol and arabinitol are obtained from C5 and C6 sugars. (Clark & Deswarte, 2015; Käldström et al., 2017). Other platform molecules typically derived from polysaccharides are sorbitol, xylitol, chloromethyl furfural, and itaconic acid. Various methods, such as hydrolysis, fermentation, dehydration, oxidation, and hydrogenation, can produce platform molecules. Hydrogen plays

a vital role in the production of biofuels from platform molecules (Käldström et al., 2017).

Several studies have been conducted on using cocoa residue in producing platform molecules with applications in the food, pharmaceutical, and cosmetic industries. Its low cost and high availability make it attractive to the industrial sector and the scientific community, which has led to the development of technologies and the generation of patents that can add value to the cocoa production chain (Acosta et al., 2018; Vasquez et al., 2019).

There are several studies of microwave pretreatments of lignocellulosic materials for pyrolytic activation and less for hydrolytic activation; the studies that there are on microwave hydrolysis are by using traditional acid or alkaline solvents or with new generation solvents such as deep eutectic or ionic liquids, commonly as pretreatments before enzymatic saccharification, with advantages and disadvantages depending much on the nature of the lignocellulosic material (Liu et al., 2017; Zhu et al., 2016). However, there are no studies of microwave pretreatments without chemical solvents that optimize the production of sugars from depolymerization of hemicellulose during pretreatment as a unique effect of using microwaves with different temperature conditions.

In addition, using tools such as metabolomics and analytical techniques is necessary to identify the changes presented in the material because of pretreatment. There are several efficient qualitative and quantitative spectroscopic techniques for studying the identification and discrimination of variables. However, an excellent alternative is infrared spectroscopy, a fast, non-destructive technique. This technique, in addition to being efficient in the presentation of results, does not require complex sample preparation or the use of solvents or chemical reagents, so it is also considered compatible with green chemistry (Amirvaresi et al., 2021; Parastar et al., 2020). In both near-infrared (NIR) and mid-infrared (MIR) spectroscopy, the vibration of chemical bonds, when excited by the absorption of infrared radiation, expresses properties of the functional groups of the structures of each chemical compound, resulting in

characteristic spectra. This is why vibrational spectroscopy presents the opportunity to analyze a matrix since the fingerprints obtained contain all the sample components. However, this can also represent a complexity if one wants to explore it with the naked eye since subtle differences can be ignored or misinterpreted. Because these data are complex and extensive, it is necessary to combine infrared spectroscopy techniques with multivariate statistical methods in available computational formats, which have shown success in discriminating parameters in an unbiased, reliable, and fast manner (Cozzolino et al., 2012; Krahmer et al., 2015; Santos et al., 2021).

Multivariate statistics in metabolomics is a tool that provides the opportunity to analyze matrices with the measurement of small molecules called metabolites. For this purpose, various mathematical models are used to provide exploratory analysis and classification of samples or to develop calibration models. On certain occasions, the data are subjected to preprocessing, such as baseline corrections and peak alignment (Amirvaresi et al., 2021; Cozzolino et al., 2012). Principal component analysis is ideal for reducing the dimensionality of data by identifying interrelationships between samples that form clusters. PLS-DA is an adaptable method for data of a linear nature; it is used to predict variables, it is based on the covariance between the spectra matrices, and it is not sensitive to the collinearity of the variables since it is based on orthogonal components (Santos et al., 2021).

Therefore, the general objective of this study was to elucidate the influence of microwave-assisted hydrothermal pretreatment of cocoa pod husk by combining statistical and spectroscopic tools as an alternative for obtaining platform molecules.

Likewise, the specific objectives of the research are presented, which were:

1) The objective of this research was to determine the antioxidant activity and chemical characterization of the main components of the CPH (*Theobroma cacao* L. variety Carmelo). To contribute to the knowledge of this variety and present

elements of valorization of a highly available residue generated by the cocoa agroindustry.

2) To determine the effect of different microwave-assisted hydrothermal pretreatments of cocoa pod husks using Principal Component Analysis (PCA) as an unsupervised classification technique and Partial Least Squares Regression Discriminant Analysis (PLS-DA), Orthogonal Projections to Latent Structures (OPLS-DA) and Sparse Least Squares (sPLS-DA) as supervised methods in combination with Near and Mid Infrared Spectroscopy (NIR-MIR) and Nuclear Magnetic Resonance (NMR), as well as the elucidation of molecules obtained by one- and two-dimensional NMR.

3) The present study aims to optimize the conditions for the extraction of carbohydrates from CPH and HMC-CPH by MA-HTP, using DBB response surface analysis and ^{1}H NMR quantification, to make a green pretreatment for this lignocellulosic material more efficient.

1. LITERATURE REVIEW

1.1. Platform Molecules

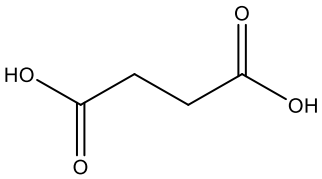
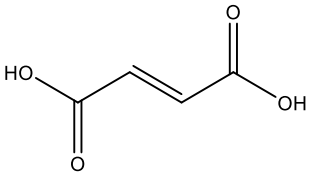
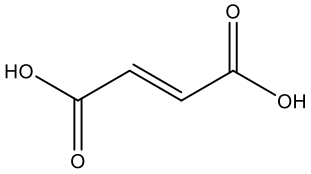
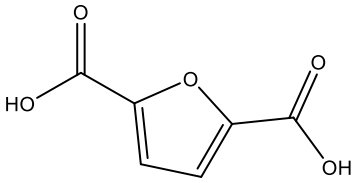
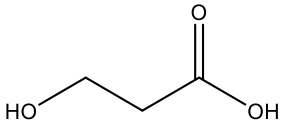
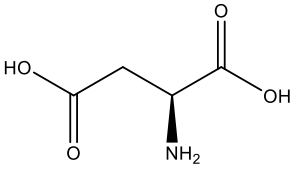
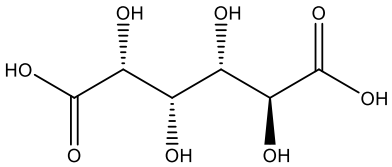
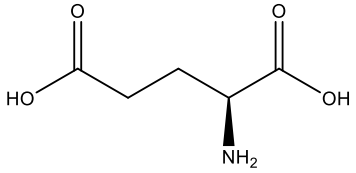
1.1.1. Platform molecule conceptualization

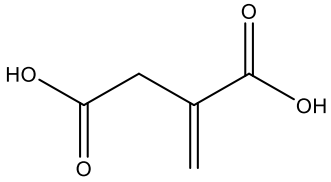
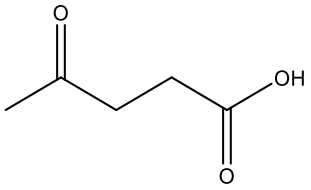
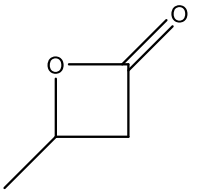
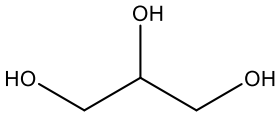
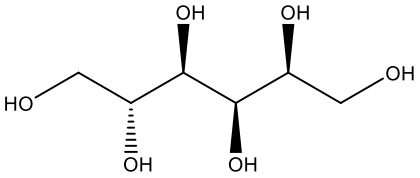
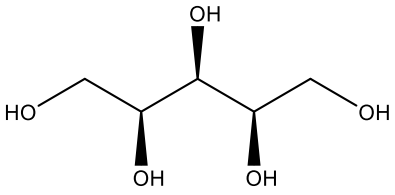
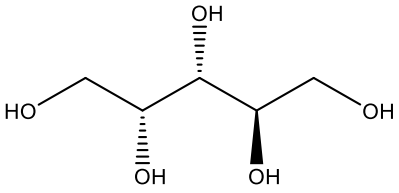
Platform molecules are chemical compounds composed of different functional groups that serve as the basis for producing hundreds of value-added products. These molecules are of biological origin, but their conceptualization comes from the petrochemical industry, which has traditionally been named "building blocks" to primary and intermediate products derived from petroleum (Bozell & Petersen, 2010; Shinde et al., 2020). From these products, which number between 8 and 12 commodities and around 30 intermediates, hundreds of products of fundamental use to society are obtained (Gérardy et al., 2020; Werpy & Petersen, 2004). As a result, the petrochemical industry is recognized as a significant driver of the global economy; however, the depletion of fossil resources, the energy dependence of countries, the effects of climate change, and other problems resulting from their extraction and use are perceived as an urgent global problem to be solved. Using renewable and sustainable raw materials to replace petrochemical industry products with bio-based platform molecules is a viable solution (Blanco Cejas et al., 2023).

1.1.2. Main biologically based platform molecules

In 2004, the U.S. Department of Energy published a report presenting a shortlist of 12 biobased molecules derived from sugars, shown in Table 1, which they identified as the "best" building blocks for obtaining high-value products, previously filtered from a list of 30 to 50 candidates, which in turn were obtained from a list of about 300 molecules (Gérardy et al., 2020; Werpy & Petersen, 2004).

Table 1 Building blocks based on sugar.

Molecule	Structure
1,4-Diacids (Succinic, Fumaric and Malic)	  
2,5-Furan dicarboxylic acid	
3-Hydroxypropionic acid	
Aspartic acid	
Glucaric acid	
Glutamic acid	

Itaconic acid	
Levulinic acid	
3-Hydroxybutyrolactone	
Glycerol	
Sorbitol	
Xylitol/Arabinitol	 

Glucose monosaccharides are hydrolyzed and fermented for ethanol production, and pentoses are ideally used for furfural production. Levulinic acid (LA), furfural and 5-hydroxymethylfurfural (HMF), 1,4-acids (succinic, fumaric and maleic), furan-2,5-dicarboxylic acid, 3-hydroxy propionic acid, aspartic acid, glucaric acid, glutamic acid, itaconic acid, levulinic acid, 3-hydroxybutyrolactone, glycerol, sorbitol, xylitol, and arabinitol are derived from C5 and C6 sugars (Clark &

Deswarte, 2015; Källström et al., 2017). Other platform molecules commonly obtained from polysaccharides include sorbitol, xylitol, chloromethyl furfural, and itaconic acid. Various methods, including hydrolysis, fermentation, dehydration, oxidation, and hydrogenation, can produce platform molecules. Hydrogen plays a vital role in the production of biofuels from platform molecules (Källström et al., 2017). The platform molecules will be derived from lignocellulosic biomass and should come from sustainable and environmentally friendly processes (Shinde et al., 2020).

1.2. Lignocellulosic biomass

1.2.1. Sources and provenance of lignocellulosic biomass

Lignocellulosic biomass is an abundant and sustainable renewable resource generated by nature; the photosynthesis of all terrestrial plants produces it so that it can be obtained at a low cost for use in the production of renewable energy and chemicals (Gérardy et al., 2020). Nonfood lignocellulosic materials are available in agriculture, forestry, industrial wood residues, agro-industrial, municipal solid waste, and energy plantations dedicated to this use (Meraj et al., 2023). Agro-industrial activities alone generate large amounts of lignocellulosic wastes that, if not properly managed, can cause environmental problems, as their disposal and conversion are not accessible due to the lack of appropriate technologies (Fong Lopez et al., 2019; Kumari & Singh, 2018; Zheng et al., 2017). Globally, the generation of lignocellulosic waste is approximately 220,000,000 tons per year (Kumari & Singh, 2018). In 2017, approximately 37,522,440 tons of agro-industrial waste was reported globally (Batista Meneses et al., 2022).

1.2.2. Composition and Structure of Lignocellulosic Materials

Lignocellulosic materials consist of cellulose, hemicellulose, and lignin, as well as volatile and non-volatile extractable compounds such as gums, waxes, resins, starch, sterols, pitch, terpenes, tannins, flavonoids, and chlorophylls; traces of pectins, non-structural sugars, nitrogenous material, and inorganic elements are

also present. Although their primary components are the same, these polymers vary in amount depending on their origin and species, ranging from 50 to 60% cellulose and hemicellulose and 20 to 30% lignin (Araya Farias et al., 2019; Galbe & Wallberg, 2019). Glucan contents extracted from cellulose, xylans from hemicellulose, and lignin contents from different biomass sources have been reported, all in varying proportions depending on the source (Table 2) (Dey et al., 2023; Roy, 2022a).

Table 2 Glucan, xylan, and lignin content of various lignocellulosic materials (adapted from).

Biomass	**Glucan (%, w/w)**	**Xylan (%, w/w)**	**Lignin (%, w/w)**
Cocoa pod husk	35.0	11.0	26.3
Nutshells	25-30	25-30	30-40
Peanut shells	20.9	12.4	42.7
Bamboo shoot shell	38.2	25.7	25.6
Corn stover	32.8-38.0	26.0-28.9	18.2-19
Corn Cobs	45	35	15
Sugar Cane Bagasse	42.3-43.5	20.9-22.7	20.5-20.8
	42.0	25	20.0
Wheat bran	17.5	19.9	12.5
Wheat straw	26.5-39.0	17.6-18.8	20.4-22.3
	29-35	26-32	16-21
Rice Straw	32.1	24	18
Sweet sorghum	45	27	21
Sorghum biomass	33.6	27.4	23.5
Sugarcane biomass	35.9	21.2	20.8
Oil palm biomass	33.7	20.3	17.1

Banana waste	13.2	14.8	14
Sponge gourd fibers	66.59	17.44	15.46
Grasses	32.9-35.5	20.2-22.4	23.2-25.8
Bamboo	40.0	21.6	26.5
Hardwoods	46.9	18.6	28.3
	40-55	24-40	18-25
Softwoods	45-50	25-35	25-35
Eucalyptus wood	44.4	14.3	24.5
Poplar wood	40.0-43.2	13.0-17.8	25.2-28.0
Birchwood	46.6	22.4	23.7
Newspaper	40-55	25-40	18-30
Paper mulberry	39.4	22.1	24.9
Sweetgum	40.2	15.7	25.2

This material has a complex and recalcitrant structure; the cell wall is made up of fibers, which in turn are made up of microfibrils, which are made up of bundles of cellulose polymers wrapped in hemicellulose and lignin, as shown in Figure 1 (Roy, 2022a).

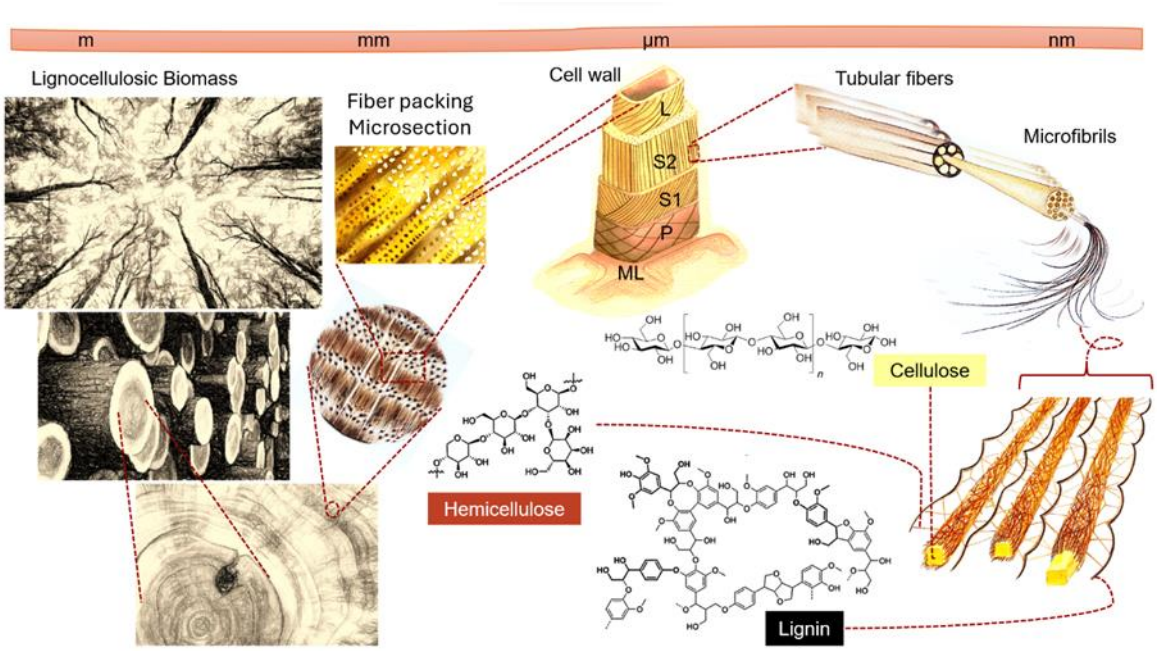


Figure 1. Cellular elements of wood as a source of lignocellulosic biomass (ML=middle lamella, P=primary cell wall layer, S1=secondary cell wall layer 1, S2=secondary cell wall layer 2, and L=lumen).

1.2.3. Cellulose

Cellulose is the major component of lignocellulosic materials. It is a long-chain linear polysaccharide composed of D-glucose units linked by β-1,4 bonds arranged in cellobiose subunits (Figure 2). It has a reducing end and a non-reducing end, and the cellulose chains are connected by hydrogen bridges (weak bonds) and van der Waals forces (hydrophobic forces) (Liao et al., 2020; Zheng et al., 2017). It has high crystallinity because cellulose polymers are stacked in microfibrils, forming flat lamellae with crystalline and amorphous zones. Its degree of polymerization results in strong molecular interactions (Liao et al., 2020; Xia et al., 2024).

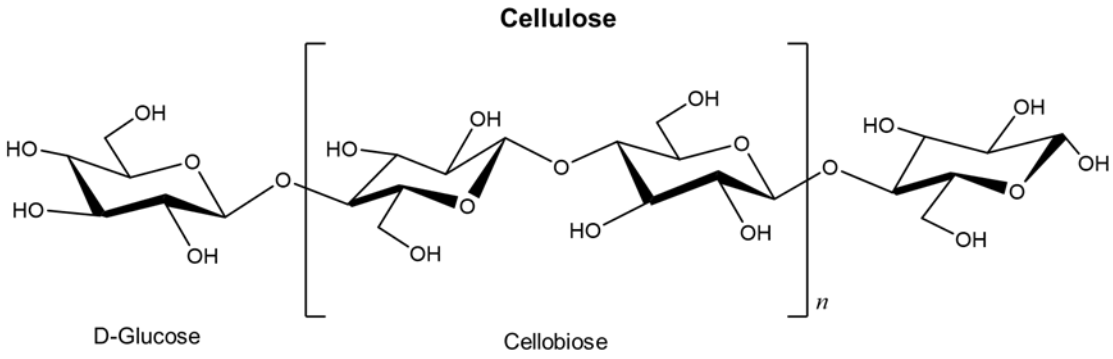


Figure 2. Cellulose structure monosaccharides (Adapted from (Shinde et al., 2020).

Cellulose in its natural form is called native cellulose or cellulose I. A cellulose fiber may consist of one to two billion cellobiose molecules. These are aggregated in monocyclic crystalline assemblies oriented in the same direction within microfibrils. The length of native cellulose is not known with certainty, but there is information that it exceeds the value of 104 nm (Salmén, 2015). This length also depends on the source of origin and the depolymerization treatment; for example, a wood cellulose fiber, after acid hydrolysis, reaches a degree of polymerization from 300 to 1700, and a cotton fiber can have a degree of polymerization from 800 to 10000 (Klemm et al., 2005).

1.2.4. Hemicellulose

Hemicellulose is a branched polysaccharide composed of pentoses (arabinose and xylose), hexoses (glucose, mannose, galactose, rhamnose), and acetylated sugars (Figure 3), which differ according to their origin (Roy, 2022a). There are less branched species that are more resistant to alkalis, such as glucomannan, and the more branched and soluble types, such as arabinogalactans, which typically have pectic and uronic acid residues (Salmén, 2015). Unlike cellulose and lignin, pretreatment makes hemicelluloses more easily modified (Roy, 2022a).

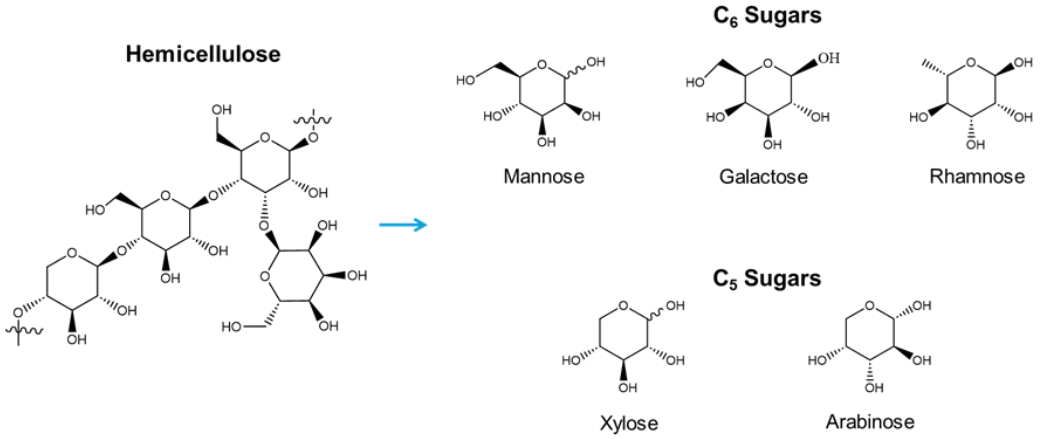


Figure 3 Hemicellulose derived C5 and C6 monosaccharides (Adapted from (Shinde et al., 2020).

1.2.5. Lignin

Lignin is the most abundant aromatic polymer in nature; it is a heteropolymer composed of phenylpropane units (monolignols), p-coumaryl (4-hydroxycinnamyl), sinapyl (3,5-dimethoxy-4-hydroxycinnamyl), and coniferyl (3-methoxy-4-hydroxycinnamyl) alcohols (Figure 4). Lignin is rich in potentially useful aromatic precursors, in addition to its numerous functional groups and high calorific value, making it a viable input to produce fuels, functional materials, and chemicals (He et al., 2024; Meng et al., 2023; Zheng et al., 2017). Lignin is tightly bound to cellulose and hemicellulose, providing stiffness and protection to the plant cell wall (Muharja et al., 2021).

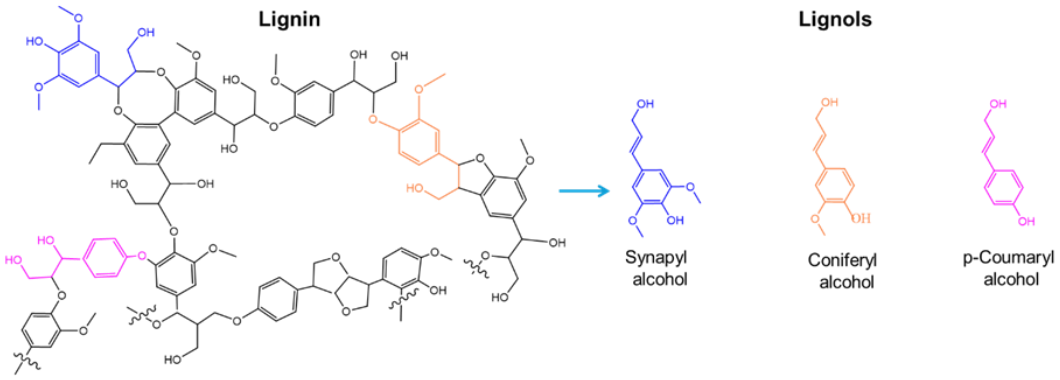


Figure 4 Lignin and lignols structures (Adapted from (Shinde et al., 2020).

To utilize lignin, it is necessary to consider its isolation, depolymerization, and conditioning of the obtained molecule. However, the isolation of lignin, with its low solubility and low reactivity due to its complex structure, hinders its conversion into value-added molecules. Therefore, its utilization remains limited. Currently, at least 98% of industrial lignin (paper industry) is underutilized for the production of heat and electricity, and the remaining 2% is transformed for acceptable applications in the sector (Tišma et al., 2021).

The irregular structure of the different lignocellulosic polymers, the complexity of the lignin molecules, and the crystallinity of the cellulose polymer, as well as its structural and compositional properties, make it a chemically complex, rigid raw material with very low solubility. These properties that characterize lignocellulosic materials provide self-protection to plants and create problems for their conversion and utilization (Kustov et al., 2022; Meraj et al., 2023). This is a great challenge since lignocellulosic biomass, besides being considered a source of opportunity to reduce dependence on oil and environmental pollution through its use, could also represent an ecological problem since large amounts of lignocellulosic waste are produced in the world, with an estimated annual production of 181,500 million tons, a large percentage of which is not adequately managed (Kustov et al., 2022; Meraj et al., 2023; Rocha Balbino et al., 2023). Biomass utilization processes are integrated into the biorefinery structure.

1.3. Biorefinery

Biomass conversion processes are brought together in a complex system called biorefinery. Thus, the conceptualization of biorefinery is the installation of biomass conversion processes and equipment to produce biofuels, chemicals, and materials that replace petrochemicals, supporting non-dependence on increasingly unsustainable fossil resources (Shinde et al., 2020). Therefore, the primary goal of biorefining is to transition to a sustainable economic system that uses resources efficiently, reduces waste generation, and uses them to obtain new products (Tišma et al., 2021).

1.3.1. Classification of a Biorefinery

Biorefineries are divided into first, second, and third-generation biorefineries. Generally, first-generation biorefineries use edible grains to produce ethanol, competing with the food industry. Second-generation plants also use edible grains but diversify into products other than ethanol, such as starch, corn oil, flour, high-fructose corn syrup, etc. Finally, third-generation plants aim to use agricultural and forestry biomass to produce various products (Shinde et al., 2020). The intermediate products that can be obtained from lignocellulosic biomass and the building blocks or platform molecules for forming new polymers through biorefinery processes are shown in Figure 5.

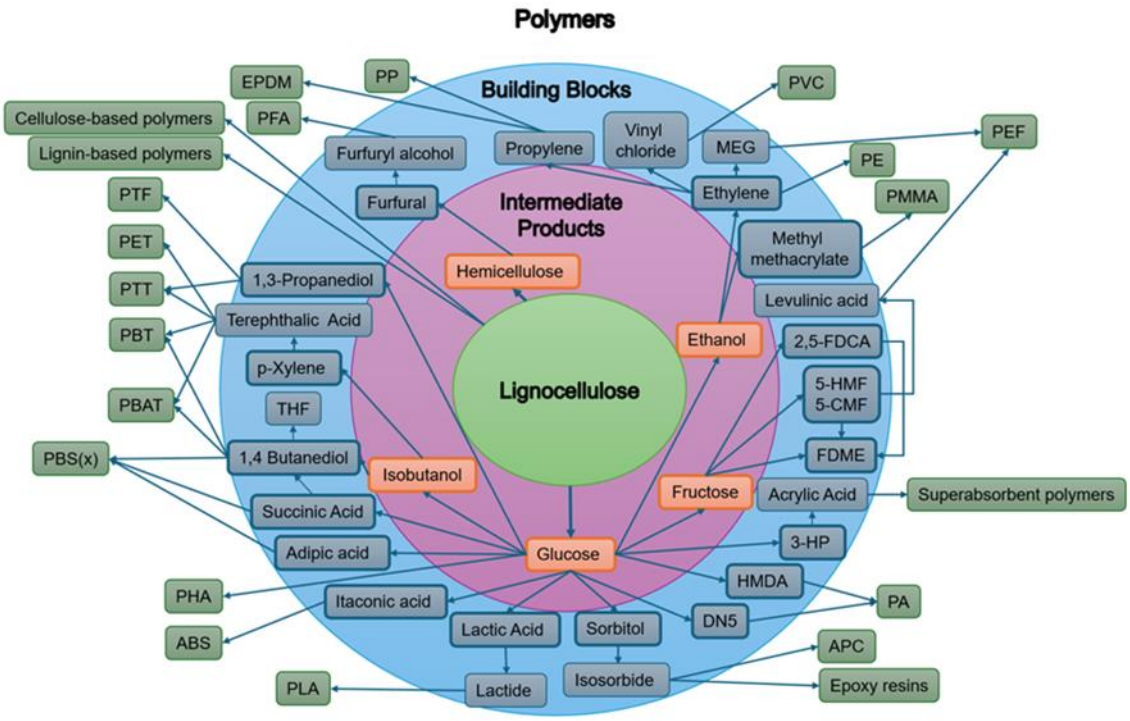


Figure 5. Intermediates, platform molecules, and polymers from lignocellulosic biomass (Adapted from (Skoczinski et al., 2022).

1.4. Pretreatments

The critical step for biomass utilization is pretreatment. Pretreatments are used for second and third-generation biorefinery processes (Meng et al., 2023). These pretreatments can be classified into physical, chemical, biological, thermal, physicochemical, and combined methods (Table 3) (Sun et al., 2022). The choice of pretreatment depends on the type and composition of the lignocellulosic biomass, as it can be forest origin, crop residues, food processing residues, energy crops, waste of biological origin, domestic, commercial, or municipal organics (Tišma et al., 2021).

The criteria to be taken into account for pretreatment to be considered adequate are the deconstruction of the three-dimensional and crystalline structure of cellulose, the efficiency of carbohydrate conversion, the low or non-existent formation of toxic products and fermentation inhibitors if the final objective is liquid biofuels, the low degradation of sugars, the possibility of recovering lignin for later

use, cost-effectiveness and low energy consumption (Alvira et al., 2010; Meng et al., 2023).

Table 3 Biomass pre-treatment methods

Pretreatment type	Method	Citation
Thermal	Pyrolysis	(Kumar Mishra et al., 2024; Liu et al., 2019; Zafeer et al., 2023)
	Carbonization	
	Torrefaction	
	Flash carbonization	
	Gasification	
Physical	Ball milling	(Real Pérez et al., 2024)
	Crumble	(Dilts, 2007)
	Disk milling	(Kumar et al., 2009; Roy, 2022b; Roy et al., 2020; Yang et al., 2023)
	Hammer mill.	
	Knife milling	
	Two roll mills.	
Chemical	Acids	(Brienza et al., 2023; Sun et al., 2022)
	Alkalis	(Balasubramani et al., 2024; Hassan et al., 2018)
	Oxidation	(Brienza et al., 2023)
	Ionic liquids	(Achinivu et al., 2024; Singh et al., 2024)

	Natural deep eutectic solvents (DES)	(Sharma et al., 2022; Xiao et al., 2024)
Biological	Microbial degradation	(Sun et al., 2022)
	Enzymatic hydrolysis	
Physic-chemical	Liquid hot water (LHW)	(Sun et al., 2022; Yan et al., 2016)
	Steam explosion	(Kumari & Singh, 2018)
	Alkali explosion	(Kumar et al., 2020)
	Ammonia fiber explosion (AFEX)	(Rahandi Lubis et al., 2024)
	Hydrothermal	(Saini et al., 2024)
Combined methods	Microwave-assisted	(Fan et al., 2024; Liu et al., 2019; Shinde et al., 2020; Torres Mayanga et al., 2019; Yue et al., 2022)
	Ultrasonic assisted	
	Thermochemical	
	Enzymatic hydrolysis	

1.4.1. Thermal pretreatments

Combustion or thermal pre-treatment of biomass is the most straightforward technology and is generally used to generate electricity or heat. However, its electrical efficiency is very low (20-35%); in addition, high temperatures are used to produce it, which produces high emissions of sulfur and nitrogen oxides, which are pollutants (Liu et al., 2019). These processes are known as pyrolysis, carbonization, and torrefaction. Pyrolysis is carried out at temperatures between 500 and 900°C in an inert medium, and carbonization is carried out under ambient conditions; it starts at a temperature of 100°C and is completed up to about 300°C, while torrefaction occurs between 200 and 300°C. These processes extract volatile organic matter, and the product is biochar (Kumar Mishra et al., 2024).

1.4.2. Physical pretreatments

Another type of pre-treatment is physical pre-treatment, which is mainly performed by mechanical processes that reduce the particle size of the biomass, providing greater volumetric density by increasing the accessible surface area and viability for subsequent hydrolysis treatment by decreasing the crystallinity of the cellulose, as well as providing better handling of the raw material and facilitating its management. However, high energy consumption is required for biomass grinding or shredding, which makes its large-scale application more expensive (Kumar et al., 2009; Yang et al., 2023).

1.4.3. Chemical pretreatments

Chemical pretreatments have been widely used because they offer advantages such as efficiency, simplicity, low cost, and thermal stability, which makes it easy to control the reaction conditions (Rusănescu et al., 2024; Yang et al., 2023). Exogenous agents such as acids, alkalis, and organic solvents are used in chemical pretreatments. These include methods such as those commonly used for pulping; kraft, sulfite, and soda, as well as acids, oxidants, organolsolv, ionic liquids, and deep eutectic solvents (Brienza et al., 2023).

1.4.4. Acid pretreatments

Acid pretreatments can be concentrated or diluted; concentrates are used to achieve the complete conversion of polysaccharides to fermentable sugars but require a washing and neutralization process before fermentation, making it a more polluting process (Brienza et al., 2023). In addition to being highly corrosive, it has the disadvantage of ultimately breaking down sugars into furans, organic acids, and condensation products that may be undesirable for further fermentation (Baruah et al., 2018; Brodeur et al., 2011). However, these intermediates can also be attractive to produce polymers and other chemicals of high commercial value (Bielski & Grynkiewicz, 2021; Gérardy et al., 2020). Commonly used acids are inorganic acids such as sulfuric acid (H_2SO_4), hydrochloric acid (HCl), phosphoric

acid (H_3PO_4), and nitric acid (HNO3). However, organic acids such as formic acid (CH2O2), acetic acid (CH_3COOH), and oxalic acid (C_2H2O_4) are also used (Brienza et al., 2023).

1.4.5. Alkaline pretreatments

If the depolymerization of lignocellulosic material is fermented for liquid biofuel production, alkaline pretreatments are better than acidic ones, as they degrade fermentable sugars to a lesser extent. In the same tenor, oxidative pretreatments are preferable as they produce fewer byproducts that can interfere with further fermentation (Hassan et al., 2018). In alkaline pretreatment, sodium hydroxide (NaOH), potassium hydroxide (KOH), and organic solvents are commonly used to remove hemicellulose and lignin and promote cellulose defibration (Balasubramani et al., 2024).

1.4.6. Oxidative pretreatments

Oxidative pretreatments are mainly used for pulp bleaching for paper; they use hydrogen peroxide, oxygen, ozone, and peroxy-acids and can be carried out in an acid or alkaline medium (Brienza et al., 2023). Lignin oxidation, in contrast to acid and alkaline treatments, involves several depolymerization pathways in which phenolic aldehydes, acids, or quinones can be formed by breaking carbon-carbon bonds of lignin side chains, and carboxylic and aliphatic acids can be formed by breaking aromatic rings, and biphenyl compounds can be formed by condensation reactions between lignin radicals (Ma et al., 2015). Various organic solvents such as methanol, ethanol, acetone, and tetrahydrofuran have been successfully used for biomass pretreatment (Wang et al., 2024).

Despite the advantages of these chemical pretreatments, they have the disadvantage of being corrosive, toxic, and environmentally unfriendly (Hassan et al., 2018).

1.4.7. Physic-chemical pretreatments

Ionic liquids have been considered a more environmentally friendly biomass pretreatment alternative to conventional chemical methods. Those based on alkanol ammonium acetate are effective in lignin depolymerization, and protic ionic liquids such as ethanol ammonium acetate have been recognized as effective in pretreating lignocellulosic biomass with high yields and sugar recovery (Singh et al., 2024; Socha et al., 2013). They have the advantage of being flammable, thermally stable, and handling low vapor pressures. Still, their industrial viability has been limited by their inability to be easily recycled, use of energy-intensive water washes, generation of toxic waste streams, and high economic costs. However, cost-reducing protic ionic liquid processes continue to be developed (Achinivu et al., 2024).

Another alternative for lignocellulose deconstruction is deep eutectic solvents (DES), which are low energy, low toxicity, highly efficient, and considered suitable for large-scale use. They can be efficiently recovered, purified, and recycled and are, therefore, considered an environmentally friendly method. They are also highly selective and have the potential to produce bio-based functional materials (Sharma et al., 2022; Xiao et al., 2024).

On the other hand, physicochemical pretreatments remove lignin and reduce cellulose crystallinity more efficiently than physical, chemical, and biological pretreatments (Rahandi Lubis et al., 2024). The most used physicochemical pretreatment is liquid hot water (LHW); this method uses water as a solvent at high temperature and pressure, and organic acids extracted from biomass by water serve as catalysts that help to break down the cell wall with minimal inhibitory products (Yan et al., 2016).

Another method used is the sudden steam explosion method, which consists of exposing the lignocellulosic material to direct contact with steam at high pressure for some time and then suddenly releasing the steam, causing the cell matrix to break by explosion or autohydrolysis (Kumari & Singh, 2018). This method is

industrially scalable, environmentally friendly, and effective, but it has the disadvantages of requiring high energy levels and producing inhibitory compounds for future fermentation (Agbor et al., 2011).

Another method is alkali blasting, which is considered a high-performance method of delignification. Alkali is used at high temperatures to dissolve the lignin and reduce the crystallinity of the cellulose. Among the disadvantages of the method are high pressures and temperatures, as well as long times, which make it costly (Kumar et al., 2020).

On the other hand, there is the ammonia fiber expansion method (AFEX). As its name suggests, ammonia is used at high pressure, which suddenly increases the temperature, breaking the ester and ether bonds of hemicellulose and lignin. It is also effective in reducing the crystallinity of cellulose. It is recommended for lignocellulosic materials with low lignin content (Rahandi Lubis et al., 2024).

Finally, hydrothermal pretreatment produces mainly Xylo oligosaccharides among the physicochemical pretreatments. Compared to other methods, it may be less efficient. Still, it is the most environmentally sustainable method because it uses only water as a solvent, so it is also less corrosive and much less costly. After hydrothermal, two fractions are obtained: a solid fraction rich in cellulose and lignin for further pretreatment and a liquid fraction enriched with sugars derived from hemicellulose (Saini et al., 2024).

1.4.8. Combined pretreatments

Combined pretreatments are generally between hydrothermal pretreatment and using technologies such as microwave or ultrasound-assisted, or by flash steam explosion or enzymatic hydrolysis. This combination of methods provides simple techniques with short processing times, are low corrosion and therefore environmentally friendly, and generates fewer undesirable by-products or inhibitors for liquid biofuel production (Yue et al., 2022). The hydrogen bonds present in the lignocellulosic matrix, mainly those of hemicellulose, weaken with

increasing water temperature, while the ionization constant increases, producing more hydrogen ions ($H3O^+$) and alkali hydroxide ions (OH^-), which give water catalytic properties and depolymerize the glycosidic bonds (Fan et al., 2024; Torres Mayanga et al., 2019). This results in converting acetyl groups into organic acids that accelerate the depolymerization of hemicellulose polysaccharides (Yue et al., 2022).

1.5. Green chemistry

Social and environmental sustainability is sought in the pretreatment of lignocellulosic materials, as in all processes in a biorefinery. Globally, the scientific community has focused its efforts over the past two decades on addressing sustainability issues, and the concept of green chemistry has gained acceptance and interest. Green chemistry is defined as the design of chemical products and processes that reduce or eliminate the use and generation of hazardous substances. In 1998, Paul Anastas and John Warner published the twelve principles of green chemistry (Collins, 2017), listed in Table 4. The application of green chemistry ranges from designing new reactions and products to developing new processes and concepts. For example, the idea of circular chemistry is where a circular economy is established in a sustainable future. However, it has been repeatedly stated that for the twelve principles to work, a multisectoral and coordinated approach between academic research and the chemical industry is required (Ardila Fierro & Hernandez, 2021).

Table 4 The twelve principles of green chemistry (adapted from (Erythropel et al., 2018)).

Prevent waste	**Preventing waste is always preferable, even if it can be recycled or treated later.**
Atom Economy	Correctly design synthesis methods to reduce by-product formation.

Less hazardous synthesis	Search for synthesis methods that produce less toxic substances for humans and the environment.
Design benign chemicals	Develop chemical products that maintain their efficacy while reducing their toxicity.
Harmless solvents & auxiliary materials	Design synthesis or separation processes that reduce or eliminate solvents or auxiliary reagents.
Designed for energy efficiency	Design methods or processes with energy requirements that have a lower environmental and economic impact.
Use of renewable resources	Seek to use renewable resources where technically and economically feasible.
Reduce derivatives	Reduce waste by reducing derivatization steps and reagent usage.
Catalysis	Strengthen new avenues of research for more selective and reusable catalysts instead of stoichiometric reagents.
Design for degradation	Design chemical products that do not remain in the environment at the end of their useful life.
Real-time analysis for pollution prevention	Development of analytical methods for real-time process monitoring to prevent by-product formation.
Inherently benign chemistry for accident prevention	Reduce the potential for chemical accidents by using less hazardous chemicals and processes and reducing fumes, explosions, and fires.

1.5.1. Microwave-assisted pretreatment.

An alternative for implementing green chemistry principles is the microwave-assisted pretreatment of lignocellulosic material. Conditionally, microwave technology can enhance a chemical process based on efficient heating of a material by the effect of electromagnetic irradiation of the material (Wang et al., 2024). The electromagnetic spectrum's microwave region lies between infrared and radio wave frequencies. Therefore, microwave electromagnetic irradiation occurs at frequencies from 0.3 to 300 GHz, corresponding to wavelengths from 1 cm to 1 m (Kappe & Stadler, 2005).

1.5.2. Types of radiation

Conventional microwaves and laboratory equipment operate at a frequency of 2.45 GHz and a wavelength of 12.25 cm (Nüchter et al., 2004), which is too low compared to ultraviolet radiation, visible light, X-rays, and gamma rays. Microwave electromagnetic radiation does not break molecular bonds or induce chemical reactions by directly absorbing energy. This can be explained by comparing the types of radiation with the energies of the different kinds of bonds shown in Tables 5 and 6. Microwave irradiation produces only dielectric heating, which depends on the ability of the solvent or reagent to absorb the energy and convert it to heat (Gabriel et al., 1998; Mingos & Baghurst, 1991).

Table 5 Types of radiation and energy of bond types.

Irradiation	Frequency (MHz)	Quantum Energy (eV)	Bond type	Energy Binding (ev)
Radio Frequency	1	4.0×10^{-9}	**Hydrogen bonding**	0.04-0.44
Microwave	2450	0.0016	**C-C single bond**	3.61

Infrared Light	3.0×10^{6}	0.012	**C-O single bond**	3.74
Visible	6.0×10^{8}	2.5	**C-H bond**	4.28
UltraViolet	1.0×10^{9}	4.1	**O-H bond**	4.80
X-rays	3.0×10^{13}	1.24×10^{5}	**C=C double bond**	6.35
Gamma rays	3.0×10^{14}	1.24×10^{6}	**C=O double bond**	7.71

1.5.3. Microwave electromagnetic radiation.

Electromagnetic waves are composed of electric and magnetic fields (Figure 6), with the electric component causing heating. This heating occurs by two mechanisms: 1) dipolar polarization and 2) ionic conduction. Dipolar polarization is caused by the effect of the oscillation of the applied electric field and the search for alignment of the dipoles of the sample with this field; in this process, molecular friction and dielectric loss energy are dissipated, which is released to the medium in the form of heat. On the other hand, ionic conduction is caused by the oscillation of the ions under the influence of the microwave field; this agitation causes the collision of the ions with neighboring atoms, generating movement or agitation that produces heat (Stass et al., 2000).

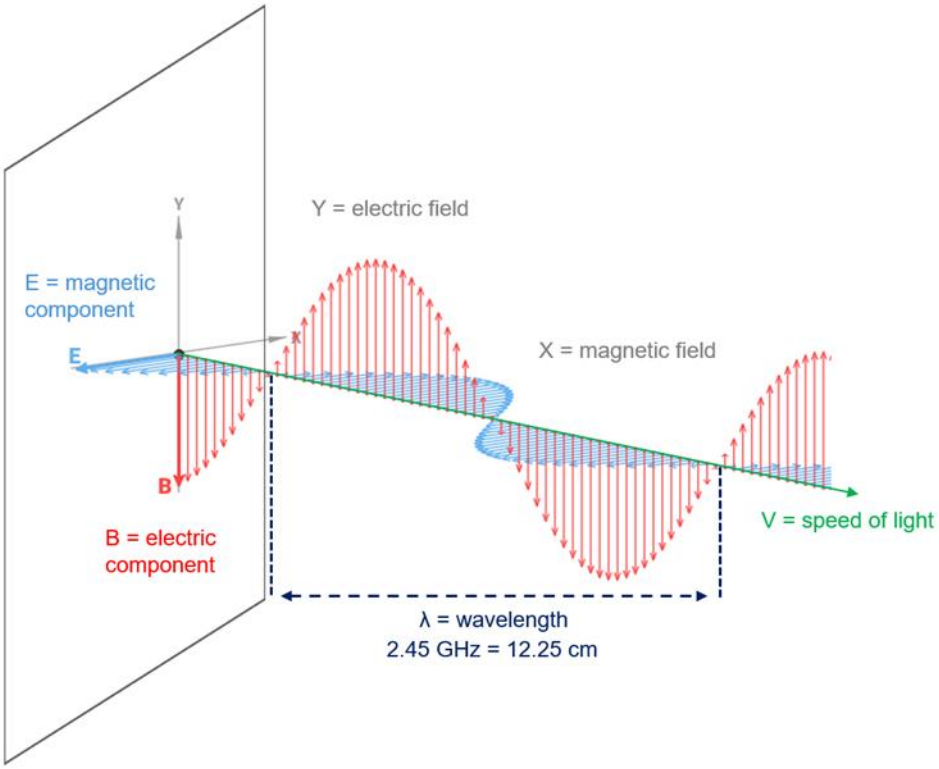


Figure 6 Electromagnetic radiation (adapted from (Kappe & Stadler, 2005).

Microwaves interact with matter by absorption, transmission, and reflection. Therefore, microwaves' efficiency depends on the material's nature; for example, polar organic solvents are more efficient in heating the medium in response to microwave irradiation (Kappe & Stadler, 2005).

Therefore, microwave irradiation outperforms conventional heating methods in terms of heating rate, higher reaction or treatment rate, lower activation energy consumption, and higher energy efficiency, and can be achieved in small devices. And what is outstanding is that the thermodynamic properties are different from conventional heating, such as selective and volumetric heating, and can be easily controlled (Wang & Wang, 2016).

The use of this technology has been extended to very diverse applications, such as organic and inorganic synthesis, catalysis, pretreatment of lignocellulosic materials, polymerization, soil remediation, air purification, wastewater treatment, sludge treatment for biogas production, phenolic compounds, pharmaceuticals,

polluted wastes from agro-industry (Jones et al., 2002; Wang & Wang, 2016; Yuan et al., 2006; Zhang et al., 2007; Zhihui et al., 2005).

1.5.4. Microwave-assisted extraction of lignocellulosic materials.

According to a study of pectin extraction from cocoa pod husk, microwave-assisted extraction of lignocellulosic materials is an attractive option as it reduces the extraction time and solvent amount by two to six times. This study concludes that extraction with short irradiation times and mild conditions protects product quality. Furthermore, it can be combined with organic solvents, which also supports the principles of green chemistry (Pangestu et al., 2020a).

The literature on microwave-assisted extraction of lignocellulosic materials provides results that highlight cocoa pod husks as a rich source of metabolites and intermediates or platform molecules for potential use in functional foods and nutraceuticals (Nguyen et al., 2022; Sarah et al., 2022). Several studies support that cocoa pod husks are rich in polyphenols, methylxanthines, dietary fiber, and phytosterols or bioactive compounds and can be extracted with relative ease and effectiveness using different methods (Belwal et al., 2022).

1.6. *Theobroma cacao* L.

1.6.1. *Theobroma cacao* origins

Cacao (*Theobroma cacao* ssp.) is a tree species native to the northern tropical regions of South America; however, Central America has been recognized as the first center of domestication and cultivation (Motamayor et al., 2002). Recent archaeological evidence indicates that cacao was used in the upper Amazon as early as 5300 years ago. The first records come from the Amazon and upper Orinoco basins. It is believed to have been cultivated in this region as early as 1600, with the highest production in the tributaries of Guayaquil, Ecuador (Jaimez et al., 2022). It is known to have been cultivated in Mexico for more than 2000 years (Motamayor et al., 2002) and was of great cultural and economic importance to the Mayans and Aztecs due to the high functionality of its fruit.

Cacao beans were used as currency, while the cocoa-based beverage was used for religious and medicinal purposes and was described as the food of the gods by the botanist Carolus Linnaeus in the 17th century (McAnany & Murata, 2007).

1.6.2. *Theobroma cacao* economic importance

After the commercial potential of cocoa spread to Europe, this crop became the basis of the economy of the Amazon region in the 19th century. At that time, *Theobroma cacao*, a Central American criollo variety, was grown in Venezuela, Trinidad, Jamaica, Haiti, and the Windward Islands. In 1900, practically 80% of world production came from the West, but by the beginning of the 21st century, at least 86% came from the Eastern Hemisphere (Bailey & Meinhardt, 2016; Ruf & Schroth, 2003). Thus, this tropical crop is one of the most important in the world, producing valuable products such as chocolate, butter, and cocoa powder, which are marketed very successfully (McAnany & Murata, 2007). Currently, at least 5 million tons of dry cocoa beans are produced annually, with four main varieties: Criollo, Trinitario, Forastero, and Nacional, classified according to geographical origin and bean morphology, with Forastero having the highest commercial production (about 70%) (Sánchez et al., 2023b).

1.6.3. *Theobroma cacao* production worldwide

According to statistics published by the International Cocoa Organization, the leading producers in the world are Côte d'Ivoire, Ghana, and Ecuador, which account for approximately 75% of the annual production, with Africa being the leading producer, followed by the Americas. This information is detailed in Table 7, which shows estimated cocoa production for the last three years.

Table 6 Cocoa Bean Production-Cocoa year 2022/23 (Adapted from ICCO Quarterly Bulletin of Cocoa Statistics, Vol. XLIX, No.2).

Provenance	2020/2021	2021/2022	2022/2023

	(Thousand tonnes)	(%)	(Thousand tonnes)	(%)	(Thousand tonnes)	(%)
Africa	**4054**	**77.3**	**3589**	**74.5**	**3727**	**74.8**
Cameroon	*292*		*295*		*290*	
Côte d'Ivoire	*2248*		*2121*		*2200*	
Ghana	*1047*		*683*		*750*	
Nigeria	*290*		*280*		*280*	
Others	*177*		*210*		*207*	
America	**933**	**17.8**	**963**	**20**	**988**	**19.8**
Brazil	*200*		*220*		*210*	
Ecuador	*365*		*365*		*400*	
Others	*368*		*378*		*378*	
Asia & Oceania	**254**	**4.8**	**266**	**5.5**	**265**	**5.3**
Indonesia	*170*		*180*		*180*	
Papua New Guinea	*42*		*42*		*42*	
Others	*42*		*44*		*43*	
World total	**5242**	**100.0**	**4818**	**100.0**	**4980**	**100.0**

1.6.4. *Theobroma cacao* cultivation in Mexico

Cocoa plantations in Mexico occupy an area of approximately 47,000 hectares, mainly in the states of Tabasco and Chiapas. An annual production of between 24,000 and 35,000 tons represents about 0.50% of world production, which

places the country in the thirteenth or fourteenth place in this economy (Maza Villalobos et al., 2024).

1.6.5. *Theobroma cacao* fruit

As mentioned, cocoa's value corresponds to its fruit's great functionality. This fruit, known as cacao, consists of seeds, beans, or grains covered by a shell or testa. They are covered by a white pulp known as mucilage, and attached to the fruit's outer shell are the placenta and the funiculus. The outer shell is formed by the endocarp, mesocarp, and exocarp, known as the cocoa pod husk (Figure 7).

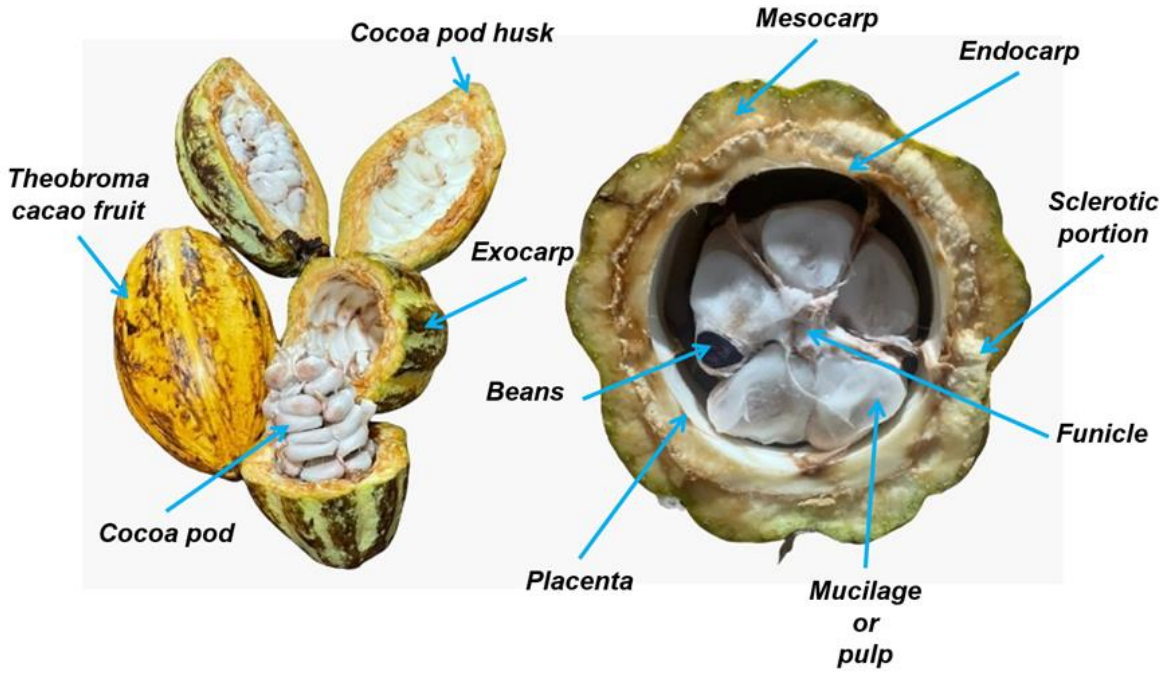


Figure 7 Components of the cacao fruit.

1.6.6. Waste generation from cocoa production.

The shelled seeds are those used to produce chocolate and its derivatives. The seed and fruit shells remain agro-industrial residues from this process (Okiyama et al., 2017). It is estimated that from each ton of dry cocoa beans marketed, about 20 million tons are left as residual biomass, this is about 90% of the total weight of the cocoa fruit; within this residual biomass is considered the cocoa pod husk, the seed shell and the pulp sweat, this has environmental and economic impacts

(Sánchez et al., 2023a). The most significant fraction of these residues is the cocoa pod husk, as it is estimated that approximately 48 million tons are produced annually worldwide. This by-product is obtained from the first step of the cocoa production line, which consists of opening the fruit and separating the pod from the pod, also known as pulping. This fraction is followed by that of the seed shell, with 700 thousand tons per year; this by-product is obtained from the fifth step of the cocoa production process, which consists of roasting the seeds (Quiceno Suarez et al., 2024a). Figure 8 shows the cocoa processing line and the residues obtained.

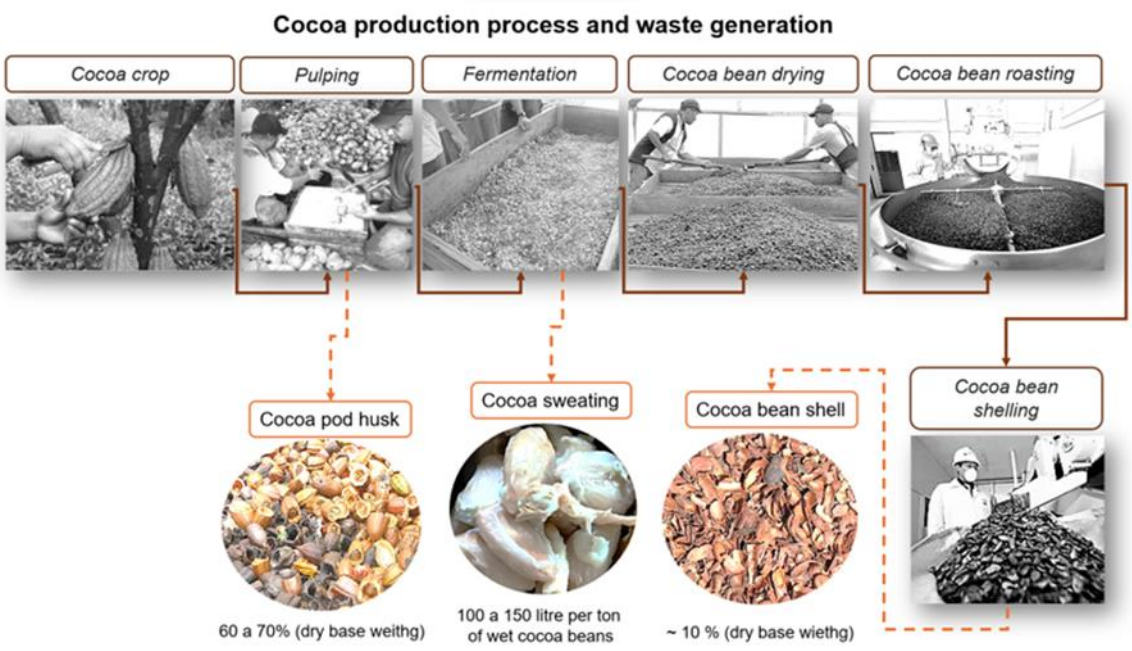


Figure 8 Waste generation from the cocoa production line.

Family farmers dominate the world's cocoa production, and 40 to 50 million people benefit from this crop. These people could benefit from using the residues in addition to the cocoa itself. This industry has also been attractive to the scientific community; there is a wide variety of research on the use of cocoa pod husk, and particular interest has been placed on the development of technologies with a biotechnological approach to obtain metabolites or molecules, as well as the generation of patents that can add value to the cocoa production chain (Vasquez et al., 2019).

1.6.7. Valorization of cocoa waste.

Some industries have considered using cocoa residues as fuel because these residues have a higher calorific value than chips from various types of wood. The cocoa industry in Brazil uses the shells as fuel and wood chips, thus avoiding the high cost of effective disposal of the residue. However, the large amount of ash produced during combustion is a secondary problem since about 10% of what is burned is ash. Figures indicate that 43,000 tons of ash could be generated per year. This brings us back to the search for new alternative uses (Fontes et al., 2019). In this sense, studies have been carried out on using these ashes. One option that has seemed viable for many years is soap production since amounts of about 57% potassium have been reported in this ash. These soaps are known as "black soaps"; they show good properties conferred to those of a soap, such as solubility, consistency, cleaning abilities, etc. (Taiwo & Osinowo, 2001).

Several studies indicate that cocoa pod husk is also a good animal feed due to their high fiber content, and they are also used for fertilizer production due to their rich mineral content, which are common uses for this material (Okiyama et al., 2017).

Another proposed application for cocoa pod husk is to obtain activated carbon, and it has been reported that the absorption capacity of this material can be improved by acid leaching with hydrochloric acid and physical activation at 850°C. In addition, the liquid residue after acid leaching can be recovered for use as fertilizer to make comprehensive use of the proposed products and co-products (Tsai & Huang, 2018).

Another study proposed to find the proportion of cocoa pod husk phytochemical extracts that effectively inhibit the growth of pathogenic bacteria in food, such as *Escherichia coli*, *Staphylococcus aureus*, and *Salmonella sp*. The results showed that the extracts contain phenol, flavonoid, triterpenoid, tannin, and alkaloid, which have selective inhibitory effects on the studied bacteria (Kayaputri et al., 2020).

Recent research has tested using more environmentally friendly technologies such as microwave-assisted extraction to obtain uronic acid, total phenolic content, and antioxidant yield from the cocoa shell, optimizing the effects of pH, time, temperature, and solid/liquid ratio. A significant influence of pH on antioxidant extraction was reported, as alkaline conditions showed higher yields. Optimal conditions were determined to be 5 min, pH 12, 97°C, and S/L 0.04 g/mL, which showed better results than conventional chemical extraction. Thirty-two extracts rich in proteins, polysaccharides, and polyphenols were determined, showing significant antioxidant results. Microwave-assisted extraction is proposed as an efficient technique for obtaining bioactive molecules from cocoa shells with potential use in the food industry (Mellinas et al., 2020).

The bromatological composition of cocoa pod husk has highlighted that it is a lignocellulosic material rich in fiber, protein, and fatty acids (oleic, linoleic, elaidic, and stearic). Its volatile compounds are composed of pyrazines, esters, alcohols, and aldehydes, and it is rich in polyphenols, epicatechin, theobromine, and caffeine. Therefore, cocoa by-products have great potential for extraction and biotransformation into precious compounds, as this residue has a low cost and high availability that is attractive to the industrial sector. Likewise, their use reduces the phytosanitary, environmental, and economic problems resulting from the inadequate disposal of waste generated in cocoa production (Quiceno Suarez et al., 2024a; Vasquez et al., 2019).

1.7. References

Achinivu, E. C., Blankenship, B. W., Baral, N. R., Choudhary, H., Kakumanu, R., Mohan, M., Baidoo, E. E. K., Scown, C. D., George, A., Simmons, B. A., & Gladden, J. (2024). Biomass pretreatment with distillable ionic liquids for an effective recycling and recovery approach. *Chemical Engineering Journal*, *479*. https://doi.org/10.1016/j.cej.2023.147824

Acosta, R., Sanabria, J., & Nabarlatz, D. (2018). Biomass from Colombian Agroindustrial Activities: Characterization and Potential for Oligosaccharides Production. *CHEMICAL ENGINEERING TRANSACTIONS*, *65*, 667-671. https://doi.org/DOI:10.3303/CET1865112

Agbor, V. B., Cicek, N., Sparling, R., Berlin, A., & Levin, D. B. (2011). Biomass pretreatment: fundamentals toward application. *Biotechnol Adv*, *29*(6), 675-685. https://doi.org/10.1016/j.biotechadv.2011.05.005

Ahmed, A., Khan, M. N., Ahmad, A., Khan, S. A., & Sohail, M. (2019). Optimization of pectinase production from Geotrichum candidum AA15 using response surface methodology. *Pakistan Journal of Botany*, *51(2)*(2), 743-750. https://doi.org/10.30848/pjb2019-2(41)

Alvira, P., Tomas Pejo, E., Ballesteros, M., & Negro, M. J. (2010). A review of pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis. *Bioresour Technol*, *101*(13), 4851-4861. https://doi.org/10.1016/j.biortech.2009.11.093

Amirvaresi, A., Nikounezhad, N., Amirahmadi, M., Daraei, B., & Parastar, H. (2021). Comparison of near-infrared (NIR) and mid-infrared (MIR) spectroscopy based on chemometrics for saffron authentication and adulteration detection. *Food Chem*, *344*, 128647. https://doi.org/10.1016/j.foodchem.2020.128647

Araya Farias, M., Husson, E., Saavedra-Torrico, J., Gerard, D., Roulard, R., Gosselin, I., Rakotoarivonina, H., Lambertyn, V., Remond, C., & Sarazin, C. (2019). Wheat Bran Pretreatment by Room Temperature Ionic Liquid-Water Mixture: Optimization of Process Conditions by PLS-Surface Response Design. *Front Chem*, *7*, 585. https://doi.org/10.3389/fchem.2019.00585

Ardila Fierro, K. J., & Hernandez, J. G. (2021). Sustainability Assessment of Mechanochemistry by Using the Twelve Principles of Green Chemistry. *ChemSusChem*, *14*(10), 2145-2162. https://doi.org/10.1002/cssc.202100478

Bailey, B. A., & Meinhardt, L. W. (2016). *Cacao DiseasesA History of Old Enemies and New Encounters* (B. A. Bailey & L. W. Meinhardt, Eds.). Springer International Publishing. https://doi.org/10.1007/978-3-319-24789-2

Balasubramani, V., Nagarajan, K. J., Karthic, M., & Pandiyarajan, R. (2024). Extraction of lignocellulosic fiber and cellulose microfibrils from agro waste-palmyra fruit peduncle: Water retting, chlorine-free chemical treatments,

physio-chemical, morphological, and thermal characterization. *Int J Biol Macromol*, *259*(Pt 2), 129273. https://doi.org/10.1016/j.ijbiomac.2024.129273

Baruah, J., Nath, B. K., Sharma, R., Kumar, S., Deka, R. C., Baruah, D. C., & Kalita, E. (2018). Recent Trends in the Pretreatment of Lignocellulosic Biomass for Value-Added Products. *Frontiers in Energy Research*, *6*. https://doi.org/10.3389/fenrg.2018.00141

Batista Meneses, D., Montes de Oca Vásquez, G., Vega Baudrit, J. R., Corrales Castillo, J., & C., M. A. L. (2022). Pretreatment methods of lignocellulosic wastes into value-added products: recent advances and possibilities. *Biomass Conversion and Biorefinery 12*, 547-564. https://doi.org/https://doi.org/10.1007/s13399-020-00722-0

Belwal, T., Cravotto, C., Ramola, S., Thakur, M., Chemat, F., & Cravotto, G. (2022). Bioactive Compounds from Cocoa Pod Husk: Extraction, Analysis and Applications in Food Production Chain. *Foods 11*(798), 1-13. https://doi.org/ https://doi.org/10.3390/foods11060798

Bielski, R., & Grynkiewicz, G. (2021). Furan platform chemicals beyond fuels and plastics. *Green Chemistry*, *23*(19), 7458-7487. https://doi.org/10.1039/d1gc02402g

Blanco Cejas, J., Martín, S., Linares, M., Iglesias, J., & Moreno, J. (2023). Life cycle assessment applied to bio-based platform molecules: Critical review of methodological practices. *Journal of Cleaner Production*, *414*. https://doi.org/10.1016/j.jclepro.2023.137513

Bozell, J. J., & Petersen, G. R. (2010). Technology development for biobased products from biorefinery carbohydrates—the US Department of Energy's "Top 10" revisited. *Green Chemistry*, *12*(4). https://doi.org/10.1039/b922014c

Brienza, F., Cannella, D., Montesdeoca, D., Cybulska, I., & Debecker, D. P. (2023). A guide to lignin valorization in biorefineries: traditional, recent, and forthcoming approaches to convert raw lignocellulose into valuable materials and chemicals. *RSC Sustainability*, *2*(1), 37-90. https://doi.org/10.1039/d3su00140g

Brodeur, G., Yau, E., Badal, K., Collier, J., Ramachandran, K. B., & Ramakrishnan, S. (2011). Chemical and physicochemical pretreatment of lignocellulosic biomass: a review. *Enzyme Res*, *2011*, 787532. https://doi.org/10.4061/2011/787532

Clark, J., & Deswarte, F. (2015). *Introduction to Chemicals from Biomass* (Second Edition ed.).

Clement, M. J., Corma, A., & Iborra, S. (2014). Conversion of biomass platform molecules into fuel additives and liquid hydrocarbon fuels. *Green Chemistry*, *16*(2). https://doi.org/10.1039/c3gc41492b

Collins, T. J. (2017). Review of the twenty-three-year evolution of the first university course in green chemistry: teaching future leaders how to create sustainable societies. *Journal of Cleaner Production*, *140*, 93-110. https://doi.org/10.1016/j.jclepro.2015.06.136

Coma, M., Martinez-Hernandez, E., Abeln, F., Raikova, S., Donnelly, J., Arnot, T. C., Allen, M. J., Hong, D. D., & Chuck, C. J. (2017). Organic waste as a sustainable feedstock for platform chemicals. *Faraday Discuss*, *202*(0), 175-195. https://doi.org/10.1039/c7fd00070g

Cornejo, A., Alegria Dallo, I., García Yoldia, Í., Sarobea, Í., Sánchez, D., Eduardo, O., Funcia, I., Gila, M. J., & Martínez Merino, V. (2019). Pretreatment and enzymatic hydrolysis for the efficient production of glucose and furfural from wheat straw, pine, and poplar chips. *Bioresource Technology 288*(121583). https://doi.org/https://doi.org/10.1016/j.biortech.2019.121583

Cozzolino, D., McCarthy, J., & Bartowsky, E. (2012). Comparison of near-infrared and mid-infrared spectroscopy to discriminate between wines produced by different Oenococcus oeni strains after malolactic fermentation: A feasibility study. *Food Control*, *26*(1), 81-87. https://doi.org/10.1016/j.foodcont.2012.01.003

Dahunsi, S. O., Adesulu Dahunsi, A. T., & Izebere, J. O. (2019). Cleaner energy through Cocoa (Theobroma cacao) pod husk liquefaction: Pretreatment and process optimization. *Journal of Cleaner Production*, *226*, 578-588. https://doi.org/10.1016/j.jclepro.2019.04.112

Dey, P., Bhattacharjee, S., Yadav, D. K., Hmar, B. Z., Gayen, K., & Bhowmick, T. K. (2023). Valorization of waste biomass for synthesizing carboxy-methyl-cellulose as a sustainable edible coating on fruits: A review. *International Journal of Biological Macromolecules*, *253*(Pt 7), 127412. https://doi.org/10.1016/j.ijbiomac.2023.127412

Dilts, M. D. (2007). *Application of the roller mill and hammermill for biomass fractionation* (Publication Number UMI Number: 1447558) Iowa State University]. Ames, Iowa.

Erythropel, H. C., Zimmerman, J. B., Petitjean, L., Melnikov, F., de Winter, T. M., Lam, C. H., Lounsbury, A. W., Mellor, K. E., Tu, Q., Pincus, L. N., Janković, N. Z., Falinski, M. M., Shi, W., Coish, P., Plata, D. L., & Anastas, P. T. (2018). The Green ChemisTREE: 20 years after taking root with the 12 principles. *Green Chemistry*, *20*(9), 1929-1961. https://doi.org/10.1039/c8gc00482j

Fan, X., Ren, M., Zhou, C., Kong, F., Hua, C., Fakayode, O. A., Okonkwo, C. E., Li, H., Liang, J., & Wang, X. (2024). Total utilization of lignocellulosic biomass with xylooligosaccharides production priority: A review. *Biomass and Bioenergy*, *181*. https://doi.org/10.1016/j.biombioe.2023.107038

Fong Lopez, M. D. C., Rigal, M., Rigal, L., Vilarem, G., & Vandenbossche, V. (2019). Influence of temperature and soda concentration in a thermo-

mechanochemical pretreatment for bioethanol production from sweet corn coproducts. *Industrial Crops & Products*, *133*, 317-324. https://doi.org/doi:https://doi.org/10.1016/j.indcrop.2019.03.044.

Fontes, C. M. A., Silva, R. B., & Lima, P. R. L. (2019). Characterization and Effect of Using Bottom and Fly Ashes from Co-combustion of Cocoa Waste as Mineral Addition in Concrete. *Waste and Biomass Valorization*, *10*(1), 223-233. https://doi.org/10.1007/s12649-017-0031-x

Gabriel, C., Gabriel, S., Grant, E. H., Halstead, B. S. J., & Mingos, M. P. (1998). Dielectric parameters are relevant to microwave dielectric heating. *Chemical Society Reviews*, *27*, 213-223.

Galbe, M., & Wallberg, O. (2019). Pretreatment for biorefineries: a review of common methods for efficiently utilizing lignocellulosic materials. *Biotechnol Biofuels*, *12*, 294. https://doi.org/10.1186/s13068-019-1634-1

Gatt, E., Rigal, L., & Vandenbossche, V. (2018). Biomass pretreatment with reactive extrusion using enzymes: A review. *Industrial Crops & Products*, *122*, 329-339. https://doi.org/https://doi.org/10.1016/j.indcrop.2018.05.069

Geng, W., Narron, R., Jiang, X., Pawlak, J. J., Chang, H.-m., Park, S., Jameel, H., & Venditti, R. A. (2019). Lignin content and structure influence hemicellulose alkaline extraction for non-wood and hardwood lignocellulosic biomass. *Cellulose*, *26*(5), 3219-3230. https://doi.org/10.1007/s10570-019-02261-y

Gérardy, R., Debecker, D. P., Estager, J., Luis, P., & Monbaliu, J.-C. M. (2020). Continuous Flow Upgrading of Selected C(2)-C(6) Platform Chemicals Derived from Biomass. *Chemical Reviews*, *120*(15), 7219-7347. https://doi.org/10.1021/acs.chemrev.9b00846

Hassan, S. S., Williams, G. A., & Jaiswal, A. K. (2018). Emerging technologies for the pretreatment of lignocellulosic biomass. *Bioresource Technology*, *262*, 310-318. https://doi.org/10.1016/j.biortech.2018.04.099

He, C., Shen, F., Tian, D., Huang, M., Zhao, L., Yu, Q., & Shen, F. (2024). Lewis acid/base-mediated deep eutectic solvents intensify lignocellulose fractionation to facilitate enzymatic hydrolysis and lignin nanosphere preparation. *International Journal of Biological Macromolecules*, *254*(Pt 2), 127853. https://doi.org/10.1016/j.ijbiomac.2023.127853

Jaimez, R. E., Barragan, L., Fernandez Nino, M., Wessjohann, L. A., Cedeno Garcia, G., Sotomayor Cantos, I., & Arteaga, F. (2022). *Theobroma cacao* L. cultivar CCN 51: a comprehensive review on origin, genetics, sensory properties, production dynamics, and physiological aspects. *PeerJ*, *10*, e12676. https://doi.org/10.7717/peerj.12676

Jones, D. A., Lelyveld, T. P., Mavrofidis, S. D., Kingman, S. W., & Miles, N. J. (2002). Microwave heating applications in environmental engineering—a review. *Resources, Conservation and Recycling*, *34*, 75–90.

Käldström, M., Lindblad, M., Lamminpää, K., Wallenius, S., & Toppinen, S. (2017). Carbon Chain Length Increase Reactions of Platform Molecules Derived from C5 and C6 Sugars. *Industrial & Engineering Chemistry Research*, *56*(45), 13356-13366. https://doi.org/10.1021/acs.iecr.7b01904

Kappe, O., & Stadler, A. (2005). Microwave Theory. In O. Kappe & A. Stadler (Eds.), *Microwaves in Organic and Medicinal Chemistry*. WILEY-VCH Verlag GmbH & Co.

Kayaputri, L., Djali, M., Sukri, N., & Fazaryasti, R. H. (2020). The antimicrobial effectiveness of the cacao shell and cacao husk combination inhibits pathogenic bacteria in food products. *International Conference on Food and Bio-Industry*, *443*, 1-9. https://doi.org/doi:10.1088/1755-1315/443/1/012077

Klemm, D., Heublein, B., Fink, H. P., & Bohn, A. (2005). Cellulose: fascinating biopolymer and sustainable raw material. *Angew Chem Int Ed Engl*, *44*(22), 3358-3393. https://doi.org/10.1002/anie.200460587

Krahmer, A., Engel, A., Kadow, D., Ali, N., Umaharan, P., Kroh, L. W., & Schulz, H. (2015). Fast and neat--determination of biochemical quality parameters in cocoa using near-infrared spectroscopy. *Food Chem*, *181*, 152-159. https://doi.org/10.1016/j.foodchem.2015.02.084

Kruyeniskia, J., Ferreira, P. J. T., Videira Sousa Carvalho, M. d. G., Vallejos, M. E., Felissia, F. E., & Area, M. C. (2019). Physical and chemical characteristics of pretreated slash pine sawdust influence its enzymatic hydrolysis. *Industrial Crops & Products 130*, 528-536. https://doi.org/https://doi.org/10.1016/j.indcrop.2018.12.075

Kumar, B., Bhardwaj, N., Agrawal, K., Chaturvedi, V., & Verma, P. (2020). Current perspective on pretreatment technologies using lignocellulosic biomass: An emerging biorefinery concept. *Fuel Processing Technology*, *199*. https://doi.org/10.1016/j.fuproc.2019.106244

Kumar Mishra, R., Singh, B., & Acharya, B. (2024). A comprehensive review on activated carbon from pyrolysis of lignocellulosic biomass: An application for energy and the environment. *Carbon Resources Conversion*. https://doi.org/10.1016/j.crcon.2024.100228

Kumar, P., Barrett, D. M., Delwiche, M. J., & Stroeve, P. (2009). Methods for Pretreatment of Lignocellulosic Biomass for Efficient Hydrolysis and Biofuel Production. *Ind. Eng. Chem. Res.*, *48*, 3713–3729.

Kumari, D., & Singh, R. (2018). Pretreatment of lignocellulosic wastes for biofuel production: A critical review. *Renewable and Sustainable Energy Reviews*, *90*, 877-891. https://doi.org/10.1016/j.rser.2018.03.111

Kustov, L. M., Kustov, A. L., & Salmi, T. (2022). Processing of lignocellulosic polymer wastes using microwave irradiation. *Mendeleev Communications*, *32*(1), 1-8. https://doi.org/10.1016/j.mencom.2022.01.001

Liao, Y., de Beeck, B. O., Thielemans, K., Ennaert, T., Snelders, J., Dusselier, M., Courtin, C. M., & Sels, B. F. (2020). The role of pretreatment in the catalytic valorization of cellulose. *Molecular Catalysis*, *487*, 1-34. https://doi.org/10.1016/j.mcat.2020.110883

Liu, E., Das, L., Zhao, B., Crocker, M., & Shi, J. (2017). Impact of Dilute Sulfuric Acid, Ammonium Hydroxide, and Ionic Liquid Pretreatments on the Fractionation and Characterization of Engineered Switchgrass. *BioEnergy Research*, *10*(4), 1079-1093. https://doi.org/10.1007/s12155-017-9868-x

Liu, Y., Nie, Y., Lu, X., Zhang, X., He, H., Pan, F., Zhou, L., Liu, X., Ji, X., & Zhang, S. (2019). Cascade utilization of lignocellulosic biomass to high-value products. *Green Chemistry*, *21*(13), 3499-3535. https://doi.org/10.1039/c9gc00473d

Ma, R., Xu, Y., & Zhang, X. (2015). Catalytic oxidation of biorefinery lignin to value-added chemicals to support sustainable biofuel production. *ChemSusChem*, *8*(1), 24-51. https://doi.org/10.1002/cssc.201402503

Maza Villalobos, S., Nicasio Arzeta, S., Benitez Malvido, J., Ramírez Marcial, N., Alvarado Sosa, E., & Rincón Arreola, D. (2024). Effect of land-use history on tree taxonomic and functional diversity in cocoa agroforestry plantations. *Agriculture, Ecosystems & Environment*, *367*. https://doi.org/10.1016/j.agee.2024.108952

McAnany, P. A., & Murata, S. (2007). America's First Connoisseurs of Chocolate. *Food and Foodways. Explorations in the History and Culture of Human Nourishment*, *15*(1-2), 7-30. https://doi.org/10.1080/07409710701229565

Mellinas, A. C., Jiménez, A., & Garrigós, M. C. (2020). Optimization of microwave-assisted cocoa bean shell waste extraction and evaluation of its antioxidant, physicochemical, and functional properties. *Lwt*, *127*. https://doi.org/10.1016/j.lwt.2020.109361

Meng, X., Wang, Y., Conte, A. J., Zhang, S., Ryu, J., Wie, J. J., Pu, Y., Davison, B. H., Yoo, C. G., & Ragauskas, A. J. (2023). Applications of biomass-derived solvents in biomass pretreatment - Strategies, challenges, and prospects. *Bioresour Technol*, *368*, 128280. https://doi.org/10.1016/j.biortech.2022.128280

Meraj, A., Singh, S. P., Jawaid, M., Nasef, M. M., Alomar, T. S., & AlMasoud, N. (2023). A Review on Eco-friendly Isolation of Lignin by Natural Deep Eutectic Solvents from Agricultural Wastes. *Journal of Polymers and the Environment*, *31*(8), 3283-3316. https://doi.org/10.1007/s10924-023-02817-x

Mingos, D. M. P., & Baghurst, D. R. (1991). Applications of Microwave Dielectric Heating Effects to Synthetic Problems in Chemistry. *Chem. SOC. Rev.*, *20*, 1-47.

Motamayor, J. C., Risterucci, A. M., Lopez, P. A., Ortiz, C. F., Moreno, A., & Lanaud, C. (2002). Cacao domestication I: the origin of the cacao

cultivated by the Mayas. *Heredity (Edinb)*, *89*(5), 380-386. https://doi.org/10.1038/sj.hdy.6800156

Muharja, M., Darmayanti, R. F., Palupi, B., Rahmawati, I., Fachri, B. A., Setiawan, F. A., Amini, H. W., Rizkiana, M. F., Rahmawati, A., Susanti, A., & Putri, D. K. Y. (2021). Optimization of Microwave-Assisted Alkali Pretreatment for Enhancement of Delignification Process of Cocoa Pod Husk. *Bulletin of Chemical Reaction Engineering & Catalysis*, *16*(1), 31-43. https://doi.org/10.9767/bcrec.16.1.8872.31-43

Nguyen, V. T., Tran, N. T. H., & Tran, T. G. (2022). Central composite experimental design for ultrasound-assisted extraction optimization of alkaloid compounds and antioxidant properties from cocoa pod husk *(Theobroma cacao* L.). *Journal of Food Processing and Preservation*, *46*(11). https://doi.org/10.1111/jfpp.17084

Nüchter, M., Ondruschka, B., Bonrath, W., & Gum, A. (2004). Microwave-assisted synthesis – a critical technology overview. *Green Chem.*, *6*(3), 128-141. https://doi.org/10.1039/b310502d

Okiyama, D. C. G., Navarro, S. L. B., & Rodrigues, C. E. C. (2017). Cocoa shell and its compounds: Applications in the food industry. *Trends in Food Science & Technology*, *63*, 103-112. https://doi.org/10.1016/j.tifs.2017.03.007

Pangestu, R., Amanah, S., Juanssilfero, A. B., Yopi, & Perwitasari, U. (2020). Response surface methodology for microwave-assisted extraction of pectin from cocoa pod husk (*Theobroma cacao)* mediated by oxalic acid. *Journal of Food Measurement and Characterization*, *14*, 2126–2133 https://doi.org/ https://doi.org/10.1007/s11694-020-00459-4

Parastar, H., van Kollenburg, G., Weesepoel, Y., van den Doel, A., Buydens, L., & Jansen, J. (2020). Integrating handheld NIR and machine learning to “Measure & Monitor” chicken meat authenticity. *Food Control*, *112*. https://doi.org/10.1016/j.foodcont.2020.107149

Quiceno Suarez, A., Cadena Chamorro, E. M., Ciro Velásquez, H. J., & Arango Tobón, J. C. (2024). By-products of the cocoa agribusiness: high value-added materials based on their bromatological and chemical characterization. *Rev. Facultad Ncional de Agronomia 77(1)*, 10585-10599. https://doi.org/https://doi.org/10.15446/rfnam.v77n1.107738

Rahandi Lubis, M. A., Seng Hua, L., Efri, M., Souvia, R., Petar, A., Robi, A., Luboš, K. á., & Bambang, N. (2024). *Biomass Conversion and Sustainable Biorefinery* (M. A. R. L. S. H. L. E. M. S. R. P. A. R. A. L. u. K. B. Nurhadi, Ed.). Springer Nature. https://doi.org/10.1007/978-981-99-7769-7

Real Pérez, C., Alcalá González, M. D., Romero Sarria, F., Hidalgo López, M. d. C., & Córdoba Gallego, J. M. (2024). Investigating the room- and cryo-milling impact in lignocellulosic biomass and its consequence over pyrolysis and oxidative treatments. *Journal of Cleaner Production*, *437*. https://doi.org/10.1016/j.jclepro.2024.140761

Rocha Balbino, T., Sanchez Munoz, S., Diaz Ruiz, E., Moura Rocha, T., Mier Alba, E., Custodio Inacio, S., Castro Alonso, M. J., de Carvalho Santos-Ebinuma, V., Brandao Pereira, J. F., Santos, J. C., & Silverio da Silva, S. (2023). Lignocellulosic biorefineries as a platform for producing high-value yeast derived pigments - A review. *Bioresour Technol*, *386*, 129549. https://doi.org/10.1016/j.biortech.2023.129549

Roy, S. (2022). *PRE-TREATMENT METHODS OF LIGNOCELLULOSIC BIOMASS FOR BIOFUEL PRODUCTION*. CRC Press is an imprint of Taylor & Francis Group, LLC. https://doi.org/DOI: 10.1201/9781003203414

Roy, S., Lee, K., Lacey, J. A., Thompson, V. S., Keiser, J. R., & Qu, J. (2020). Material Characterization-Based Wear Mechanism Investigation for Biomass Hammer Mills. *ACS Sustainable Chemistry & Engineering*, *8*(9), 3541-3546. https://doi.org/10.1021/acssuschemeng.9b06450

Ruf, F., & Schroth, G. (2003). Chocolate Forests and Monocultures: A Historical Review of Cocoa Growing and Its Conflicting Role in Tropical Deforestation and Forest Conservation. In *Chocolate Forests and Monocultures* (pp. 107-134).

Rusănescu, C. O., Ciobanu, M., Rusănescu, M., & Dinculoiu, R. L. (2024). Pretreatments Applied to Wheat Straw to Obtain Bioethanol. *Applied Sciences*, *14*(4). https://doi.org/10.3390/app14041612

Saini, R., Singhania, R. R., Patel, A. K., Chen, C. W., Piechota, G., & Dong, C. D. (2024). Sustainable production of cellulose and hemicellulose-derived oligosaccharides from pineapple leaves: Impact of hydrothermal pretreatment and controlled enzymatic hydrolysis. *Bioresour Technol*, *398*, 130526. https://doi.org/10.1016/j.biortech.2024.130526

Salmén, L. (2015). Wood morphology and properties from molecular perspectives. *Annals of Forest Science*, *72*(6), 679-684. https://doi.org/10.1007/s13595-014-0403-3

Sánchez, M., Laca, A., Laca, A., & Díaz, M. (2023a). Cocoa Bean Shell as Promising Feedstock for the Production of Poly(3-hydroxybutyrate) (PHB). *Applied Sciences*, *13*(975), 1-14. https://doi.org/https://doi.org/10.3390/app13020975

Sánchez, M., Laca, A., Laca, A., & Díaz, M. (2023b). Cocoa Bean Shell: A By-Product with High Potential for Nutritional and Biotechnological Applications. *Antioxidants (Basel)*, *12*(1028). https://doi.org/10.3390/antiox12051028

Santos, I. A., Conceicao, D. G., Viana, M. B., Silva, G. J., Santos, L. S., & Ferrao, S. P. B. (2021). NIR and MIR spectroscopy for quick detection of the adulteration of cocoa content in chocolates. *Food Chem*, *349*, 129095. https://doi.org/10.1016/j.foodchem.2021.129095

Sarah, M., Hasibuan, I. M., Misran, E., & Maulina, S. (2022). Optimization of Microwave-Assisted Pectin Extraction from Cocoa Pod Husk. *Molecules*,

27(6544), 3-39. https://doi.org/https://doi.org/10.3390/molecules27196544

Shahzadi, T., Mehmood, S., Irshad, M., Anwar, Z., Afroz, A., Zeeshan, N., Rashid, U., & Sughra, K. (2014). Advances in lignocellulosic biotechnology: A brief review on lignocellulosic biomass and cellulases. *Advances in Bioscience and Biotechnology*, *05*(03), 246-251. https://doi.org/10.4236/abb.2014.53031

Sharma, V., Tsai, M. L., Chen, C. W., Sun, P. P., Patel, A. K., Singhania, R. R., Nargotra, P., & Dong, C. D. (2022). Deep eutectic solvents as promising pretreatment agents for sustainable lignocellulosic biorefineries: A review. *Bioresour Technol*, *360*, 127631. https://doi.org/10.1016/j.biortech.2022.127631

Shinde, S. H., Hengne, A., & Rode, C. V. (2020). Lignocellulose-derived platform molecules. In *Biomass, Biofuels, Biochemicals* (pp. 1-31). https://doi.org/10.1016/b978-0-444-64307-0.00001-9

Singh, K., Mehra, S., & Kumar, A. (2024). Recent advances in catalytic conversion of lignin to value-added chemicals using ionic liquids and deep eutectic solvents: a critical review. *Green Chemistry*, *26*(3), 1062-1091. https://doi.org/10.1039/d3gc03439a

Skoczinski, P., Carus, M., Tweddle, G., Ruiz, P., de Guzman, D., Ravenstijn, J., Käb, H., Hark, N., & Dammer, L. (2022). *Bio-based Building Blocks and Polymers* (nova-Institut für politische und ökologische Innovation GmbH, Ed.). https://doi.org/10.52548/cmzd8323

Socha, A. M., Plummer, S. P., Stavila, V., Simmons, B. A., & Singh, S. (2013). Comparison of sugar content for ionic liquid pretreated Douglas-fir woodchips and forestry residues. *Biotechnology for Biofuels*, *6:61*. http://www.biotechnologyforbiofuels.com/content/6/1/61

Song, B., Wu, Z., Yu, Y., & Wu, H. (2020). Hydrothermal Reactions of Biomass-Derived Platform Molecules: Distinct Effect of Aprotic and Protic Solvents on Primary Decomposition of Glucose and Fructose in Hot-Compressed Solvent/Water Mixtures. *Industrial & Engineering Chemistry Research*, *59*, 7336-7345. https://doi.org/https://dx.doi.org/10.1021/acs.iecr.0c01514

Stass, D. V., Woodward, J. R., Timmel, C. R., Hore, P. J., & McLauchlan, K. A. (2000). Radiofrequency magnetic ®eld e□ects on chemical reaction yields. *Chemical Physics Letters*, *329*, 15-22.

Sun, D., Lv, Z. W., Rao, J., Tian, R., Sun, S. N., & Peng, F. (2022). Effects of hydrothermal pretreatment on the dissolution and structural evolution of hemicelluloses and lignin: A review. *Carbohydr Polym*, *281*, 119050. https://doi.org/10.1016/j.carbpol.2021.119050

Taiwo, O. E., & Osinowo, F. A. O. (2001). Evaluation of various agro-wastes for traditional black soap production. *Bioresource Technology*, *79*, 95-97.

Tišma, M., Bucić Kojić, A., & Planinić, M. (2021). Bio-based Products from Lignocellulosic Waste Biomass. *Chemical & Biochemical Engineering Quarterly*, *35*(2), 139-156. https://doi.org/10.15255/cabeq.2021.1931

Torres Mayanga, P. C., Lachos Perez, D., Mudhoo, A., Kumar, S., Brown, A. B., Tyufekchiev, M., Dragone, G., Mussatto, S. I., Rostagno, M. A., Timko, M., & Forster Carneiro, T. (2019). Production of biofuel precursors and value-added chemicals from hydrolysates resulting from hydrothermal processing of biomass: A review. *Biomass and Bioenergy*, *130*. https://doi.org/10.1016/j.biombioe.2019.105397

Tsai, W. T., & Huang, P. C. (2018). Characterization of acid-leaching cocoa pod husk (CPH) and its resulting activated carbon. *Biomass Conversion and Biorefinery*, *8*(3), 521-528. https://doi.org/10.1007/s13399-018-0302-1

Vasquez, Z. S., de Carvalho Neto, D. P., Pereira, G. V. M., Vandenberghe, L. P. S., de Oliveira, P. Z., Tiburcio, P. B., Rogez, H. L. G., Goes Neto, A., & Soccol, C. R. (2019). Biotechnological approaches for cocoa waste management: A review. *Waste Manag*, *90*, 72-83. https://doi.org/10.1016/j.wasman.2019.04.030

Wang, N., & Wang, P. (2016). Study and application status of microwave in organic wastewater treatment – A review. *Chemical Engineering Journal*, *283*, 193-214. https://doi.org/10.1016/j.cej.2015.07.046

Wang, N., Xu, A., Liu, K., Zhao, Z., Li, H., & Gao, X. (2024). Performance of green solvents in microwave-assisted pretreatment of lignocellulose. *Chemical Engineering Journal*, *482*. https://doi.org/10.1016/j.cej.2024.148786

Werpy, T., & Petersen, G. (2004). *Top Value-Added Chemicals From Biomass*.

Xia, Y., Li, X., Zhuang, J., Wang, W., Abbas, S. C., Fu, C., Zhang, H., Chen, T., Yuan, Y., Zhao, X., & Ni, Y. (2024). Exploitation of function groups in cellulose materials for lithium-ion batteries applications. *Carbohydr Polym*, *325*, 121570. https://doi.org/10.1016/j.carbpol.2023.121570

Xiao, T., Hou, M., Guo, X., Cao, X., Li, C., Zhang, Q., Jia, W., Sun, Y., Guo, Y., & Shi, H. (2024). Recent progress in deep eutectic solvent(DES) fractionation of lignocellulosic components : A review. *Renewable and Sustainable Energy Reviews*, *192*. https://doi.org/10.1016/j.rser.2023.114243

Yan, L., Ma, R., Li, L., & Fu, J. (2016). Hot Water Pretreatment of Lignocellulosic Biomass: An Effective and Environmentally Friendly Approach to Enhance Biofuel Production. *Chemical Engineering & Technology*, *39*(10), 1759-1770. https://doi.org/10.1002/ceat.201600394

Yang, Y., Zhang, M., Zhao, J., & Wang, D. (2023). Effects of particle size on biomass pretreatment and hydrolysis performances in bioethanol conversion. *Biomass Conversion and Biorefinery*, *13*(14), 13023-13036. https://doi.org/10.1007/s13399-021-02169-3

Yuan, S., Tian, M., & Lu, X. (2006). Microwave remediation of soil contaminated with hexachlorobenzene. *J Hazard Mater*, *137*(2), 878-885. https://doi.org/10.1016/j.jhazmat.2006.03.005

Yue, P., Hu, Y., Tian, R., Bian, J., & Peng, F. (2022). Hydrothermal pretreatment for the production of oligosaccharides: A review. *Bioresour Technol*, *343*, 126075. https://doi.org/10.1016/j.biortech.2021.126075

Zafeer, M. K., Menezes, R. A., Venkatachalam, H., & Bhat, K. S. (2023). Sugarcane bagasse-based biochar and its potential applications: a review. *Emergent Materials*, *7*(1), 133-161. https://doi.org/10.1007/s42247-023-00603-y

Zhang, Y. M., Wang, P., Han, N., & Lei, H. F. (2007). Microwave Irradiation: A Novel Method for Rapid Synthesis of D, L-Lactide. *Macromolecular Rapid Communications*, *28*(4), 417-421. https://doi.org/10.1002/marc.200600668

Zheng, Y., Shi, J., Tu, M., & Cheng, Y.-S. (2017). Principles and Development of Lignocellulosic Biomass Pretreatment for Biofuels. In E. Inc. (Ed.), *Advances in Bioenergy* (Vol. 2, pp. 1-68). https://doi.org/10.1016/bs.aibe.2017.03.001

Zhihui, A., Peng, Y., & Xiaohua, L. (2005). Degradation of 4-chlorophenol by microwave irradiation enhanced advanced oxidation processes. *Chemosphere*, *60*(6), 824-827. https://doi.org/10.1016/j.chemosphere.2005.04.027

Zhu, Z., Rezende, C. A., Simister, R., McQueen-Mason, S. J., Macquarrie, D. J., Polikarpov, I., & Gomez, L. D. (2016). Efficient sugar production from sugarcane bagasse by microwave-assisted acid and alkali pretreatment. *Biomass and Bioenergy*, *93*, 269-278. https://doi.org/10.1016/j.biombioe.2016.06.017

2. ANTIOXIDANT ACTIVITY AND MINERAL AND LIGNOCELLULOSIC PROFILE OF COCOA INDUSTRY RESIDUES FOR POTENTIAL UTILIZATION

ABSTRACT

Mexico is the thirteenth largest cocoa-producing country in the world, producing 20,000 tons of cocoa per year, or 0.4% of world production. The states of Chiapas and Tabasco are the largest producers. Sixteen million tons of cocoa pod hulls (CPH) are produced worldwide each year, and their chemical properties make them a potential source of value-added products. The objective is to determine the antioxidant activity, mineral profile, and quantification of the leading chemical constituents of CPH (*Theobroma cacao* L. var. Carmelo). This study quantified and determined the primary organic and inorganic constituents of CPH (Theobroma cacao L. var. Carmelo) using analytical techniques (TAPPI, ASTM, and other methods reported in the literature). An extractable content in organic solvents of 21% was obtained. The cellulose content was 30.5%, hemicellulose 19%, lignin 23.9%, and ash 7.7%. The phenolic content was 24.59 ± 0.93 mg EAG/g, and the flavonoid content was 2.35 ± 0.24 mg EC/g. The antioxidant activity was determined to be 304.84 ± 57.59 mg ET for ABTS and 145.80 ± 3.84 for FRAP. About two-thirds of the carbohydrates have the potential to be used for sweeteners, thickeners, and ethanol. In contrast, about one-third of the lignin can be used for pesticides, emulsifiers, resins, polyurethanes, adhesives, and other applications.

Keywords: *Theobroma cacao* L. Carmelo variety; cocoa pod husk; phenolic content; flavonoid content; ABTS-FRAP.

Doctoral Thesis: Doctorado en Ciencias Agroalimentarias, Universidad Autónoma Chapingo.

Author: Edna Elena Suárez Patlán, M. C.
Advisor: Holber Zuleta Prada, Dr.

2.1. Introduction

Theobroma cacao L. has been cultivated in Central America since pre-Columbian times and has spread from the Amazon Basin of Guyana and the Upper Orinoco to southern Mexico and the tropics of Africa and Asia (Espinosa García et al., 2015; Porto de Souza Vandenberghe et al., 2022). The motivation for this crop today is to respond to the growing demand for cocoa, mainly in Europe, the United States, and Asia, and to contribute to the economic and social development of the agricultural sector in more than 70 countries in the tropics, with approximately 5 million producers worldwide (Alonso Báez et al., 2020; Gómez Hoyos et al., 2020; Porto de Souza Vandenberghe et al., 2022). The International Cocoa Organization reports that the world's major cocoa producers are Côte d'Ivoire, Ghana, and Ecuador. In the 2020/2021 harvest, 5.24 million tons of cocoa will be produced, with African countries contributing about 75% of the world's annual production, with Côte d'Ivoire being the most representative with about 44%. Mexico is the thirteenth largest cocoa-producing country in the world, with Chiapas and Tabasco being the states with the highest production in the country, with just over 20,000 tons of cocoa per year, representing 0.4% of world production (Alonso Báez et al., 2020). The cocoa fruit is composed of approximately 10% of the seed, as this is the product that is marketed; the other 90% of the fruit remains as residual biomass, which is composed of the cocoa bean shell (CBS), mucilage, or pulp, placenta, and cocoa pod husk (CPH). The CBS is the covering that covers the cocoa seed, representing between 10 and 20% of the total weight of the seed; the CPH is the covering of the ripe fruit; once the seed is removed, the pulp and placenta represent about 5% of the weight of the fruit, these parts represent the largest volume of residual biomass, between 70 and 80% of the weight of the fruit, so it is estimated that more than 16 million tons of CPH per year around the world (Sánchez et al., 2023b; Vasquez et al., 2019). CPH is an undervalued resource; it tends to remain in growing areas and accumulates in large piles, so its decomposition can cause environmental and phytosanitary problems, such as plant diseases that directly affect crop yields (Nelson et al., 2021). This shell is composed of epicarp, mesocarp, a sclerotic

lamina, and endocarp, whose chemical composition is cellulose, hemicellulose, lignin, pectins, and ash (Campos Vega et al., 2018; Grob et al., 2021). Cellulose is a linear polymer of D-glucose units linked by β-1,4 bonds. Hemicellulose is a branched polysaccharide composed of pentoses and hexoses. Lignin is an aromatic heteropolymer composed of phenylpropane units (p-coumarilic, sinapyl, and coniferyl alcohols) (Zheng et al., 2017). These compounds can be depolymerized into molecules that serve as the basis for obtaining value-added chemical products; for example, the fermentation of cellulose and hemicellulose sugars could yield biofuels or sweeteners from lignin, pesticide dispersants, emulsifiers, or their incorporation into polyurethanes, polyesters, etc. (Asiedu et al., 2019; Chávez Sifontes & Domine, 2013; Porto de Souza Vandenberghe et al., 2022). Likewise, its pectin content, composed mainly of uronic acid, has potential as a food additive (Paz Cedeno et al., 2022). Finally, phenolics and flavonoids with oxido-reducing or antioxidant effects could be used in functional foods (Valadez Carmona et al., 2017). Cocoa pod husk is a potential, renewable, low-cost, readily available resource. It is attractive to the industrial sector and of interest to the scientific community, which is seeking alternatives for valorization and the development of technologies and patents for their use. This could bring additional benefits to the cocoa production chain (Porto de Souza Vandenberghe et al., 2022). Three types of cocoa are recognized: two morpho-geographical groups, Criollo and Forastero cocoa, and a third derived from the natural cross between Criollo and Forastero, Trinitary. Research has mainly focused on Forastero and Trinitary, while little is known about Mexican Criollo (Ricaño Rodríguez et al., 2018). In Mexico, there is a variety called "Carmelo," which is recognized under Breeder's Title 1036, granted to the producer C. Carlos Hernández Echeverría, who, through a breeding process obtained this variety, which is considered homogeneous, distinct, and stable (DOF: 29/09/2015). The Carmelo variety has Criollo as its genetic origin (Alonso Báez et al., 2020). In this sense, the little information that exists on Mexican Criollo cocoa is mostly related to the cocoa seed, so there is even less information on the by-products concerning vegetative varieties, and much less is known about the Carmelo

variety due to the breeder's title, which is valid until 2031, this variety has been little propagated and less studied. Therefore, this research aimed to determine the antioxidant activity and chemical characterization of the main components of the pod's shell of *Theobroma cacao* L. variety Carmelo. To contribute to the knowledge of this variety and present elements of valorization of a highly available residue generated by the cocoa agroindustry.

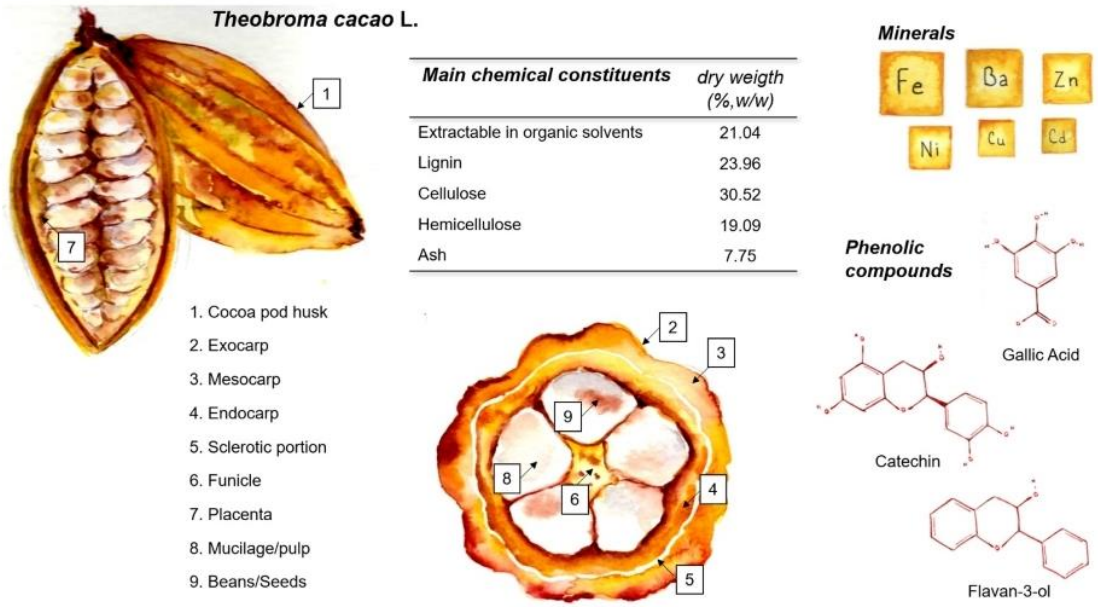


Figure 9 Chemical characterization of *Theobroma cacao* L. var. Carmelo.

2.2. Materials and methods

2.2.1. Raw material and study area

The raw material used in this study was the cocoa pod husk of *Theobroma cacao* L. variety Carmelo (CPH), which consists of the exocarp, mesocarp, sclerotic layer, and endocarp. The samples were collected from two locations: Rancheria Río Seco in the municipality of Cunduacan [latitude: 18° 7'55.90" N, longitude: 93° 18'4.49" W] and the farm Jesús María in the city of Comalcalco [latitude: 18° 11'0.22" N, longitude: 93° 14'28.02" W], both located in the state of Tabasco, Mexico. The collection sites were 10 and 13 meters above sea level (Figure 1). C. Carlos Hernández Echeverría provided the plant material for this study.

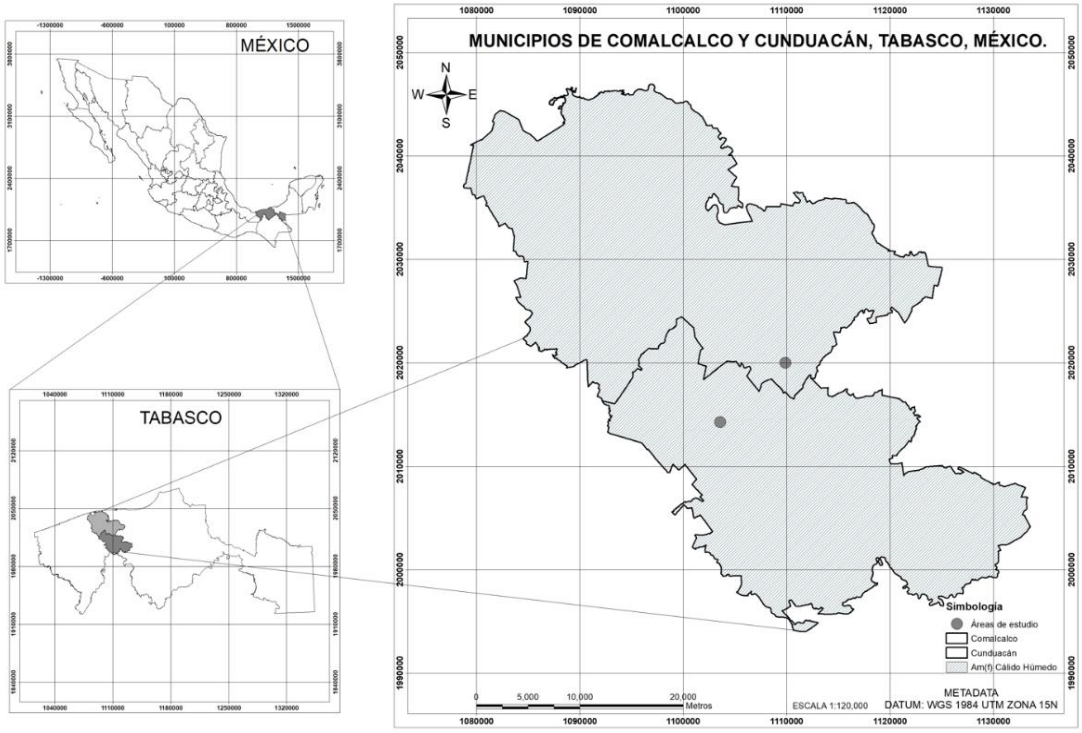


Figure 10 Location of the Theobroma cacao L. variety Carmelo collection meters above sea level (Figure 1).

2.2.2. Reagents

Reagents: 96% sulfuric acid (H_2SO_4), 99% glacial acetic acid (CH_3COOH), 80-85% sodium chlorite ($NaClO_2$), 99% sodium hydroxide (NaOH), 99.5% sodium carbonate (Na_2CO_3), Folin-Ciocalteu phenol reagent, diammonium salt (ABTS) 2,2'-Azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) 98%, and (±)-6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid, also known as Trolox, were purchased from SUPELCO (Merck KGaA, Saint Louis, Missouri, USA) and were used without further purification.

2.2.3. Cocoa pod husk preparation

The cocoa pod husk (CPH) was dried at room temperature with partial sun exposure for 18 days. Subsequently, they were ground in a laboratory mill (Thomas Wiley 4, Pennsylvania, USA). The ground material was then sieved

using U.S. STD 45 and 60 mesh sieves (W.S. Tyler, Ohio, USA) to obtain particle size classifications of 355 µm and 250 µm, respectively. The meal retained on the 250 µm particle size 60 mesh was used for all chemical determinations. Finally, the samples were weighed appropriately for each determination and dried to a constant weight in a conventional oven at 102°C.

2.2.4. Main chemical components

The determination of lignocellulose was based on the use of two standards of the Technical Association of the Pulp and Paper Industry (TAPPI), one standard of the American Society for Testing and Materials (ASTM), and a widely used chlorite-acetic acid method for the disintegration of lignocellulosic material (Rowell et al., 2012). Standard T 204 cm 97 was used with modifications for sample preparation, solvent extraction, and quantification of extracts (Mohamad & Rushdan, 2014). The solvents used in this standard were replaced by hexane, dichloromethane, and methanol to eliminate the use of benzene, which is highly toxic and carcinogenic, as specified in the standard. For the determination of lignin, standard T 222 om-.02 was used (Mohamad & Rushdan, 2014), and for the quantification of ash, standard T 211 om-02 (Mohamad & Rushdan, 2014) was used. Based on the method proposed by Rowell et al. (Rowell et al., 2012) and standard ASTM D 103-77 (Blankenhorn et al., 1985), the holocellulose and cellulose contents were determined, respectively. Finally, the hemicellulose content was quantified by weight difference. All determinations were performed in quadruplicate, and the results were expressed as mean percentages ± standard deviation.

2.2.5. Minerals Profile

To profile and quantify the minerals present in cocoa pod husks, 0.5 g of CPH was placed in a microwave Teflon cell (Xpress Plus CEM), followed by the addition of 10 mL of 69% HNO_3. The sample was then introduced into a CEM microwave instrument, model Mars 6, using the US EPPA 3051 method for sample digestion. The resulting digestion was volumized to 25 mL with deionized water to facilitate

the identification and quantification of the concentration of each metal. Subsequently, an iCAP Q ICP-MS instrument (Thermo Scientific, Bremen, Germany) was used to identify and quantify the concentration of each metal, employing a calibration curve of standard solutions. Mineral determinations were conducted in triplicate, and the results were expressed as the mean in µg/g ± standard deviation.

2.2.6. Methanolic extract

An extract was prepared with a 0.5 g dry base of CPH and 25 mL of 80% methanol to determine antioxidant properties, adjusted to pH 3 with 0.1 N HCl. 1 N, this mixture was subjected to vortex agitation at 3000 rpm for 3 minutes, followed by sonication for 15 minutes, then placed in a rotary incubator at 30°C and 150 rpm for 30 minutes; finally, it was centrifuged at 3000 rpm for 20 minutes, the excess was transferred to a volumetric flask and made up to 25 mL with 80% methanol. This extract was used to determine the antioxidant properties of the cacao pod husk. The method used was reported by Hernández et al. (Hernández Rodríguez et al., 2019), who indicated that, since phenolic compounds are polar substances, they can be extracted using organic solvents that have been acidified to promote protonation.

2.2.7. Phenolic content

The total phenolic content of CPH was determined by the method reported by Hernández et al. (Hernández Rodríguez et al., 2016). Aliquots of 25 µg of methanolic extract of CPH were prepared, to which 25 µL of distilled water and 20 µL of Folin-Ciocalteau reagent were added, then 30 µL of 20% Na_2CO_3 solution was added and allowed to react for 30 min in the absence of light, the absorbance was measured at a wavelength of 760 nm in a µQuant microplate reader (Biotek, Instruments Inc., USA). The result was expressed as milligram gallic acid equivalents per gram of cocoa pod dry weight (mg EAG/g).

2.2.8. Flavonoid content

The total flavonoid content was determined according to the method of Hernández *et al*. (Hernández Rodríguez et al., 2019) 0.5 mL of CPH methanolic extract was mixed with 2.5 mL of distilled water and 0.15 mL of 5% $NaNO_2$ solution. The mixture was allowed to stand for 6 minutes, after which 0.3 mL of 10% $AlCl_3$ $6H_2O$ was added. Following a 5-minute rest, 1 mL of 5% NaOH was added, and the solution was vortexed at 3000 rpm for 3 minutes. Subsequently, 200 µL of the aliquot was added to the microplate in quadruplicate, and absorbances were measured at 510 nm using the µQuant microplate reader (Biotek Instruments Inc., USA). The result was expressed as milligram catechin equivalents per gram of cocoa pod husk dry matter (mg EC/g).

2.2.9. ABTS Assay

The method reported by Arzeta et al. (Arzeta Ríos et al., 2020) was used for the corresponding assay. First, solution one was prepared by combining 7.4 mM ABTS and 2.6 mM sodium persulfate in equal volumes of 10 mL and then incubating the mixture for 16 hours at room temperature without light. Next, 600 µL of solution one was taken and topped up to 10 mL with absolute methanol. In a microplate reader µQuant (Biotek, Instruments Inc., USA), 20 µL of CPH extract, 20 µL of Trolox solution in concentrations ranging from 5 to 60 µL, and 180 µL of solution one were placed. The absorbances were then read at 734 nm after 10 minutes, and the data from the calibration curve were used to determine the results, which are expressed in µmol of Trolox equivalents per gram of CPH dry base (µmol ET/g) required to trap the ABTS radical.

2.2.10. FRAP Assay

The FRAP assay followed the method described by Hernández *et al*. (Hernández Rodríguez et al., 2019). The FRAP reagent was prepared using a mixture of sodium acetate buffer (300 mM, pH 3.6), ten mM TPTZ dilution (40 mM HCl), and 20 mM ferric chloride dilution in a 10:1:1 ratio. An aliquot (20 µL) of the extract

was mixed with 180 µL of the FRAP dilution and 60 µL of distilled water. The absorbance was measured at 595 nm using a µQuant microplate reader (Biotek, Instruments Inc., USA). The results were expressed as µmol Trolox equivalents per gram of CPH dry weight (µmol ET/g). The determinations of total phenolic content, flavonoid content, and the ABTS assay were performed in quadruplicate, and the results were expressed as the mean ± standard deviation.

2.3. Results and discussion

The Carmelo variety is characterized by its elongated fruit with an absent to weak basal constriction, an obtuse apex shape, medium length, a large diameter, a moderately elongated length-to-diameter ratio, an orange exocarp, a moderately rough surface, medium depth of furrows, medium thickness, light cream flesh, a low number of white seeds, and a sweet taste. The determination of extractable in polar solvents, lignocellulosic constituents, mineral profile and content, quantification of phenolics, flavonoids, and antioxidant activity of *Theobroma cacao* L. variety Carmelo pod husk is shown in Table 7.

Table 7 Composition of the major chemical constituents of the cocoa pod husk of *Theobroma cacao* L. variety Carmelo

Chemical Constituent		***Average ± σ This study***	***Average Other research***	***Reference***
Removable components	Hexane extractable	0.68 ± 0.01	ND	This work
	Dichloromethane extractables	1.32 ± 0.05	ND	This work
	Methanol extractables	19.0 ± 1.29	ND	This work
	Total extractable	**21.0** ± 1.35	23.7	(Titiloye et al., 2013)
Lignocellulosic (%)	Lignin	**23.9** ± 0.21	24.2	(Sandesh et al., 2020)
			27.9	(Redgwell et al., 2003)
	Cellulose	**30.5** ± 1.03	14.7	(Daud et al., 2014)
			30.4	(Titiloye et al., 2013)
			28.2	(Sandesh et al., 2020)
			35.4	(Daud et al., 2014)
			35.0	(Campos Vega et al., 2018)

	Hemicellulose	**19.1** ± 1.44	16.7 22.4	(Sandesh et al., 2020) (Redgwell et al., 2003)
	Ashes	**7.75** ± 0.10	8.4 / 8.3 6.7 / 10 6.4 / 8.4 6.7 / 13	(Martínez et al., 2012) (Campos Vega et al., 2018) (Lu et al., 2018) (Vasquez et al., 2019)
Minerals (µg/g)	Al	**22.8** ± 2.46	ND	This work
	V	0.06 ± 0.00	ND	This work
	Cr	0.37 ± 0.01	ND	This work
	Fe	**164.4** ± 1.79	58 466	(Vriesmann et al., 2011) (Aregheore, 2002)
	Co	0.15 ± 0.01	ND	This work
	Ni	**10.7** ± 0.30	3.10 / 6.27 0.55	(Barraza et al., 2021) (Moyin Jesu, 2007)
	Cu	**6.91** ± 0.24	6.18	(Vriesmann et al., 2011)
	As	0.04 ± 0.00	ND	This work
	Cd	1.00 ± 0.02	0.60 / 0.80 0.16 / 0.27	(Gramlich et al., 2016)
	Ba	**63.6** ± 0.36	98 41.5 / 47.0	(Barraza et al., 2021) (Barraza et al., 2021)
	Zn	**63.4** ± 0.86	42.9 / 45.8 90 39	(Gramlich et al., 2016) (Aregheore, 2002) (Vriesmann et al., 2011)
	Hg	0.03 ± 0.02	ND	This work
	Pb	0.18 ± 0.01	ND	This work
(mg GAE/g) *	Total phenolic compounds	**24.6** ± 0.93	24.8 / 27.3 17.6 19.8 34.9 35.5 16.6 3.23 2.06 3.52	(Nieto Figueroa et al., 2020) (Jamaluddin et al., 2022) (Jamaluddin et al., 2022) (Jamaluddin et al., 2022) *(Castro Vargas et al., 2019)* (Siqueira Melo et al., 2011) (Valadez Carmona et al., 2017) (Martínez et al., 2012) (Martínez et al., 2012)
(mg CE/g) **	Flavonoids	**2.35** ± 0.24	1.10 / 1.30 0,97 0,56	(Nieto Figueroa et al., 2020) (Valadez Carmona et al., 2017) (Castro Vargas et al., 2019)

			227.70 /147.00	(Nieto Figueroa et al., 2020)
(µmol TE/g) ***	ABTS	**304.8** ± 57.2	30.60	(Valadez Carmona et al., 2017)
			24.13	(Martínez et al., 2012)
			42 / 24	(Lu et al., 2018)
(µmol TE/g) ***	FRAP	**145.8** ± 3.84	521.73	(Jamaluddin et al., 2022)
			1.98	(Martínez et al., 2012)

** mg gallic acid equivalents/g (dry basis)*
*** mg catechin equivalents/g (dry basis)*
**** Trolox Equivalent/g (dry basis)*

2.3.1. Extractables in organic solvents

In the determination of extractives from lignocellulosic materials using ethanol-benzene solvents, fats, waxes, phytosterols, low molecular weight carbohydrates, resins, and salts can be obtained (Mohamad & Rushdan, 2014). The total content of cocoa pod husk extracts obtained in this study was 21%, of which 19% corresponds to those soluble in methanol, which could be low molecular weight carbohydrates and salts, 1.32% corresponds to those soluble in dichloromethane, and a percentage of 0.68% are those extracted in hexane, which according to the TAPPI (Mohamad & Rushdan, 2014), standard could be waxes and fats. Despite the exclusion of benzene as a solvent of low polarity compounds, the content of extractable determined in this work is close to the 23.66% reported by Titiloye *et al*. (Titiloye et al., 2013) for cocoa pod husk, who did not specify the composition of these extractables. In this sense, it can be observed that the substitution of benzene did not alter the release of CPH extractables for subsequent lignocellulosic determinations, in which it is necessary to have the material free of these extractables.

2.3.2. Lignin content

The lignin content in this study was 23.96%, similar to the 24.16% reported by Sandesh *et al*. (Sandesh et al., 2020) and close to the 27.9% reported by Redgwell et al. (Redgwell et al. using the Klason lignin method. On the other hand, the lignin content reported by Daud *et al.* (Daud et al., 2014) was 14.7%, much lower than that found in the present study and those cited above. According to

Barhoum *et al*. (Barhoum et al., 2020), the lignin content reported for other lignocellulosic materials, such as wood, is usually between 10 and 25% depending on the species; some values reported for eucalyptus and pine are 21.5% and 20% respectively, very similar to those reported in this study. Lignin has high value-added applications as a fertilizer, dispersant, and chelating agent. Because of its high carbon content, it is a potential precursor for functional carbon materials such as activated carbon, carbon electrodes, etc. The paper industry generates 75 million tons of lignin annually from wood fibers in pulp production (Yao et al., 2022). However, this industry has been replacing wood fibers with recycled fibers as an environmental strategy (Jochem et al., 2021), so producing lignin from cocoa pod husk could be a viable option.

2.3.3. Cellulose and hemicellulose content

The cellulose content of the raw material studied was 30.52%, which is similar to the 30.41% reported by Titiloye *et al*. (Titiloye et al., 2013) and close to the 28.25% reported by Sandesh et al. (Sandesh et al., 2020). Other cellulose contents of cocoa pod husks found in the literature are 35.4% and 35%, reported by Daud et al. (Daud et al., 2014) and Campos et al. (Campos Vega et al., 2018) respectively, using the chlorite method as in this work. The cellulose content of other lignocellulosic materials, such as pine and eucalyptus wood, is 40 and 54%, respectively (Barhoum et al., 2020), much higher than that of cocoa pod husk. The hemicellulose content of the raw material studied was 19.09%, which is between the values of 16.75% and 22.4% of cocoa pod hemicelluloses reported by Sandesh *et al*. (Sandesh et al., 2020) and Redgwell *et al*. (Redgwell et al., 2003). According to Barhoum *et al.* (Barhoum et al., 2020), lignocellulosic wood biomass has a hemicellulose content ranging from 11% to 35.9% for hardwoods and 24% to 27% for softwoods. Agricultural and forestry residues represent a viable option for producing alternative fuels. (Devi et al., 2021) Cocoa pod husks of the Carmelo variety have the potential as a source of cellulose and hemicellulose, comparable to cocoa husks from Chiapas, Mexico, which were

studied and reported by Hernández *et al*. (Hernández Mendoza et al., 2021) as having potential for bioethanol production.

2.3.4. Ash content

The ash content reported in this study is 7.75%, similar to that reported by Martinez *et al.,* (Martínez et al., 2012) of 8.42 and 8.32% for cacao pod husk of different origins, and between the ranges of 6.7 to 10.02%, 6.4 to 8.4%, and 6.7 to 13% reported by Campos *et al*., (Campos Vega et al., 2018), Lu *et al*., (Lu et al., 2018) and Vásquez *et al*., (Vasquez et al., 2019) respectively. This ash content is an average value of the mineral salts that are part of the lignocellulosic composition of this material; it is about 25 times higher than the ash content reported for other highly available lignocellulosic materials, such as pine wood of different species, which ranges from 0.26 to 0.52% (Bernabé Santiago et al., 2013). CPH ash has been studied as a fertilizer replacement in partial or total application (Campos Vega et al., 2018). CPH has also been studied as a compost and has been reported to increase soil pH and improve soil fertility by incorporating most of the essential elements, improve cocoa growth, provide a valuable alternative to fungicides, and reduce the severity of black pod disease when properly managed and not just left as waste in the growing areas (Doungous et al., 2018).

2.3.5. Mineral Profile

About a high ash content present in cocoa pod husk, this material has been widely recognized to be rich in mineral nutrients, concentrations of K (2768 - 112.04 mg/100 g), Mg (110.9 - 21.23 mg/100 g), Ca (254 - 6.15 mg/100 g), Zn (39.74 - 7.22 mg/100 g), Mn (35.72 - 7. 32 mg/100 g), Na (10.5 - 0.47 mg/100 g), Cu (8.55 - 6.18 mg/100 g), Fe (5.8 - 5.04 mg/100 g) and Se (0.01 mg/100 g) (Vargas Arana et al., 2022; Vriesmann et al., 2011). Minerals such as Mn, Cu, Zn, Ni, and Mo are required by plants as micronutrients, but others are considered toxic to human health, such as As, Cd, Cr, and Pb (Barraza et al., 2021). Cadmium accumulates in the cocoa bean, its derivatives and by-products, and other foods. Cadmium is

a metal toxic to human health, has no biological function, occurs naturally in soil, and is increasing worldwide due to anthropogenic inputs (Gramlich et al., 2016). Cocoa pod husk has been used locally in some West African countries to produce "black" soap. It has also been extensively used in composts, and its potential as organic matter for soil replenishment, fertilization, and biological control has been extensively studied (Doungous et al., 2018; Moyin Jesu, 2007; Olubunmi Kayode et al., 2018). However, the possibility of using it as an additive in the food industry has been raised, so it was important to measure some other minerals that have not been reported and could be considered toxic to human health besides those already known. In this study, 13 concentrations of minerals were determined in cocoa pod husk of the Carmelo variety, presented in descending order of concentration; with the highest values between 164 and 63 µg/g, Fe > Ba > Zn were found, followed by those with medium values between 22 and 1 µg/g, Al > Ni > Cu > Cu > Cd and finally those with low values between 0.37 and 0.03 µg/g, Cr > Pb > Co> V > As > Hg. Vriesmann *et al*. (Vriesmann et al., 2011) coincidentally reported higher Fe content about Zn, with values of 58 and 39 µg/g, respectively. Aregheore *et al*. (Aregheore, 2002) agree with this trend, reporting 466 and 90 µg/g for Fe and Zn, respectively, but with a value for Zn much higher than those reported by the generality. Barraza *et al*. (Barraza et al., 2021) reported Zn concentrations of 41.5 and 47 mg kg^{-1} and Ba concentrations of 98 mg kg^{-1} (Gramlich et al., 2016), reported Zn concentrations of 42.9 and 45.8 mg kg^{-1} in the shells of two varieties of cacao pods husk from Ecuador, Bolivia, and Honduras, respectively; the latter values are close to those reported for Zn in this study; however, those for Ba are much higher. Concerning Cu, concentrations of 0.55 mg Kg^{-1} and 6.18 mg/100 g of cacao pod husks were reported by Moyin (Moyin Jesu, 2007) and Vriesman *et al*., (Vriesmann et al., 2011) respectively, which are similar to those reported in this study. Barraza *et al*. (Barraza et al., 2021) reported mean Ni concentrations of 3.10 and 6.27 mg kg^{-1} in two varieties of cocoa pod husk from Ecuador, which are significantly lower than those found in this study for the Carmelo variety. On the other hand, Gramlich *et al*. (Gramlich et al., 2016) reported Cd concentrations ranging from 0.6 to 0.8 and 0.16 to 0.27

mg/kg^{-1} for cocoa pod husk and seeds, respectively, observing that the concentration of Cd in the shells was significantly higher than that in the seeds. In this sense, the value reported in this study for Cd is relatively higher. Finally, no comparative references in cocoa pod husk exist for Cr, Pb, Co, V, As, and Hg. Still, the concentrations determined by this study are low for the remaining minerals that make it up and about Cd, as a reference of contaminant and toxic minerals. The concentrations of minerals in this material will vary according to the origin, type of soil, type of crop, crop management, and surrounding activity association (Barraza et al., 2021; Gramlich et al., 2016). However, if this material were to be considered for any food use, it would be necessary to verify that the volume contribution of contaminant minerals does not exceed the 2.5 µg/kg body weight tolerable monthly intake concentration established by the Joint FAO/WHO Expert Committee on Food Additives and the EFSA Panel on Contaminants in the Food Chain (European Food Safety Authority, 2012).

2.3.6. Total phenolic and flavonoid content

The total phenolic content determined in this study for cacao nibs of variety Carmelo extracted with methanol-water was 24.59 mg EAG/g dry base sample; this result is comparable to the range of values reported by Nieto *et al*. (Nieto Figueroa et al. 2020) from 24.8 to 27.3 mg EAG/g of cacao nibs from Tabasco Mexico dried by different methods. It is within the range reported by Jamaluddin *et al*. (Jamaluddin et al., 2022) for phenolic contents of acetone (17.62 mg EAG/g), ethanol (19.83 mg EAG/g), and methanol (34.88 mg EAG/g) extracts of cocoa pod husk from Malaysia. These values are comparable to 35.53 mg EAG/g and 16.57 mg EAG/g of mango and grape peel reported by Castro *et al*. (Castro Vargas et al., 2019) and Siqueira et al. (Siqueira Melo et al., 2011) respectively. On the other hand, Valadez *et al*. (Valadez Carmona et al., 2017) reported a phenolic content of 323.7 mg EAG/100 g of cocoa pod husk from Chiapas, Mexico, which is similar to the phenolic content of ethanol (206.67 - 227 mg EAG/100 g) and methanol-acetone (352.67 - 365.33 mg EAG/100 g) extracts of cocoa pod husk from two locations in Ecuador reported by Martinez *et al*.,

(Martínez et al., 2012). These values are significantly lower than those found in the present study and those reported by Nieto *et al*. (Nieto Figueroa et al., 2020) and Jamaluddin et al. (Jamaluddin et al., 2022). However, the trend of higher phenolic content in methanol extracts is maintained in all cases. The phenolic content of cocoa nib extracts depends on the variety or genotype, the origin of the raw material, the extraction method (technology, conditions, and solvents), and the pre-treatment. The time and method of drying have a significant influence, as dehydration of the raw material at low temperatures results in higher levels of phenolics, flavonoids, and antioxidant activity (Sánchez et al., 2023b; Valadez Carmona et al., 2017). Regarding the flavonoid content of the cocoa pod husk of the Carmelo variety of this study, a value of 2.35 mg EC/g dry base sample was determined, which is close to the value of 1.1 - 1.3 mg ER/g reported by Nieto *et al*. (Nieto Figueroa et al., 2020) and higher than the value of 97 mg EE/100 g reported by Valadez *et al*., (Valadez Carmona et al., 2017) for cacao nibs from Chiapas, Mexico. In addition to the flavonoid content of 56 mg EE/100 g in mango peel reported by Castro *et al*. (Castro Vargas et al., 2019), cocoa pod husk is mainly composed of catechin, quercetin, (-)-epicatechin, gallic acid, coumaric acid, and protocatechinic acid (Lu et al., 2018; Valadez Carmona et al., 2017). The polyphenols in CPH are bioactive compounds that represent a potential source for antioxidant-rich foods (Campos Vega et al., 2018).

2.3.7. Antioxidant Activity

The value of cocoa pod husk may be mainly due to their high content of phenolics and antioxidants, which could benefit human health by being used as an additive in foods or other value-added products (Jamaluddin et al., 2022). Evaluating antioxidant activity requires different test methods, as a single method can provide basic information, and combining several methods can describe the antioxidant properties in more detail. This is because each method measures the ability of a plant's antioxidants to scavenge specific radicals through different electron transfer pathways, either by inhibiting lipid peroxidation or by chelating metal ions (Cádiz Gurrea et al., 2014; Martínez et al., 2012). This study used two

antioxidant evaluation methods to assess the antioxidant capacity of Carmelo variety cocoa pod husk, ABTS and FRAP. ABTS was found to have a higher antioxidant capacity than FRAP, which is attributed to the low selectivity of the $ABTS^{+}$ radical in reacting with hydroxylated aromatic compounds in general. At the same time, FRAP is based on the ability of phenolic compounds to reduce Fe^{3+} to Fe^{2+} in the presence of 2,4,6-tripyridyl-s-triazine. The antioxidant capacity values obtained in this study were 304.84 µM ET/g and 145.80 µM ET/g for ABTS and FRAP, respectively. These results are close to the ABTS antioxidant capacity (227.7 - 147 µM ET/g) reported by Nieto *et al*. (Nieto Figueroa et al., 2020) and to the ABTS antioxidant capacity of grape skin (240 µM ET/g) reported by Siqueira *et al*., (Siqueira Melo et al., 2011) and higher than the ABTS antioxidant capacity of carrot skin beta-carotene (79.15 µM ET/g) reported by Jayesree *et al*., (Jayesree et al., 2021). These values are higher than the ABTS antioxidant capacity of cocoa pod husk (30.6 µM ET/g, 24.13 µM ET/g, and 42 - 24 µM ET/g) reported by Valadez *et al*. (Valadez Carmona et al., 2017), Martínez *et al*., (Martínez et al., 2012) and Lu *et al*., (Lu et al., 2018) respectively. Regarding the FRAP assay, a wide range of values was found in the literature for the antioxidant capacity of cacao pod husk, ranging from 521.73 µM ET/g to 1.98 µM ET/g (Jamaluddin et al., 2022; Lu et al., 2018; Martínez et al., 2012). There is a linear correlation between phenolic compounds and the antioxidant capacity of a plant (Cádiz Gurrea et al., 2014). According to Campos *et al*. (Campos Vega et al., 2018), the antioxidant activity of CPH has been studied in green chemistry as nanoparticles with larvicidal activity and against resistant bacteria, resulting in potentially viable products.

2.4. Conclusions

Lignocellulose content, antioxidant capacity, and mineral profile of CPH (*Theobroma cacao* L. Carmelo variety) were determined for the first time. With the data obtained in this study and those reported in the literature, it was found that the Carmelo variety does not present any peculiarities concerning other varieties. CPH is a viable source of lignin, and the cellulose and hemicellulose

contents have potential as a source of carbohydrates to produce biofuels; however, the high content of polyphenols and its antioxidant capacity characterize it as a potential source for the food and pharmaceutical industry. Although heavy metals were detected at trace levels in this study, the mineral profile should always be analyzed as the area of origin determines it. Contrary to cocoa beans, CPH has been little studied regarding antioxidant capacity. Therefore, the present study contributes to understanding a by-product with valorization potential.

Acknowledgments

The authors thank the Consejo Nacional de Humanidades, Ciencias y Tecnología (CONAHCYT) for the Postgraduate Grant no. 763652, the Dirección General de Investigación y Posgrado (DGIP) of the Universidad Autónoma Chapingo for funding projects: 20007-DTT-90, 22003 EIP, the Laboratorio de Productos Naturales del Departamento de Química of the Universidad Autónoma Chapingo, the Instituto Nacional de Investigaciones Forestales Agrícolas y Pecuarias (INIFAP), the Marine Geochemistry and LANNBIO laboratories of CINVESTAV-IPN, Unidad Mérida. HZP thanks C. Carlos Hernández Echeverría for providing cocoa pod husk of the Carmelo variety.

2.5. References

Alonso Báez, M., Avendaño Arrazate, C. H., Alonso López, B. L., Iracheta Donjuan, L., Martínez Bolaños, M., Ruíz Cruz, P., & Ortíz Curiel, S. (2020). Nutritional removal of cacao fruit (*Theobroma cacao* L.) in Mexico [Remoción nutrimental del fruto de cacao (*Theobroma cacao* L.) en México]. *Agro Productividad*, *13*(8), 53-59. https://doi.org/10.32854/agrop.vi.1713

Aregheore, E. M. (2002). Chemical Evaluation and Digestibility of Cocoa (*Theobroma cacao)* Byproducts Fed to Goats. *Tropical Animal Health and Production*, *34(4)*, 339-348.

Arzeta Ríos, A. J., Guerra Ramírez, D., Reyes Trejo, B., Ybarra Moncada, M. C., & Zuleta Prada, H. (2020). Microwave heating effect on total phenolics and antioxidant activity of green and mature coconut water. *International Journal of Food Engineering*, *16(12)*(20190378), 1-6. https://doi.org/10.1515/ijfe-2019-0378

Asiedu, N. Y., Neba, F. A., & Addo, A. (2019). Modeling the attainable regions for catalytic oxidation of renewable biomass to specialty chemicals: Waste biomass to carboxylic acids. *South African Journal of Chemical Engineering*, *30*, 1-14. https://doi.org/10.1016/j.sajce.2019.07.003

Barhoum, A., Jeevanandam, J., Rastogi, A., Samyn, P., Boluk, Y., Dufresne, A., Danquah, M. K., & Bechelany, M. (2020). Plant celluloses, hemicelluloses, lignins, and volatile oils synthesize nanoparticles and nanostructured materials. *Nanoscale*, *12*(45), 22845-22890. https://doi.org/10.1039/d0nr04795c

Barraza, F., Schreck, E., Uzu, G., Leveque, T., Zouiten, C., Boidot, M., & Maurice, L. (2021). Beyond cadmium accumulation: Distribution of other trace elements in soils and cacao beans in Ecuador. *Environ Res*, *192*, 110241. https://doi.org/10.1016/j.envres.2020.110241

Bernabé Santiago, R., Ávila Calderón, L. E. A., & Rutiaga Quiñones, J. G. (2013). Chemical components of the wood of five pine species of Morelia, Michoacán. *Madera y Bosques*, *19(2)*, 21-35.

Blankenhorn, P. R., Bowersox, T. W., Kuklewski, K. M., & Stirnely, G. L. (1985). EFFECTS OF ROTATION, SITE, AND CLONE ON THE CHEMICAL COMPOSITION OF POPULUS HYBRIDS. *Wood and Fiber Science*, *17(3)*, 351-360.

Cádiz Gurrea, M. L., Lozano Sanchez, J., Contreras Gámez, M., Legeai Mallet, L., Fernández Arroyo, S., & Segura Carretero, A. (2014). Isolation, comprehensive characterization, and antioxidant activities of *Theobroma cacao* extract. *Journal of Functional Foods*, *10*, 485-498. https://doi.org/10.1016/j.jff.2014.07.016

Campos Vega, R., Nieto Figueroa, K. H., & Oomah, B. D. (2018). Cocoa (*Theobroma cacao* L.) pod husk: Renewable source of bioactive

compounds. *Trends in Food Science & Technology*, *81*, 172-184. https://doi.org/10.1016/j.tifs.2018.09.022

Castro Vargas, H. I., Ballesteros Vivas, D., Ortega Barbosa, J., Morantes Medina, S. J., Aristizabal Gutierrez, F., & Parada-Alfonso, F. (2019). Bioactive Phenolic Compounds from the Agroindustrial Waste of Colombian Mango Cultivars 'Sugar Mango' and 'Tommy Atkins'-An Alternative for Their Use and Valorization. *Antioxidants (Basel)*, *8*(2). https://doi.org/10.3390/antiox8020041

Chávez Sifontes, M., & Domine, M. E. (2013). LIGNIN, STRUCTURE AND APPLICATIONS: DEPOLYMERIZATION METHODS FOR OBTAINING AROMATIC DERIVATIVES OF INDUSTRIAL INTEREST. *Avances en Ciencias e Ingeniería*, *4(4)*, 15-46. http://www.exeedu.com/publishing.cl/av_cienc_ing/

Daud, Z., Awang, H., Mohd Kassim, A. S., Mohd Hatta, M. Z., & Mohd Aripin, A. (2014). Cocoa Pod Husk and Corn Stalk: Alternative Paper Fibres Study on Chemical Characterization and Morphological Structures. *Advanced Materials Research*, *911*, 331-335. https://doi.org/10.4028/www.scientific.net/AMR.911.331

Devi, A., Niazi, A., Ramteke, M., & Upadhyayula, S. (2021). Techno-economic analysis of ethanol production from lignocellulosic biomass-a comparison of fermentation, thermo catalytic, and chemocatalytic technologies. *Bioprocess Biosyst Eng*, *44*(6), 1093-1107. https://doi.org/10.1007/s00449-020-02504-4

Doungous, O., Minyaka, E., Longue, E. A. M., & Nkengafac, N. J. (2018). Potentials of cocoa pod husk-based compost on Phytophthora pod rot disease suppression, soil fertility, and *Theobroma cacao* L. growth. *Environ Sci Pollut Res Int*, *25*(25), 25327-25335. https://doi.org/10.1007/s11356-018-2591-0

Espinosa García, J. A., Uresti Gil, J., Vélez Izquierdo, A., Moctezuma López, G., Inurreta Aguirre, H. D., & Góngora González, S. F. (2015). Productivity and profit potential of cocoa (*Theobroma cacao* L.) in the Mexican tropics. *Revista Mexicana de Ciencias Agrícolas*, *6-5*(1052), 1051-1063.

Cadmium dietary exposure in the European population, 37 § 2551 1 (2012). www.efsa.europa.eu/efsajourna

Gómez Hoyos, C., Mazo Márquez, P., Penagos Vélez, L., Serpa Guerra, A., Eceiza, A., Urbina, L., Velásquez Cock, J., Gañán Rojo, P., Vélez Acosta, L., & Zuluaga, R. (2020). Cocoa shell: an industrial by-product for the preparation of suspensions of holocellulose nanofibers and fat. *Cellulose*, *27*(18), 10873-10884. https://doi.org/10.1007/s10570-020-03222-6

Gramlich, A., Tandy, S., Andres, C., Chincheros Paniagua, J., Armengot, L., Schneider, M., & Schulin, R. (2016). Cadmium uptake by cocoa trees in agroforestry and monoculture systems under conventional and organic

management. *Science of the Total Environment*, *580*, 677-686. https://doi.org/10.1016/j.scitotenv.2016.12.014

Grob, L., Ott, E., Schnell, S., & Windhab, E. J. (2021). Characterization of endocarp powder derived from the cocoa pod. *Journal of Food Engineering*, *305*. https://doi.org/10.1016/j.jfoodeng.2021.110591

Hernández Mendoza, A. G., Saldaña Trinidad, S., Martínez Hernández, S., Pérez Sariñana, B. Y., & Láinez, M. (2021). Optimization of alkaline pretreatment and enzymatic hydrolysis of cocoa pod husk (*Theobroma cacao* L.) for ethanol production. *Biomass and Bioenergy*, *154*. https://doi.org/10.1016/j.biombioe.2021.106268

Hernández Rodríguez, G., Espinosa Solares, T., Hernández Eugenio, G., Villa García, M., Reyes Trejo, B., & Guerra Ramírez, D. (2016). Influence of Polar Solutions on the Extraction of Phenolic Compounds from Capulín Fruits (Prunus serotina). *J. Mex. Chem. Soc.*, *60(2)*, 73-78.

Hernández Rodríguez, G., Espinosa Solares, T., Pérez López, A., Salgado Escobar, I., & Guerra Ramírez, D. (2019). Antioxidant capacity of capulin (Prunus serotina subsp. capuli (Cav). McVaugh) fruit at different stages of ripening. *Ecosistemas y Recursos Agropecuarios*, *6(16)*(16), 35-44. https://doi.org/10.19136/era.a6n16.1947

Jamaluddin, M., Rizlan Ross, E. E., Azmi, A. F. M. N., Zubir, I., Nazir, N., Inderan, V., & Raseetha, S. (2022). Effect of Extraction Solvents on Phenolic Compounds of *Theobroma Cacao* L. By-products Using Ultrasound-Assisted Extraction. *International Journal on Advanced Science Engineering Information Technology*, *12(1)*.

Jayesree, N., Hang, P. K., Priyanka, A., Krishnamurthy, N. P., Ramanan, R. N., Turki, M. S. A., Charis, M. G., & Ooi, C. W. (2021). Valorization of carrot peel waste by water-induced hydrocolloidal complexation for extraction of carotene and pectin. *Chemosphere*, *272*, 129919. https://doi.org/10.1016/j.chemosphere.2021.129919

Jochem, D., Bösch, M., Weimar, H., & Dieter, M. (2021). National wood fiber balances for the pulp and paper sector: An approach to supplement international forest products statistics. *Forest Policy and Economics*, *131*. https://doi.org/10.1016/j.forpol.2021.102540

Lu, F., Rodriguez Garcia, J., Van Damme, I., Westwood, N. J., Shaw, L., Robinson, J. S., Warren, G., Chatzifragkou, A., McQueen Mason, S., Gomez, L., Faas, L., Balcombe, K., Srinivasan, C., Picchioni, F., Hadley, P., & Charalampopoulos, D. (2018). Valorization strategies for cocoa pod husk and its fractions. *Current Opinion in Green and Sustainable Chemistry*, *14*, 80-88. https://doi.org/10.1016/j.cogsc.2018.07.007

Martínez, R., Torres, P., Meneses, M. A., Figueroa, J. G., Pérez Álvarez, J. A., & Viuda Martos, M. (2012). Chemical, technological, and in vitro antioxidant properties of cocoa *(Theobroma cacao* L.) co-products. *Food Research International*, *49*(1), 39-45. https://doi.org/10.1016/j.foodres.2012.08.005

Mohamad, J. S., & Rushdan, I. (2014). Effect of bleaching on coir fibre pulp and paper properties (Kesan pelunturan terhadap sifat-sifat pulpa dan kertas sabut kelapa). *J. Trop. Agric. and Fd. Sc.*, *42(1)*, 51 – 61.

Moyin Jesu, E. I. (2007). Use of plant residues for improving soil fertility, pod nutrients, root growth, and pod weight of okra (Abelmoschus esculentum L). *Bioresource Technology* *98*(11), 2057-2064. https://doi.org/10.1016/j.biortech.2006.03.007

Nelson, N., Darkwa, J., Calautit, J., Worall, M., Mokaya, R., Adjei, E., Kemausuor, F., & Ahiekpor, J. (2021). Potential of Bioenergy in Rural Ghana. *Sustainability*, *13*(1). https://doi.org/10.3390/su13010381

Nieto Figueroa, K. H., Mendoza Garcia, N. V., Gaytan Martinez, M., Wall Medrano, A., Loarca Pina, M. G. F., & Campos Vega, R. (2020). Effect of drying methods on the gastrointestinal fate and bioactivity of phytochemicals from cocoa pod husk: In vitro and silico approaches. *Food Res Int*, *137*, 109725. https://doi.org/10.1016/j.foodres.2020.109725

Olubunmi Kayode, C., Olajiire Adeoye, G., Ezekiel Adewoyin, D. T., AyanfeOluwa, O. E., Ogundeji Ogunleti, D., & Fisayo Adekunle, A. (2018). Influence of Cocoa Pod Husk-Based Compost on Nutrient Uptake of Okra (Abelmoschus esculentus(L.) MOENCH) and Soil Properties on an Alfisol. *Communications in Soil Science and Plant Analysis*, *49*(17), 2113-2122. https://doi.org/10.1080/00103624.2018.1499108

Paz Cedeno, F. R., Solorzano Chavez, E. G., Dias, L. M., Otaviano, C. A., Bustamante, L. J. A., Monti, R., Martínez Galán, J. P., & Masarin, F. (2022). Composition and Chemical Structure of Hemicelluloses and Polysaccharides with Capability of Gel Formation. In *Hemicellulose Biorefinery: A Sustainable Solution for Value Addition to Bio-Based Products and Bioenergy* (pp. 111-137). https://doi.org/10.1007/978-981-16-3682-0_4

Porto de Souza Vandenberghe, L., Valladares Diestra, K., Bittencourt G., A., Murawski de Mello A., F., Sarmiento Vasquez, Z., Zwiercheczewski de Oliveira, P., Vinicius de Melo Pereira, G., & Ricardo Soccol, C. (2022). Added-value biomolecules' production from cocoa pod husks: A review. *Bioresour Technol*, *344*(Pt B), 126252. https://doi.org/10.1016/j.biortech.2021.126252

Redgwell, R., Trovato, V., Merinat, S., Curti, D., Hediger, S., & Manez, A. (2003). Dietary fiber in cocoa shell: characterization of component polysaccharides. *Food Chemistry*, *81*, 103-112.

Ricaño Rodríguez, J., Ramos Prado, J. M., Cocoletzi Vásquez, E., & Hipólito Romero, E. (2018). THE GENOMIC STUDY OF CACAO (*Theobroma cacao* L.): BRIEF RECOPILATION OF ITS CONCEPTUAL BASES. *Agro Productividad*, *11 (9)*, 29-35.

Rowell, R., Pettersen, R., & Tshabalala, M. (2012). Cell Wall Chemistry. In *Handbook of Wood Chemistry and Wood Composites, Second Edition* (pp. 33-72). https://doi.org/10.1201/b12487-5

Sánchez, M., Laca, A., Laca, A., & Díaz, M. (2023). Cocoa Bean Shell: A By-Product with High Potential for Nutritional and Biotechnological Applications. *Antioxidants (Basel)*, *12*(1028). https://doi.org/10.3390/antiox12051028

Sandesh, K., Shishir, R. K., & Vaman Rao, C. (2020). Optimization and comparison of induction heating and LPG-assisted acid pretreatment of cocoa pod for ABE fermentation. *Fuel*, *262*. https://doi.org/10.1016/j.fuel.2019.116499

Siqueira Melo, P., Bergamaschi, K. B., Tiveron, A. P., Prado Massarioli, A., Cadorin Oldoni, T. L., Celso Zanus, M., Matias de Alencar, S., & Pereira, G. E. (2011). Phenolic composition and antioxidant activity of agroindustrial residues. *Ciência Rural*, *41-6*, 1088-1093.

Titiloye, J. O., Abu Bakar, M. S., & Odetoye, T. E. (2013). Thermochemical characterization of agricultural wastes from West Africa. *Industrial Crops and Products*, *47*, 199-203. https://doi.org/10.1016/j.indcrop.2013.03.011

Valadez Carmona, L., Plazola Jacinto, C. P., Hernández Ortega, M., Hernández Navarro, M. D., Villarreal, F., Necoechea Mondragón, H., Ortiz Moreno, A., & Ceballos Reyes, G. (2017). Effects of microwaves, hot air, and freeze-drying on the phenolic compounds, antioxidant capacity, enzyme activity, and microstructure of cacao pod husks (*Theobroma cacao* L.). *Innovative Food Science & Emerging Technologies*, *41*, 378-386. https://doi.org/10.1016/j.ifset.2017.04.012

Vargas Arana, G., Merino Zegarra, C., Tang, M., Pertino, M. W., & Simirgiotis, M. J. (2022). UHPLC-MS Characterization and Antioxidant and Nutritional Analysis of Cocoa Waste Flours from the Peruvian Amazon. *Antioxidants (Basel)*, *11*(3). https://doi.org/10.3390/antiox11030595

Vasquez, Z. S., de Carvalho Neto, D. P., Pereira, G. V. M., Vandenberghe, L. P. S., de Oliveira, P. Z., Tiburcio, P. B., Rogez, H. L. G., Goes Neto, A., & Soccol, C. R. (2019). Biotechnological approaches for cocoa waste management: A review. *Waste Manag*, *90*, 72-83. https://doi.org/10.1016/j.wasman.2019.04.030

Vriesmann, L. C., de Mello Castanho Amboni, R. D., & de Oliveira Petkowicz, C. L. (2011). Cacao pod husks (*Theobroma cacao* L.): Composition and hot-water-soluble pectins. *Industrial Crops and Products*, *34*(1), 1173-1181. https://doi.org/10.1016/j.indcrop.2011.04.004

Yao, M., Bi, X., Wang, Z., Yu, P., Dufresne, A., & Jiang, C. (2022). Recent advances in lignin-based carbon materials and their applications: A review. *Int J Biol Macromol*, *223*(Pt A), 980-1014. https://doi.org/10.1016/j.ijbiomac.2022.11.070

Zheng, Y., Shi, J., Tu, M., & Cheng, Y.-S. (2017). Principles and Development of Lignocellulosic Biomass Pretreatment for Biofuels. In E. Inc. (Ed.), *Advances in Bioenergy* (Vol. 2, pp. 1-68). https://doi.org/10.1016/bs.aibe.2017.03.001

3. MIR-NIR AND NMR METABOLOMIC APPROACH TO ANALYZE THE RELEASE OF CARBOHYDRATES AND SMALL MOLECULES BY THE EFFECT OF MICROWAVE-ASSISTED MILD-HYDROTHERMAL PRETREATMENTS ON LIGNOCELLULOSE FROM *Theobroma cacao* L. POD HUSK

ABSTRACT

Microwave-assisted hydrothermal pretreatment (MA-HTP) of lignocellulosic material from cocoa pod husk (CPH) was evaluated by targeted and non-targeted metabolomics using near-infrared (MIR-NIR) and nuclear magnetic resonance (NMR) analytical methods. The objective was to evaluate five temperatures (100, 120, 150, 150, 180, and 200°C) and four particle sizes (<150 µm, 150 µm, 250 µm, and 355 µm) as discriminators of CPH depolymerization. The metabolomics MIR-NIR and NMR approach detected the dependence on temperature variables and the independence of particle size. At 200°C, MA-HTP monosaccharides and disaccharides can be obtained from the lignocellulosic matrix of CPH, as demonstrated by MIR-NIR and NMR discriminant analysis score plots, improving the discrimination of variables depending on the robustness of the statistical method. Targeted NMR analysis by rigorous matching between HSQC, TOCSY, and J-resolved entries concerning the Human Metabolome Database (HMDB) was able to identify metabolites derived from xylans, galactomannans, among others, which could evaluate different MA-HTP processes. An innovation of the work was to improve the reliability of the fit and prediction in the supervised models by applying the first derivative of the NIR-MIR matrix.

Keywords: Cocoa pod husk (CPH), Near-infrared (NIR), mid-infrared (MIR), nuclear magnetic resonance (NMR), Microwave-assisted hydrothermal pretreatment (MA-HTP), multivariate statistical analysis (MSA).

Doctoral Thesis: Doctorado en Ciencias Agroalimentarias, Universidad Autónoma Chapingo.
Author: Edna Elena Suárez Patlán, M. C.
Advisor: Holber Zuleta Prada, Dr.

3.1. Introduction

Cocoa production has served as an alternative for economic and social growth in countries that cultivate it, driven by a continual rise in global demand to meet the needs of the chocolate and its derivatives industry (Gómez Hoyos et al., 2020). Despite the numerous benefits of this industrial activity, cocoa production is linked to the substantial generation of lignocellulosic waste (Sarmiento Vásquez et al., 2021). Over the past years, an estimated 20 million tons of cocoa shell waste, known as *Theobroma cacao* L. Pod Husk or Cocoa Pod Husk (CPH), have been reported alongside 5.0 million tons of usable cacao grain annually (ICCO, 2024). These residues are typically left in field crops, where their natural degradation processes can lead to environmental and phytosanitary impacts, directly impacting crop yields (Nelson et al., 2021). CPH comprises a lignocellulosic residue composed mainly of cellulose (around 24.24 – 35%), hemicellulose (around 8-11%), lignin (around 14.6-26.38%), pectin (around 6.1-9.2%), ashes (around 6-10%) and proteins (around 4-10%) (Vasquez et al., 2019). The main physiological components of CPH include a mesocarp, epicarp, and endocarp. The mesocarp is composed of approximately 50% of cellulose. The epicarp is composed mainly of lignin. Finally, the endocarp mainly comprises hemicellulose and pectin (Lu et al., 2018). Molecularly, cellulose is composed of D-glucose unities, while hemicellulose is composed of xylan, arabinoxylan, arabinan, xyloglucan, galactomannan, and glucomannan (Priyangini et al., 2018). Lignin is mainly formed by hydroxyl-phenylpropane, p-coumaryl alcohol, sinapyl alcohol, and conipheryl alcohol units (Zheng et al., 2017). Pectin is associated with cellulose and hemicellulose and is mainly constituted by uronic acid moieties (Vriesmann et al., 2011). Principal elements identified in CPH ashes are Ca, Cu, Fe, K, Mg, Mn, Na, Se, and Zn (Lu et al., 2018; Vasquez et al., 2019). Due to their chemical above compositions, Cocoa Pod Husks (CPH), along with other lignocellulosic materials beyond hard and soft woods, constitute alternative raw materials that are being considered in the exploration of new frontiers for the valorization of lignocellulosic biomass. This entails investigating and analyzing the degradation processes of this waste byproduct of the cocoa industry to yield a

significant array of exploitable metabolites that can serve as raw materials for various chemical transformations and as substitutes for fossil fuels in green energy applications. This shift from polluting and finite fossil fuels will enhance the cocoa supply chain and mitigate its environmental and phytosanitary impacts (Brienzo, 2022; Muharja et al., 2021). The complex structural composition of polymeric materials from lignocellulosic biomass and CPHs represents a persistent limitation that poses various challenges to achieving effective depolymerization processes (Liu et al., 2017). For this reason, reliable and efficient pretreatments are currently required to solubilize lignin and hemicellulose, facilitate enzymatic hydrolysis, and produce exploitable small molecules (Roy, 2022b). Conventional pretreatment technologies for the depolymerization of lignocellulosic matrices include chemical, physical, and biological processes, as well as combinations of these, with a significant number of them developed in the field of green chemistry, such as treatments with organosolv, ionic liquids, deep eutectic solvents, and some hydrothermal pretreatments (Ruengsrichaiya et al., 2024; Swami et al., 2024; Duarte et al., 2024; Chakraborty et al., 2024). In particular, microwave-assisted hydrothermal pretreatments (MA-HTP) are considered one of the most attractive and competent technologies, with economic and environmental advantages (Mellinas et al., 2020; Muharja et al., 2021). These pretreatments manage to solubilize a large part of the hemicellulose and lignin components through uniform heating ramps, reaching temperatures between 160 and 200 °C in relatively short treatment times and with low processing costs (López Linares et al., 2019; Puke et al., 2024). Recently, there was a report on microwave-assisted hydrothermal extraction (MA-HTE) of tobacco biomass, capable of directly extracting non-structural carbohydrates (NSC) and hemicellulose (Yuan et al., 2019). On the other hand, hemicellulose can be effectively depolymerized into soluble monomers, oligosaccharides, and low molecular weight polymers through hydrothermal processes, as indicated by high temperatures. Pressurized hydrothermal treatments promote more effective water dissociation into H_3O^+ and OH^- ions due to the expected changes in Kd from 10^-14 (25°C) to 10^-11 (220°C), which will

facilitate the solubilization of both hemicellulose and a small fraction of lignin (Sun et al., 2022). Vriesmann et al. (Vriesmann et al., 2011) carried out a CPH-HTE at two different mild temperatures (50 and 100°C) to obtain fractions of 7% and 12% of pectin, respectively. Fourier Transform Infrared (FT-IR) and 13C nuclear magnetic resonance spectroscopy (NMR 13C) determined that uronic acid was the main component of the extracted pectin.

High-resolution analytical methods combined with chemometric models and multivariate statistical analysis are emerging tools for analyzing complex data matrices to obtain holistic undirected fingerprints to classify discriminant factors and identify a set of relevant metabolites related to the controlled variables in metabolomic research methodologies (García Pérez et al., 2024; Nepper Davidsen et al., 2024; Herbert Pucheta et al., 2021). Combined metabolomic analysis methods have been applied to studying *Theobroma cacao* L cacao and its by-products (Balcáza Zumaeta et al., 2023). Recent studies have proposed the use of quadrupole time-of-flight mass spectrometry coupled with the separation method (HPLC-QTOF-MS) to elucidate discriminating cocoa testa metabolites related to geographical origins, harvest years, and varieties (Cain et al., 2019). Similarly, with equivalent instrumentation, matrices of metabolomic data related to fruit ripening have been obtained as a strategy for cocoa harvesting (Gallego et al., 2022). In terms of orthogonal non-hyphenated combined analytical methods such as the use of Fourier-Transform infrared and nuclear magnetic resonance spectroscopy have been used for characterizing hot-water-soluble pectins in *Theobroma cacao* L. (Vriesmann et al., 2011). Recognizing the importance of the leading analytical platforms for metabolomic analysis: nuclear magnetic resonance (NMR) and high-resolution mass spectrometry (HRMS); as far as we know, combined near and mid-infrared spectroscopy with nuclear magnetic resonance spectroscopy is used as analytical methods to obtain, respectively, untargeted, and targeted/untargeted metabolomic data matrices to evaluate any discriminant factor related to the food product *Theobroma cacao L* CPH. has not been reported so far. The present work uses an approach of untargeted orthogonal near-infrared (NIR) and mid-infrared (MIR) metabolomics in

combination with a set of mono and multidimensional liquid-state nuclear magnetic resonance (NMR) schemes for a targeted metabolomics model - from now on referred to as the MIR-NIR and NMR model - capable of identifying features used to discriminate between five different MA-HTP conditions (100°C, 120°C, 150°C, 180°C, and 200°C) of samples sieved into four different particle sizes (355 µm, 250 µm, 150 µm, and less than 150 µm). The MIR-NIR and NMR data matrices also include a set of relevant standards that allow the creation of multivariate statistical analysis models to evaluate the efficiency of MA-HTP processes in obtaining small molecules and low molecular weight carbohydrates.

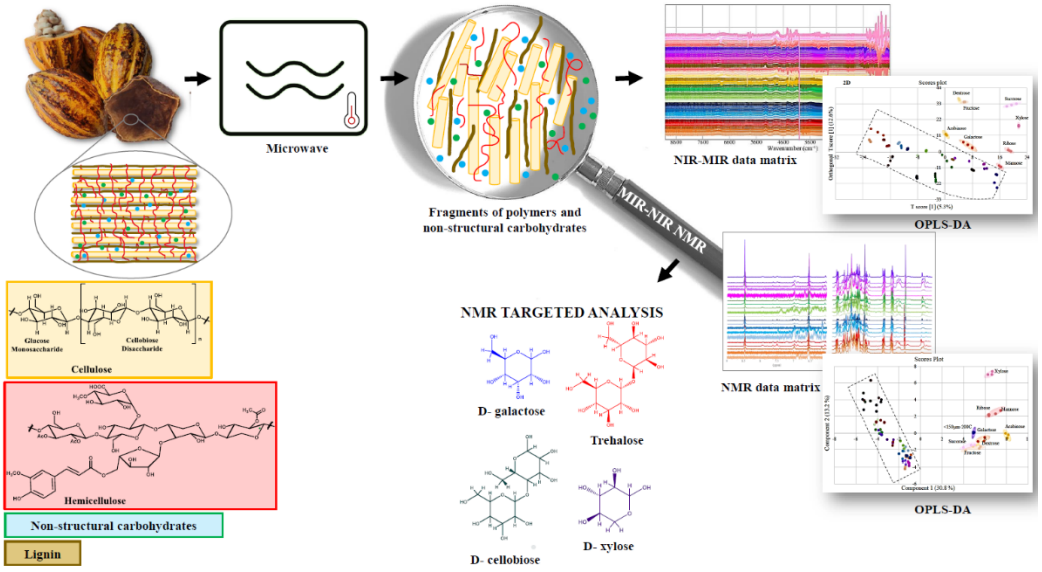


Figure 11 MIR-NIR and NMR metabolomic approach to microwave-assisted hydrothermal pretreatment of *Theobroma cacao* L pod husks.

3.2. Materials and methods

3.2.1. Materials and reagents

The raw material used for the present study comprises the exocarp, mesocarp, and endocarp of *Theobroma cacao* L., “Carmelo” variety. Cacao was collected at Rancheria Rio Seco, at Cunduacan, Tabasco County in Mexico [18°04'00"N y 93°10'00"W]. The crop field is 10-13 meters above sea level (MASL). Four particle sizes were obtained after sieving: 355 µm, 250 µm, 150 µm and below 150 µm, using mechanical crushing with a Thomas Wiley model 4 laboratory mill

(TP4274E70520A, Thomas Scientific, Swedesboro, New Jersey, U.S.A.) and further classified with a U.S. STD 45, 60, 80 and 100 laboratory sieves (W. S. Tyler, Ohio, U.S.A.), further dried in an oven at 102°C ± 3 °C and finally stored in a desiccator until their use.

Carbohydrate standards used for the metabolomics model were: D-Ribose Ribosa ≥ 99 % (CAS No. 50-69-1), L-Arabinose ≥ 99 % (CAS No. 5328-37-0), D-Mannose ≥ 99 % (CAS No. 3458-28-4), D-Glucose ≥ 99.5 % (CAS No. 50-99-7), D-Xylose ≥ 99 % (CAS No. 58-86-6, D-Fructose ≥ 99 % (CAS No. 57-48-7), Sucrose ≥ 99.5 % (CAS No. 57-50-1), all from Sigma Aldrich (Merck KGaA, Saint Louis Missouri, U.S.A.) and Galactose (CAS No. 59-23-4) from SUPELCO (Merck KGaA, Saint Louis Missouri, U.S.A.). Cellulose and lignin standards used for the metabolomics model were obtained using the reference method ASTM D 103-77. Two grams of dried holocellulose base, previously extracted from a CPH 60-sieve filtration, was hydrolyzed with NaOH base at 17.5% v/v during 45 minutes at 20°C ± 3°C and constant stirring, and a final stage of 1 hour without mechanic agitation. Afterward, reaction bulk was filtered within a Gooch Porcelain Crucible (ASTM 40-60 C) for accurate basic washing with 100 ml of NaOH at 8.3% v/v, followed by a distilled water washing accompanied with a final acidic washing with 15 ml of diluted acetic acid / distilled water, until arriving to a neutral pH. Once neutralized, the obtained solid was dried in an oven at around 105ºC heating until it obtained constant weight. Lignin standard was obtained in similar conditions, following TAPPI T 204 cm 97 y T 222 om-02.

3.3. Methods

3.3.1. Pod Husk *Theobroma cacao* CPH microwave-assisted hydrothermal pretreatments (MA-HTP)

An experimental batch is composed of 100 milligrams of CPH, each sieved at one of the four possible different particle sizes' filters: <150 µm, 150 µm, 250 µm y, 355 µm. Each batch was deposited in a 10 ml volume CEM Corporation glass reactor for microwave synthesis (Discover System TM 908005, Rhode Island,

U.S.A.). Each glass reactor was added with 0.5 ml of bi-distilled water and a magnetic stir bar. The glass reactor was sealed with a rubber cap and inserted into the CEM Corporation microwave synthesizer (Discover System TM 908005, Rhode Island, U.S.A.), installing a specific temperature according to the herein presented protocol: 100, 120, 150, 180 y 200°C. Microwave residence time was 5 minutes, with a pressure of 300 psi. In turn, the experimental matrix is composed of 20 samples labeled as follows:

<150 µm-100°C, <150 µm-120°C, <150 µm-150°C, <150 µm-180°C and <150 µm-200°C

150 µm-100°C, 150 µm-120°C, 150 µm-150°C, 150 µm-180°C and 150 µm-200°C

250 µm-100°C, 250 µm-120°C, 250 µm-150°C, 250 µm-180°C and 250 µm-200°C

350 µm-100°C, 350 µm-120°C, 350 µm-150°C, 350 µm-180°C and 350 µm200°C

They observed that the experimental unit comprises five different temperatures and four different sieving filters. Each point of the experimental unit was triplicated for statistical purposes. After 5 minutes of each MA-HTP, a two-phase heterogeneous mixture (solid and aqueous) was obtained within each microwave glass synthesizer. The solid phase corresponds to the lignocellulosic fragment not modified by the treatment. In contrast, the liquid phase comprises the hydrothermal extraction product. The supernatant is separated from the solid phase and versed in a 15 ml Eppendorf tube, adding 5 ml more bi-distilled H_2O. Said liquid phase is centrifuged at 4000 rpm in a Frontier™ centrifuge, model 5000 (FC5718R, OHAUS Corporation, Parsippany-Troy Hills, New Jersey, U.S.A.) and afterward filtered with a 50 ml Gooch Porcelain Crucible (ASTM 40-60 C). Residual precipitates within the Gooch Porcelain Crucible were washed with 5 ml of bi-distilled water. Each collected aqueous aliquot was lyophilized with a temperature of -49°C and a vacuum of 0.040 mbar until a solid phase that was analyzed with near-infrared (NIR), mid-infrared (MIR) and Nuclear Magnetic Resonance Spectroscopy (NMR).

3.3.2. Near-infrared (NIR) spectroscopy acquisition details

Lyophilized MA-HTP products and solid-state standards were each placed in a glass vial with a constant optical path for maximizing its surface contact. NIR absorbance spans were carried out with a Multipurpose analyzer Bruker Optics spectrophotometer (Rosenheim, Germany) scanning a wavelength between 830 and 2500 nm, with a 2.5 nm span resolution (wavenumbers between 12000 to 4000 cm^{-1} and a 4 cm^{-1} span resolution). Acquisition routines were performed for all samples in absorbance mode with 128 scans for blanks and sample collections. All acquisitions and data export to JCAMP-DX universal format were done using the OPUS 7.8 program.

3.3.3. Mid-infrared (MIR) spectroscopy acquisition details

Lyophilized MA-HTP products and solid-state standards were each placed in the transmission acquisition module of a Cary 630 Fourier-transform mid-infrared (FTIR)-attenuated total reflection, ATR (Agilent Technologies Inc., Santa Clara, California, U.S.A.). All scans were carried out within a wavelength between 2500 and 15384 nm, with a 2.5 nm span resolution (wavenumbers between 4000 to 650 cm^{-1} and a 4 cm^{-1} span resolution). Acquisition routines were performed for all samples in absorbance mode with 128 scans for blanks and sample collections. All acquisitions and data export to JCAMP-DX universal format were done using the MicroLab FTIR software.

3.3.4. Nuclear magnetic resonance (NMR) spectroscopy acquisition details

NMR spectroscopy was carried out with a 400/54 Premium Shielded spectrometer (Agilent Technologies Inc. California, U.S.A.), operating at 400 MHz proton frequency / 100 MHz carbon-13 frequency.

In all cases, 30 mg of solid-state standards or lyophilized MA-HTP products were dissolved in a 600 µl solution containing 10% of deuterated water (99.9% deuteration, CAS No. 7789-20-0) and 90% of an aqueous solution containing 0.05

M of 3-(trimethylsilyl)-2,2,3,3-d4-propionic acid that serves as internal standard (TSP: 98% of deuteration, CAS No. 24493-21-8) and a phosphate buffer solution at a concentration of 0.5 mM (pH= 7.4, CAS No. 7778-77-0). One dimensional proton NMR acquisition strategy was carried out with a direct polarization scheme (1D-^{1}H-DP) that calibrates the noesy presaturation scheme (1D-^{1}H-water suppression) needed to develop a water suppression[10]. Each 1D-^{1}H-water suppression experiment was done with 16 transients comprising a time domain of 16384 complex points, a spectral width of 16 ppm (6410 Hz), a relaxation recovery delay of 10 seconds, an acquisition time of 2.5 seconds, and a receiver gain of eight, producing in turn experimental times of 10 minutes per experiment. The 1D-^{1}H-water suppression spectra was the basis for non-targeted metabolomics, whilst targeted schemes were done with two-dimensional NMR homonuclear ^{1}H-^{1}H long-range correlation TOCSY excitation sculpting water suppression (spectral width of 16 ppm in both dimensions, acquisition time in the direct dimension of 15 ms, time domain in the indirect dimension of 1024 points 16 transients per increment, relaxation recovery delay of 1 second and an optimized receiver gain of 46, produced experimental times of 12 hours per experimental batch), proton J-resolved (HOMO2DJ pulse sequence comprising 16 ppm of spectral width in the F2 dimension, 256 complex points acquired in the indirect J-coupling dimension, a relaxation recovery delay of 1.5 seconds and an optimized receiver gain of 32 that produced experimental times of 5.17 hours per experimental batch) and heteronuclear correlation ^{1}H-^{13}C HSQC (^{1}H spectral width of 16 ppm, ^{13}C spectral width of 200 ppm, relaxation recovery delay of 1 second, 1024 complex points acquired for the indirect dimension, acquisition times for the direct dimension of 15 ms, and optimized receiver gain of 101 produced experimental times of 5 hours).

3.3.5. NMR experiments.

3.3.5.1. Spectroscopic preprocessing

3.3.5.2. MIR-NIR data matrix preprocessing for non-targeted metabolomics.

After both MIR and NIR data matrixes were separately JCAMP converted, all data were further transformed to a comma-separated value format (CSV) using the JCAMP-DX to CSV file converter script found at the following link: http://www.cheminfo.org/Chemistry/Cheminformatics/Convert_a_logP_to_a_color/i ndex.html. Scripts' outputs were extracted and pasted into Excel spreadsheets in a CSV format. The first derivative NIR Absorbance data matrix was obtained from Raw data with Microsoft Excel routines. The addition of the MIR-NIR data was carried out by implementing standard normalization factors. Individual CSV files were arranged in a two-variable input format (mz= frequency / into= normalized absorbance) suitable for MetaboAnalyst 5.0 software. Each MIR-NIR raw data and first derivative triplicate were saved in folders defining the discriminant factor. All variable sets were, in turn, appropriately arranged. ZIP format as standard Metaboanalyst 5.0 Statistical analysis inputs.

3.3.5.3. 1D-NMR data matrix preprocessing for non-targeted metabolomics.

NMR postprocessing for producing the MSA input variables was carried out: ppm calibration and manual phase corrections were conducted with Bruker TopSpin 4.1.1 software. Global and intermediate baseline corrections, least square or parametric time warping NMR alignments, variable size bucketing for untargeted profiling, and data matrix normalization were carried out with the NMRProcFlow software.

3.3.5.4. 2D-NMR data matrix preprocessing for targeted metabolomics.

2D- HSQC, TOCSY, and J-Resolved experiments were transformed into a Bruker format for processing routines with the Topspin 4.1.1 software. All data were Fourier transformed with at least 2048 points of zero filling for both dimensions in homo- and heteronuclear correlation spectra and a zero filling of 8192 points in the direct dimension for J-resolved spectra. Homo and heteronuclear correlation spectra assignment were used as constraints for metabolites' identification, with the use of the Human Metabolome Database 2D NMR spectra search routine.

(https://hmdb.ca/spectra/nmr_two_d/search), with a (x,y) tolerance of ± 0.03 ppm. Positive hits were considered only for the HSQC and TOCSY double match. All HMDB metabolites' candidates were confirmed or discarded by identifying a 1H chemical shift with a scalar coupling constant obtained with 2D-J Resolved spectroscopy. Comparing the experimental value with an NMRDB platform simulation confirmed these as positive hits.

(http://www.nmrdb.org/new_predictor/index.shtml?v=v2.138.0), They were obtained from the PubChem database using the sdf molecular structures' candidates as inputs.

(https://pubchem.ncbi.nlm.nih.gov/). Table 8 shows all identified metabolites with the triple match TOCSY-HSQC-J resolved, which can be resumed as follows: A first ^{1}H chemical shift pair is assigned from the homonuclear TOCSY assignment. Then, one identified proton spin presents heteronuclear connectivity with a ^{13}C spin system assigned with HSQC. An $^{n}J_{H-H}$ coupling constant between the identified proton and its ^{1}H spin pair priorly assigned by TOCSY is confirmed. Finally, a $^{1}H^{13}C$ HSQC correlation between the second ^{1}H spin and its ^{13}C one-bond distance scalar counterpart is assigned. The two heteronuclear ^{1}H-^{13}C pairs with their corresponding chemical shifts, with their proton scalar coupling value, are highlighted and reported in Table 8.

3.3.6. Multivariate statistical analysis (MSA)

Multivariate statistical analysis of MIR-NIR and NMR data matrixes independently include data matrix normalization by sum (for adjusting differences amongst samples), Log transformation, and autoscaling (mean centering divided by the standard deviation of each variable), which are applied to remove any possible variation during the experimental phase, to make features as comparable between them as possible. Normalized and scaled NIR-MIR and NMR data is then submitted to the statistical analysis workflow for obtaining unsupervised Principal Component (PCA, 13A, and 13E comprise MIR-NIR score plots; Figures 14A and 14E comprise NMR score plots), standard (PLS-DA Figures 13B and 13F comprise MIR-NIR score plots; Figures 14B and 14F comprise NMR score plots) and sparse partial least-square discriminant analysis (sPLS-DA, Figures 13C, and 13G comprise MIR-NIR score plots; Figures 14C and 14G comprise NMR score plots), as well as the Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA, Figures 13D and 13H comprise MIR-NIR score Figures plots; Figures 14D and 14H comprise NMR score plots), from the constant sum normalized MIR-NIR (Figure 13) and NMR (Figure 14) data matrixes, were developed with the MetaboAnalyst 5.0. In all cases, T2 Hotelling's regions depicted by ellipses in score plots of each model define a 95% confidence interval. The reliability of each classification in supervised models was evaluated regarding the goodness of the fit (R^2) and the goodness of prediction (Q^2).

3.4. Results and discussion

3.4.1. Spectral characteristics MIR-NIR data matrix of CPH MA-HTP sampling

Figure 12 presents MIR-NIR first derivative data matrix convolution comprising a spectral wavelength range between 1163 and 16666 nm (8600-600 cm^{-1}), showing the data from 20 MA-HT pretreatments, cellulose, lignin, as well as monomeric and dimeric carbohydrates. A NIR first derivative absorbance band of 6900 cm^{-1}, related to an O-H stretching of mono- or dimeric arabinose, fructose,

mannose, ribose, and sucrose, is observed at 5150 cm^{-1}, for MA-HTP sampling, as well as in cellulose and lignin. In contrast, the later NIR first derivative absorbance band results from a dual hydroxyl stretching and C-O distorted bending exclusively manifested in complex carbohydrates (Amirvaresi et al., 2021). A complex NIR first derivative carbohydrate pattern between 4000 – 5000 cm^{-1} is observed for the complete set of samples, mostly related to a different set of C-H and O-H vibration modes (Lin & Dence, 1992). Set region results are highly discriminant for differentiating complex and simple carbohydrates (vide infra).

Within the mid-infrared region, a set of discriminant first derivative spans are observed: i) the band around 3260 cm^{-1} wavenumber related to the typical O-H stretching, has a significant presence in all monomeric and dimeric carbohydrate standards, cellulose and high-temperature MA-HT pretreatments (180 and 200°C), with a leaser contribution for MA-HT pretreatments carried out at lower temperatures (100, 120 and 150°C); ii) a $C_{(sp3)}$-H stretching at 2920 cm^{-1}, exclusively observed for monomeric standards and sucrose and iii) a general NIR first derivative span pattern based on frequencies at 1730 cm^{-1}(C=O stretching), 1560 cm^{-1} (C=C stretching), 1440 cm^{-1} (C-H bending), 1380 cm^{-1} (O-H bending), 1240 cm^{-1} (C-O stretching), and 1040 cm^{-1} (C-O-C stretching) present in all CPH pretreatments (Adjin-Tetteh et al., 2018; Brienzo, 2022; Dahunsi et al., 2019; Gatot, 2012). Raw and first derivative absorbance MIR frequencies comprise the contribution of diverse normal vibration modes. Consequently, they do not significantly contribute to providing discriminative signals, as elsewhere reported (Lin & Dence, 1992).

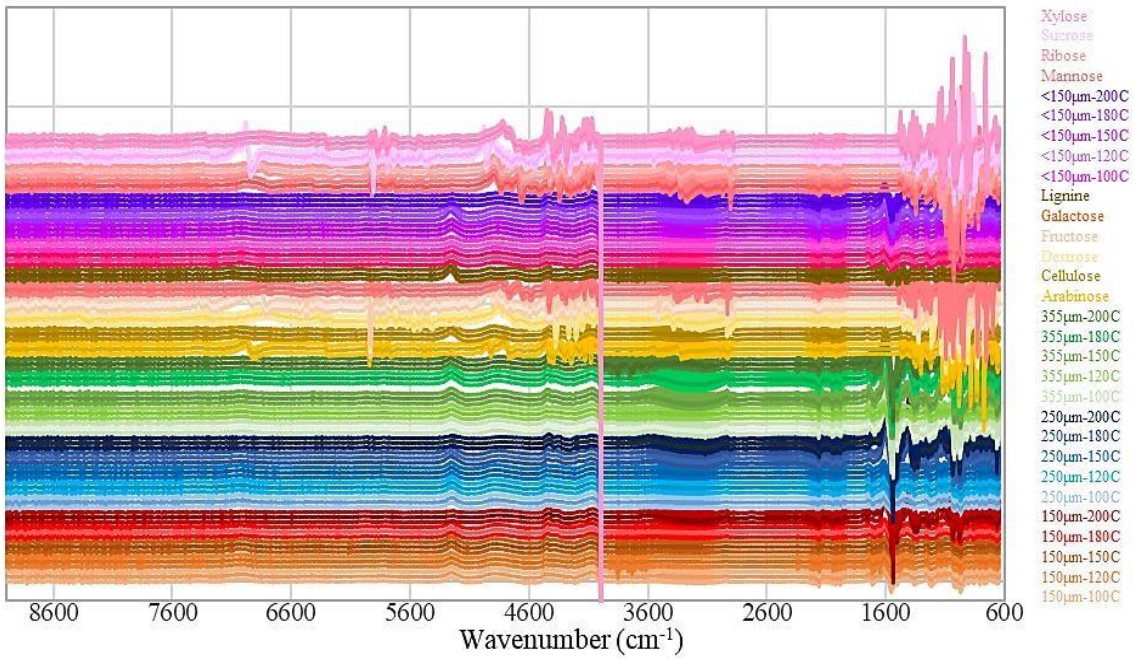


Figure 12 Stacked NIR-MIR spectra of CPH pretreatments, pentose, hexose, sucrose standards, and cellulose and lignin polymers.

3.4.2. MIR-NIR multivariate statistical analysis

Application of artificial neuronal networking / deep learning ML algorithms have been developed for data pre-processing, peak identification, peak integration, compound identification/quantification, data analysis, and data integration of complex spectroscopic and spectrometric data matrixes. The work uses equivalent multivariate statistical analysis algorithms to evaluate two data matrixes of the exact CPH MA-HTP / CHO standards sampling: MIR-NIR (Figure 13) and NMR (Figure 14).

Generally, all non-supervised and supervised MSA approaches define a separate subspace between MA-HTP and polymeric cellulose -lignin matrixes (PC1, T score [1] < 0) and monomeric and dimeric CHO moieties (PC1, T score [1] >0). As observed, most discriminative supervised artificial neuronal networking algorithms show that effective MA-HTP processes that might degrade to simpler CHO moieties will trend to shift to a PC1 (T score [1]) > 0 subspace, strongly suggesting that said observable might be an indicative of effective pretreatments. Multivariate statistical analysis of MIR-NIR data matrix starts with a non-

supervised Principal occupy a Component Analysis (PCA, figure 13A), that in turn explains data variance by first estimating and reducing data dimensionality, with a further calculation of a new coordinate system or principal components, without a class labeling (Gómez Hoyos et al., 2020). Figure 13A shows the unsupervised PCA score plot of the MIR-NIR data matrix. The first and second principal components explain 61.1% of the total variance by contributing to PC1= 53.7% and PC2=7.4%. The MIR-NIR PCA multivariate statistical analysis does not produce significative discriminations amongst MA-HT pretreatments and cellulose-lignin polymers, in turn, delimited within the dotted box, expanded in Figure 13E.

The standards of pentoses, hexoses, and disaccharide sucrose are separated and identified in Hotelling ellipses (which define a confidence level of 95%). Supervised methods such as partial least squares discriminant analysis (PLS-DA) were applied to increase the discriminant capacity of the NIR-MIR matrix under study. This discriminant method applies PLS regressions from which possible class members can be predicted, and the discriminations are evaluated with permutation tests and cross-validations (Alamar et al., 2020). Figure 13B shows the PLS-DA showing a total variance of 60.2 %, with 53 % for CP1 and 7.2 % for CP2. These analyses allowed greater separation within the group of pretreatments and polymers. However, there is no significant discrimination in this group. The standards for pentoses, hexoses, and sucrose remain bounded by Hotelling ellipses (at a 95% confidence level) except for arabinose, which appears not to discriminate effectively from the pretreatment group. Sparse partial least squares discriminant analysis (sPLS-DA) is a variant of PLS-DA that performs variable selection and classification during model calibration and discards uninformative variables. The sPLS algorithm efficiently estimates many variables (Jimenez Carvelo et al., 2021). Figure 13C shows the sPLS-DA discriminant analysis with an overall variance of 21.3 %, considering 12.5 % for PC1 and 8.8 % for PC2. This analysis shows a remarkable increase in the separation of pretreatments and polymers within the group, delimiting each pretreatment and polymer by species. Likewise, a closer approach of the galactose, arabinose,

ribose, xylose, and mannose standards to the group of pretreatments and polymers is observed. Finally, the discriminant analysis to orthogonal projections to latent structures (OPLS-DA) is a supervised model that provides dimension reduction and identification of spectral features due to its ability to distinguish between relevant and irrelevant data variations within a group, which also allows the prediction of the labels of these groups. This model leads to maximum group separation among the spectroscopic data set by applying Monte-Carlo cross-validations and results in less complex discriminant models with more accurate and reliable reductions than PLS-DA (Herbert Pucheta et al., 2021). Figure 13D shows the OPLS-DA dot plot; this analysis provided even broader separations within the group of cellulose and lignin pretreatments and polymers and defined them more effectively than sPLS-DA; this can be seen in more detail in Figure 13H, which shows a close-up of the group of pretreatments and polymers, here it can be seen that the pretreatments similar to each other are 355µm-200°C and 355µm-180°C, as well as the 250µm-200°C and 250µm-180°C pretreatments, which are together with the 150µm-200°C pretreatment in a quadrant opposite to the quadrant of the pentoses, hexoses, and sucrose standards. This could indicate that high temperatures influence the hydrothermal pretreatment of CPH. The permutation test between one predicted component (p1) and three orthogonal components (o1, o2, and o3) yielded coefficients R2X = 0.3095, R2Y = 0.9894, and Q2 = 0.9698 (Appx. 1).

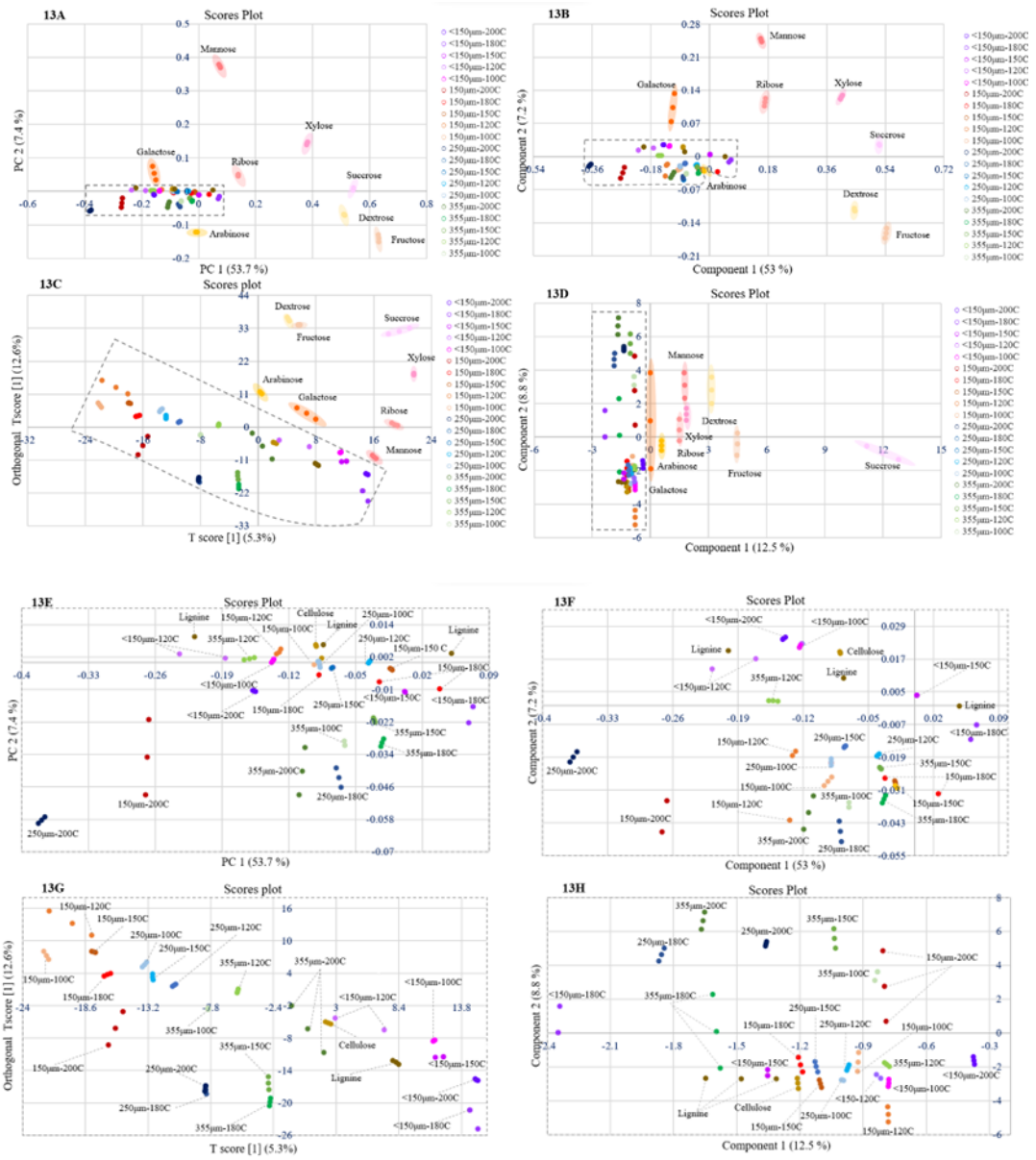


Figure 13 Supervised and unsupervised multivariate analysis based on NIR-MIR spectroscopy of hydrothermal pretreatments of CPH, arabinose, dextrose, fructose, galactose, mannose, ribose, sucrose and xylose standards, and lignin and cellulose polymers;13A) principal component analysis (PCA), 13B) partial least squares discriminant analysis (PLS-DA), 13C) sparse partial least discriminant analysis (sPLS-DA) and 13D) orthogonal projections to latent structure discriminant analysis (OPLS-DA).

3.4.3. Multivariate statistical analysis of ^{1}H NMR

Figure 14A shows the dot plot of the unsupervised PCA, which classifies the ^{1}H NMR data matrix and determines the correlation between factors. The most significant principal components were PC1, with 33.4%, and PC2, with 15.6%, for a total variance of 49%. This model allowed only limited separation in almost all pretreatments (grouped in a dashed rectangle), except the <150µm-200°C pretreatment (purple Hotelling ellipse at 95% confidence level), which was completely separated from the pretreatment group and can be seen in the quadrant of the pentoses, hexoses and sucrose standards (Hotelling ellipses at 95% confidence level), which are in quadrant one except xylose and dextrose. Figure 14B presents the PLS-DA dot plot with a total variance of 44% of the initial variance, with PC1 contributing 30.8% and PC2 13.2%. In this model, it was possible to observe more excellent discrimination between the pentoses, hexoses, and sucrose standards of the hydrothermal pretreatment group, being reiterative regarding the discrimination of the <150µm-200°C pretreatment, which is differentiated by the expression of the galactose, dextrose, fructose, and sucrose standards. In this case, all standards are in the same quadrant except for fructose. On the other hand, figure 14C shows the sPLS-DA dot plot with 43.8% overall variance composed of two principal components, PC1 contributing 29.7% and PC2 contributing 14.1% of the initial variability. This model reiterates the discrimination of the <150µm-200°C pretreatment (purple Hotelling ellipse at 95% confidence level) from the rest of the pretreatments, which in this model were compactly grouped in a dashed box where an almost generalized trend of separation between the 100, 120 and 150°C pretreatments and the 180 and 200°C pretreatments within the same compact group of pretreatments is observed (Figure 14G). This model's pentose and hexose standards were in the same quadrant as the <150µm-200°C discriminated pretreatment except for dextrose, xylose, and sucrose. Finally, figure 15D shows the OPLS-DA model; this model is reiterative in the discrimination of the hydrothermal pretreatment <150µm-200°C, which remains in the quadrant of the pentoses, hexoses, and sucrose standards except for xylose. Regarding the set of pretreatments, they are

separated from each other, preserving the grouping of species within the same group. The permutation test with p1 and o1, o2, o3, and o4 was analyzed with a supervised orthogonal projection to the latent structure discriminant model: R2X = 0.7033, R2Y = 0.9393 and Q2 = 0. 9154.

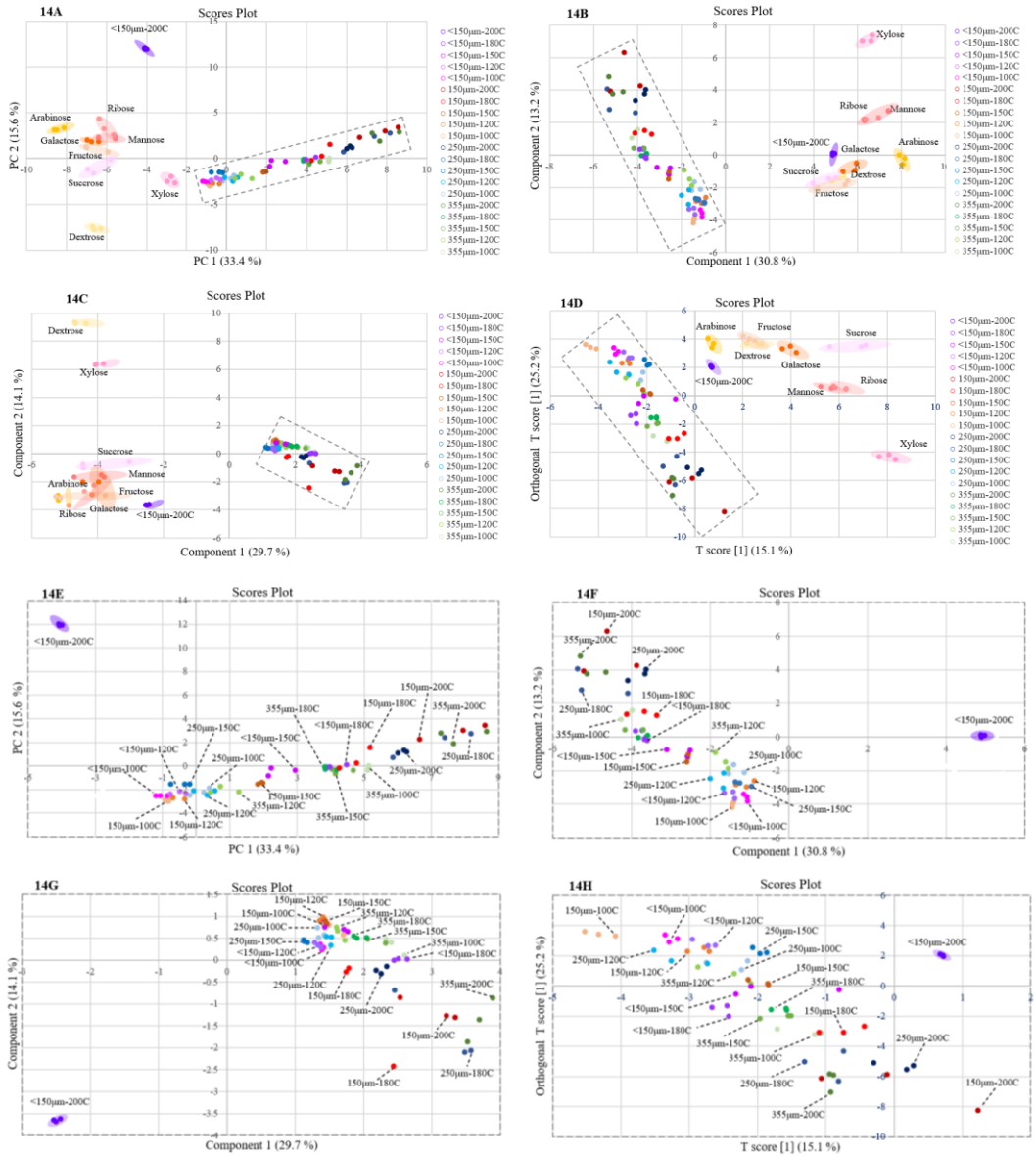


Figure 14 Multivariate analysis of NMR spectroscopy-based metabolomics data of cocoa pod husk samples after applying hydrothermal pretreatments; 14A) score plots of the unsupervised principal component analysis (PCA), 14B) supervised standard partial least squares (PLS-DA), 14C) scattered partial least squares (sPLS-DA) and 14D) orthogonal projections to latent structures discriminant analysis (OPLS-DA).

3.4.4. One-and two-dimensional NMR spectroscopic analyses

The preliminary inspection of stacked ^{1}H water-present one-dimensional NMR spectra of cocoa pod husk samples after applying hydrothermal pretreatments, considering the non-determining anomeric region for this analysis in the area from 4.50 to 3.00 ppm (Figure 15), which is a complex matrix dominated by signal overlap originated by the integrated multiplets of sugars. Here, the presence of simple monosaccharides D-fructose, D-galactose, D-xylose, and the disaccharide D-maltose is evidenced, Supporting Information (Figures Si-35 and Si-36.).

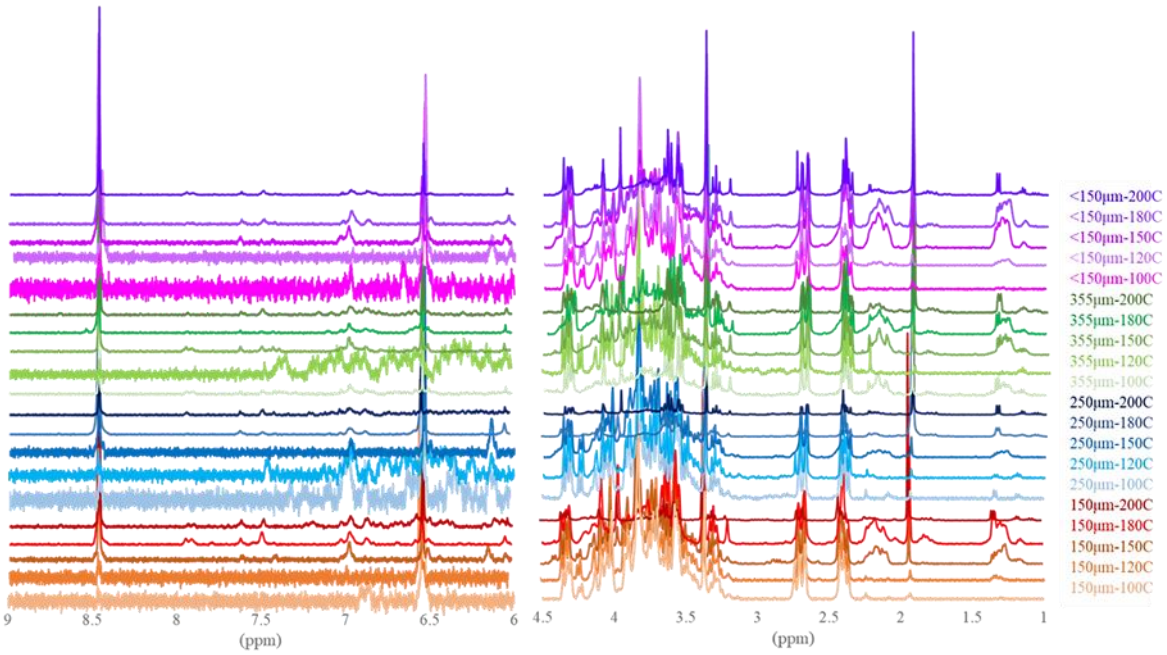


Figure 15 Stacked one-dimensional ^{1}H water_presat NMR spectra of cocoa pod husk samples after applying hydrothermal pretreatments.

3.4.5. Two-dimensional NMR spectroscopic analyses

By processing and mapping the two-dimensional NMR spectra, HSQC, TOCSY, and HOMO2DJ, 14 metabolites could be identified in the hydrothermal pretreatments of CPH (Allose, Cellobiose, Cytidine, D-fructose, D-galactose, D-maltose, D-xylose, Fructose 6-phosphate, Gluconolactone, Gluconic acid, Glycerol, Glycerol 3-phosphate, N-acetylmannosamine, and Trehalose) which can be seen in Table 8.

Table 8 Assigned ^{1}H NMR resonances with chemical shifts (δ, ppm), proton homonuclear scalar couplings (J, Hz), and signal diversity.

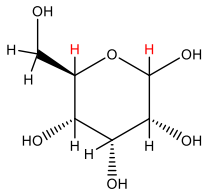
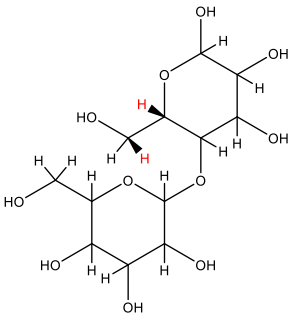
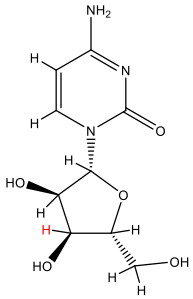
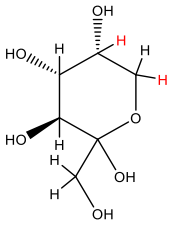
Molecular Formula (Assigned ^{1}H in red)	^{1}H δ (ppm)	^{13}C δ (ppm)	J (Hz)	Multiplicity
 Allose	4.6617	95.858	3.8095	Doublet
	3.7231	63.762	5.915	Doublet
 Cellobiose	3.567	72.39	5.8894	Sextuplet
	3.8532	62.337	5.0294	Doublet
 Cytidine	3.913	63.48	7.3914 4.9276	Quartet
 D-fructose	3.6995	66.193	8.0525 3.4701	Doublet of doublets
	3.5573	66.64	6.415 3.49	Quartet

Table 8. *Cont.*

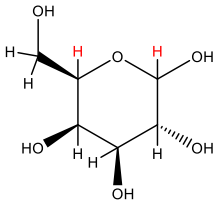
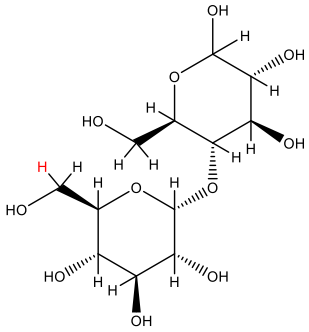
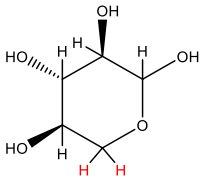
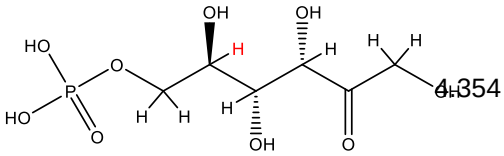
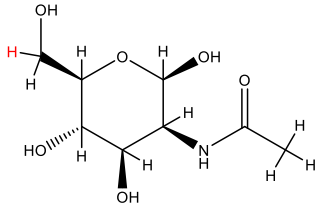
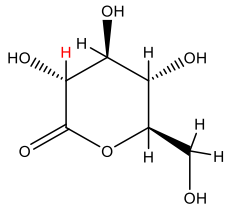
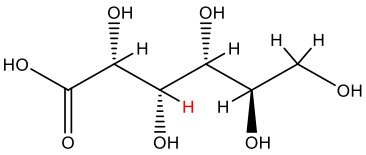
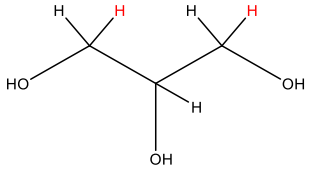
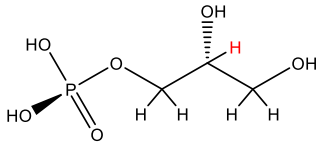
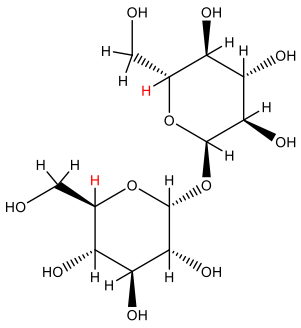
Molecular Formula (Assigned ^{1}H in red)	^{1}H δ (ppm)	^{13}C δ (ppm)	J (Hz)	Multiplicity
 D-galactose	4.3537 3.859	73.8375 72.38	3.8039 4.923	Doublet Doublet
 D-maltose	3.742	63.26	5.901	Doublet of doublets
 D-xylose	3.633 3.736	72.37 63.28	9.635 9.003 3.3477	Doublet of doublets Doublet of doublets
 Fructose 6-phosphate	4.354	76.778	6.5308 3.3379	Multiplet
 N-acetylmannosamine	3.8338	62.37	5.0656	Doublet

Table 8. *Cont.*

Molecular Formula (Assigned 1H in red)	^{1}H δ (ppm)	^{13}C δ (ppm)	J (Hz)	Multiplicity
 Gluconolactone	4.375	76.78	9.7204	Doublet
 Gluconic acid	4.1234	73.18	3.5525	Quartet
 Glycerol	3.562	65.32	2.893	Doublet
 Glycerol 3-phosphate	3.817	67.41	5.026	Quintuplet
 Trehalose	3.4556	72.42	9.2958	Multiplet

3.5. Conclusions

The present work proposes a MIR-NIR and NMR-based metabolomics approach for respectively profiling and targeting the temperature and sieving dependence within the sample treatment and preparation to efficiently degrade the lignocellulosic residues in *Theobroma cacao* L. pod husk (CPH) with microwave-assisted hydrothermal pretreatments (MA-HTP). First, the MIR-NIR first derivative absorbance data matrix can straightforwardly discriminate between simple carbohydrates and complex cellulose, hemicellulose, lignin, pectin, and CPH partial degradation moieties, employing the produced unsupervised PCA and supervised PLS-DA, sPLS-DA and OPLS-DA score plots, as simple carbohydrates are defined in a positive PC1 subspace. In contrast, polymeric CHO scaffolds are defined as negative PC1. The closer a pretreatment score approaches a PC1>0, the better degradation is expected, as observed for MA-HT pretreatments done at higher temperatures and, to a lesser extent, supervised MSA algorithms for samples sieved with a filter smaller than *150 µm*. The present work is the first to exploit a MIR-NIR first derivative absorbance data matrix for metabolomics studies. NMR-based metabolomics supports the observations done with the MIR-NIR approach, as a similar trend is observed within the produced unsupervised PCA and supervised PLS-DA, sPLS-DA, and OPLS-DA score plots: different PC1 dimensionality between simple carbohydrates and complex CHO scaffolds expected in CPH raw samples and MA-HTP and proximity between efficient CPH pretreatments and CHO standards as a fingerprint to elucidate the efficiency of the temperature-dependent microwave-assisted processes, as well as the sieving dependence in a minor extent. NMR targeted analysis employing a rigorous match between HSQC, TOCSY, and J-resolved inputs concerning the human metabolome data basis (HMDB) identifies 20 discriminant metabolites as possible biomarkers that can be tracked in MA-HTP optimizations. Future studies might evaluate further CPH MA-HTP dependencies such as microwave frequencies, pressure dependence, temperature ramps and increments, and smaller sieving processes with the present methodology.

Author contribution

E.S.P.: Investigation, sample preparation, micro-wave assisted hydrothermal pretreatments, experimental MIR-NIR and NMR data acquisition, pre-and post-processing, multivariate statistical analysis, 2D NMR spectral assignment and targeted analysis, writing original draft; B.R.T.: NMR resources; T.E.S.: formal analysis, supervision; H.Z.-P.: Supervision, CPH sample acquisition, sample preparation, writing; J.E.H.-P.: Conceptualization, methodology, investigation, formal analysis, writing, review, and editing. All authors have read and agreed to the published version of the manuscript.

Funding

E.S.P. acknowledges CONACyT graduate scholarship No. 763652, the Laboratorio de Productos Naturales del Departamento de Química of the Universidad Autónoma Chapingo, the Instituto Nacional de Investigaciones Forestales Agrícolas y Pecuarias (INIFAP), J.E.H.-P. Thanks to financial support from the Instituto Politécnico Nacional (IPN) program “Programa Institucional de Contratación de Personal Académico de Excelencia (PICPAE)” and “Estímulos al Desempeño de los Investigadores (EDI)” EDI-IPN financial support grant No. F-00318, H.Z.P. thanks financial support to the Dirección General de Investigación y Posgrado (DGIP) of the Universidad Autónoma Chapingo for funding projects: 20007-DTT-90, 22003 EIP and C. Carlos Hernández Echeverría for providing cocoa pod husk of the Carmelo variety.

3.6. References

Adjin, T. M., Asiedu, N., Dodoo, A. D., Karam, A., & Amaniampong, P. N. (2018). Thermochemical conversion and characterization of cocoa pod husks are potential agricultural waste from Ghana. *Industrial Crops and Products*, 119, 304-312. https://doi.org/10.1016/j.indcrop.2018.02.060

Alamar, P. D., Caramês, E. T. S., Poppi, R. J., & Pallone, J. A. L. (2020). Detection of Fruit Pulp Adulteration Using Multivariate Analysis: Comparison of NIR, MIR and Data Fusion Performance. *Food Analytical Methods*, 13(6), 1357–1365. https://doi.org/10.1007/s12161-020-01755-x

American Society for Testing and Materials (1977). Method of Test for Alpha-Cellulose in Wood ASTM D1103 – 60. Annual book of ASTM standards. Part 22: *Wood; Adhesives. Pennsylvania*: ASTM. Pp 346-348. 1977.

Amirvaresi, A., Nikounezhad, N., Amirahmadi, M., Daraei, B., & Parastar, H. (2021). Comparison of near-infrared (NIR) and mid-infrared (MIR) spectroscopy based on chemometrics for saffron authentication and adulteration detection. *Food Chemistry*, 344, 128647. https://doi.org/10.1016/j.foodchem.2020.128647

Balcázar-Zumaeta, C. R., Castro-Alayo, E. M., Cayo-Colca, I. S., Idrogo-Vásquez, G., & Muñoz-Astecker, L. D. (2023). Metabolomics during the spontaneous fermentation in cocoa (*Theobroma cacao* L.): An exploratory review. *Food Research International*, *163*, 112190. https://doi.org/10.1016/j.foodres.2022.112190

Brienzo, M. (Ed.). (2022). Hemicellulose Biorefinery: A Sustainable Solution for Value Addition to Bio-Based Products and Bioenergy. *Springer Nature Singapore*. https://doi.org/10.1007/978-981-16-3682-0

Cain, N., Alka, O., Segelke, T., von Wuthenau, K., Kohlbacher, O., & Fischer, M. (2019). Food fingerprinting: Mass spectrometric determination of the cocoa shell content (*Theobroma cacao* L.) in cocoa products by HPLC-QTOF-MS. *Food Chemistry*, 298, 125013. https://doi.org/10.1016/j.foodchem.2019.125013

Chakraborty, P., Kumar, R., Chakrabortty, S., Saha, S., Chattaraj, S., Roy, S., Banerjee, A., Tripathy, S. K., Kumar Ghosh, A., & Jeon, B.-H. (2024). Technological Advancements in the pretreatment of lignocellulosic biomass for effective valorization: A review of challenges and prospects. *Journal of Industrial and Engineering Chemistry*. https://doi.org/10.1016/j.jiec.2024.03.025

da Cruz Ferraz Dutra, J., Passos, M. F., Moretti, É. R., do Nascimento, L. A., da Silva, A. J., da Silva, T. F., Aguiar, R. H., dos Santos Rodrigues, L., & Mockaitis, G. (2022). Methane production from lignocellulosic biomass using Hydrothermal Pretreatment. *Biomass Conversion and Biorefinery*, *14*

Dahunsi, S. O., Adesulu, D. A. T., & Izebere, J. O. (2019). Cleaner energy through Cocoa (Theobroma cacao) pod husk liquefaction: Pretreatment and process

optimization. *Journal of Cleaner Production*, 226, 578-588. https://doi.org/10.1016/j.jclepro.2019.04.112

Duarte, L. C., Sampaio, B., & Carvalheiro, F. (2024). Organosolv pretreatment of lignocellulosic biomass. *Handbook of Biorefinery Research and Technology*, 1–28. https://doi.org/10.1007/978-94-007-6724-9_81-1

Gallego, A. M., Zambrano, R. A., Zuluaga, M., Camargo R., A. V., Candamil C., M. S., Romero V., A. P., & Arboleda V., J. W. (2022). Analysis of fruit ripening in *Theobroma cacao* pod husk based on untargeted metabolomics. *Phytochemistry*, 203, 113412. https://doi.org/10.1016/j.phytochem.2022.113412

García-Pérez, P., Becchi, P. P., Zhang, L., Rocchetti, G., & Lucini, L. (2024). Metabolomics and chemometrics: The next-generation analytical toolkit for evaluating food quality and authenticity. *Trends in Food Science & Technology*, *147*, 104481. https://doi.org/10.1016/j.tifs.2024.104481

Gatot S. H., (2012). Synthesis and characterization of sodium carboxymethylcellulose from Cacao pod husk (*Theobroma cacao* L.). *African Journal of Food Science*, 6(6). https://doi.org/10.5897/AJFS12.02 https://doi.org/10.5897/AJFS12.020

Gómez, H. C., Mazo, M. P., Penagos, V. L., Serpa, G. A., Eceiza, A., Urbina, L., Velásquez, C., J., Gañán, R. P., Vélez, A. L., & Zuluaga, R. (2020). Cocoa shell: An industrial by-product for preparing holocellulose nanofibers and fat suspensions. *Cellulose*, 27(18), 10873-10884. https://doi.org/10.1007/s10570-020-03222-6

Herbert, P. J. E., Lozada, R. J. D., Ortega, R. A. E., Hernández, L. R., & Anaya, P. C. (2021). Nuclear Magnetic Resonance Metabolomics with Double Pulsed-Field-Gradient Echo and Automatized Solvent Suppression Spectroscopy for Multivariate Data Matrix Applied in Novel Wine and Juice Discriminant Analysis. *Molecules*, 26(14), 4146. https://doi.org/10.3390/molecules26144146

ICCO. (2024, March 18). *February 2024 Quarterly Bulletin of Cocoa Statistics*. International Cocoa Organization. https://www.icco.org/february-2024-quarterly-bulletin-of-cocoa-statistics/

Jiménez, C., A. M., Martín T., S., Ortega, G., F., & Camacho, J. (2021). PLS-DA vs sparse PLS-DA in food traceability. A case study: Authentication of avocado samples. *Talanta,* 224, https://doi.org/10.1016/j.talanta.2020.121904

Kustov, L. M., Kustov, A. L., & Salmi, T. (2022). Processing of lignocellulosic polymer wastes using microwave irradiation. *Mendeleev Communications*, 32(1), 1-8. https://doi.org/10.1016/j.mencom.2022.01.001

Lin, S. Y., & Dence, C. W. (Eds.). (1992). Methods in Lignin Chemistry. *Springer Berlin Heidelberg.* https://doi.org/10.1007/978-3-642-74065-7

Liu, Y., Chen, W., Xia, Q., Guo, B., Wang, Q., Liu, S., Liu, Y., Li, J., & Yu, H. (2017). Efficient Cleavage of Lignin-Carbohydrate Complexes and Ultrafast Extraction of Lignin Oligomers from Wood Biomass by Microwave-Assisted Treatment with Deep Eutectic Solvent. *ChemSusChem*, 10(8), 1692–1700. https://doi.org/10.1002/cssc.201601795

López-Linares, J. C., García-Cubero, M. T., Lucas, S., González-Benito, G., & Coca, M. (2019). Microwave-assisted hydrothermal as Greener Pretreatment of Brewer's spent grains for biobutanol production. *Chemical Engineering Journal*, *368*, 1045–1055. https://doi.org/10.1016/j.cej.2019.03.032

Lu, F., Rodriguez, G. J., Van, D. I., Westwood, N. J., Shaw, L., Robinson, J. S., Warren, G., Chatzifragkou, A., McQueen, M. S., Gomez, L., Faas, L., Balcombe, K., Srinivasan, C., Picchioni, F., Hadley, P., & Charalampopoulos, D. (2018). Valorization strategies for cocoa pod husk and its fractions. *Current Opinion in Green and Sustainable Chemistry*, 14, 80-88. https://doi.org/10.1016/j.cogsc.2018.07.007

Mellinas, A. C., Jiménez, A., & Garrigós, M. C. (2020). Optimization of microwave-assisted cocoa bean shell waste extraction and evaluation of its antioxidant, physicochemical, and functional properties. *LWT*, *127*, 109361. https://doi.org/10.1016/j.lwt.2020.109361

Muharja, M., Darmayanti, R. F., Palupi, B., Rahmawati, I., Fachri, B. A., Setiawan, F. A., Amini, H. W., Rizkiana, M. F., Rahmawati, A., Susanti, A., & Putri, D. K. Y. (2021). Optimization of Microwave-Assisted Alkali Pretreatment for Enhancement of Delignification Process of Cocoa Pod Husk. *Bulletin of Chemical Reaction Engineering & Catalysis*, *16*(1), 31-43. https://doi.org/10.9767/bcrec.16.1.8872.31-43

Nelson, N., Darkwa, J., Calautit, J., Worall, M., Mokaya, R., Adjei, E., Kemausuor, F., & Ahiekpor, J. (2021). Potential of Bioenergy in Rural Ghana. *Sustainability*, *13*(1), 381. https://doi.org/10.3390/su13010381

Nepper-Davidsen, J., Magnusson, M., Lawton, R. J., Fletcher, D., Holmes, G., & Glasson, C. R. K. (2024). Chemometric models for high-throughput biomass grading of the Kelp Ecklonia radiata, using mid-infrared (MIR) and near-infrared (NIR) spectroscopy. *Algal Research*, *77*, 103352. https://doi.org/10.1016/j.algal.2023.103352

Pinyo, J., Luangpituksa, P., Suphantharika, M., Hansawasdi, C., & Wongsagonsup, R. (2017). Improvement of sago starch extraction process using various pretreatment techniques and their combination: Improvement of sago starch extraction process. *Starch - Stärke*, 69(9-10), 1700005. https://doi.org/10.1002/star.201700005

Priyangini, F., Walde, S. G., & Chidambaram, R. (2018). Extraction optimization of pectin from cocoa pod husks (*Theobroma cacao* L.) with ascorbic acid using response surface methodology. Carbohydrate Polymers, 202, 497-503. https://doi.org/10.1016/j.carbpol.2018.08.103

Puke, M., Godina, D., & Brazdausks, P. (2024). Catalyzed hydrothermal pretreatment of Oat Husks for integrated furfural and Lignocellulosic Residue production. *Polymers*, *16*(5), 707. https://doi.org/10.3390/polym16050707

Roy, S. (2022a). PRE-TREATMENT METHODS OF LIGNOCELLULOSIC BIOMASS FOR BIOFUEL PRODUCTION. CRC Press is an imprint of Taylor & Francis Group, LLC. https://doi.org/DOI: 10.1201/9781003203414

Ruengsrichaiya, B., Perez, C. L., & Mussatto, S. I. (2024). Deep eutectic solvents and their use for sustainable pretreatment of lignocellulosic biomass. *Handbook of Biorefinery Research and Technology*, 1–28. https://doi.org/10.1007/978-94-007-6724-9_62-1

Sarmiento, V. Z., Vandenberghe, L., Rodrigues, C., Tanobe, V. O. A., Marín, O., de Melo, P. G. V., Ghislain, R. H. L., Góes, N. A., & Soccol, C. R. (2021). Cocoa pod husk valorization: Alkaline-enzymatic pre-treatment for propionic acid production. *Cellulose*, *28*(7), 4009-4024. https://doi.org/10.1007/s10570-021-03770-5

Sun, D., Lv, Z. W., Rao, J., Tian, R., Sun, S. N., & Peng, F. (2022). Effects of hydrothermal pretreatment on the dissolution and structural evolution of hemicelluloses and lignin: A review. *Carbohydrate Polymers*, 281, 119050. https://doi.org/10.1016/j.carbpol.2021.119050

Swami, S., Suthar, S., Singh, R., Thakur, A. K., Gupta, L. R., & Sikarwar, V. S. (2024). The potential of ionic liquids as an emerging green solvent for the pretreatment of Lignocellulosic Biomass. *Environmental Science and Pollution Research*, *31*(9), 12871–12891. https://doi.org/10.1007/s11356-024-32100-y

Technical Association for the Pulp and Paper Industries (1997). Preparation of extractive free-wood. TAPPI Test Method T 204 cm-97. TAPPI PRESS. Atlanta.

Technical Association for the Pulp and Paper Industries (2002). Lignin acid is insoluble in wood and pulp. TAPPI Test Method T 222 om-02. TAPPI Press. Atlanta

Vásquez, Z. S., de Carvalho, N., D. P., Pereira, G. V. M., Vandenberghe, L. P. S., de Oliveira, P. Z., Tiburcio, P. B., Rogez, H. L. G., Góes Neto, A., & Soccol, C. R. (2019). Biotechnological approaches for cocoa waste management: A review. *Waste Management*, *90*, 72-83. https://doi.org/10.1016/j.wasman.2019.04.030

Vriesmann, L. C., de Mello, C., A. R. D., & de Oliveira, P. C. L. (2011). Cacao pod husks (*Theobroma cacao* L.): Composition and hot-water-soluble pectins. *Industrial Crops and Products*, 34(1), 1173-1181. https://doi.org/10.1016/j.indcrop.2011.04.004

Yuan, Y., Zou, P., Zhou, J., Geng, Y., Fan, J., Clark, J., Li, Y., & Zhang, C. (2019). Microwave-assisted hydrothermal extraction of non-structural carbohydrates

and hemicelluloses from tobacco biomass. *Carbohydrate Polymers*, 223, 115043. https://doi.org/10.1016/j.carbpol.2019.115043

Zheng, Y., Shi, J., Tu, M., & Cheng, Y.-S. (2017). Principles and Development of Lignocellulosic Biomass Pretreatment for Biofuels. *Advances in Bioenergy, 2,* 1-68. (Vol. 2, pp. 1–68). Elsevier. https://doi.org/10.1016/bs.aibe.2017.03.001

4. OPTIMIZATION OF MONO- AND DI-SACCHARIDE EXTRACTION FROM COCOA POD HUSK

ABSTRACT

Cocoa pod husk (CPH) is a rich source of platform molecules that can be used to create value-added products. The present study aimed to optimize microwave-assisted hydrothermal pre-treatment (MA-HTP) from CPH and hemicellulose from CPH (HMC-CPH), using a combination of response surface analysis (RSA) Box Behnken design (BBD) and proton nuclear magnetic resonance (^{1}H NMR Qu) identification and quantification, to provide an efficient protocol for the extraction of mono- and disaccharides. The methodology consisted of 15 CPH MA-HTPs and 15 HMC-CPH MA-HTPs (each in triplicate) designed by RSA-BBD; the independent variables were time, temperature, and power; the dependent variable was the concentration of extraction products. Glucose, sucrose, and fructose were identified as products of the MA-HTP extractions of CPH and HMC-CPH as determined by ^{1}H NMR. A higher sucrose content was found for CPH (45.62%) compared to HMC-CPH (17.34%) at a 95% confidence level, and a high fructose content for both CPH and HMC-CPH (37.88% and 35.37%, respectively), both CPH and HMC-CPH had minimal glucose concentrations (4.57% and 0.93% respectively). From the RSA-BBD, optimal temperature, power, and time points were predicted for CPH glucose: 135.4°C-180.6 W and 5.8 min; sucrose: 154.3°C-256.3 W and 20. 2 min; fructose: 129.5°C-173.8 W and 5.27 min. For HMC-CPH glucose: 142.2°C-204.4 W and 10.5 min; sucrose: 148.8°C-215.6 W and 14.3 min; and fructose: 151.6°C-231.6 W and 13 min.

Keywords: Response surface analysis (RSA), Box Behnken design (BBD), microwave-assisted hydrothermal pre-treatment (MA-HPT), proton nuclear magnetic resonance quantification (^{1}H NMR Qu), and hemicellulose from cocoa pod husk (HMC-CPH).

Doctoral Thesis: Doctorado en Ciencias Agroalimentarias, Universidad Autónoma Chapingo.

Author: Edna Elena Suárez Patlán, M. C.
Advisor: Holber Zuleta Prada, Dr.

4.1. Introduction

The cocoa industry generates large amounts of waste; in 2021, cocoa bean production will be 4.2 million tons (Ramos Escudero et al., 2023). Cocoa beans make up about 10% of the total weight of the fruit and are used to make chocolate (Jarrín Chacón et al., 2023). The cocoa pod husk (CPH) is the main by-product of the cocoa industry, accounting for approximately 75% of the total weight of the fruit (Lu et al., 2018; Vasquez et al., 2019), with annual estimates of 48 million tons of CPH worldwide (Quiceno Suarez et al., 2024b). Management of these residues is costly and complex, and they generally remain on the land, causing odors, soil contamination, greenhouse gas emissions, and excessive growth of pathogenic fungi that cause diseases that affect crop production (Priyangini et al., 2018; Soares & Oliveira, 2022). However, it is also an essential and challenging renewable source for biorefining and has been the focus of research interest for several decades (Perwitasari et al., 2021). CPH is rich in biologically active molecules with nutraceutical properties, consisting of carbohydrates, lignin, proteins, lipids, pectin, minerals, theobromine, phenolic compounds, and tannins (Vasquez et al., 2019). CPH pectin's Sug includes xylose, arabinose, rhamnose, galactose, mannose, glucose, and galacturonic acid, making it a good source for the food industry. (Muñoz Almagro et al., 2019). Hemicellulose is rich in xylose with neutral sugar substitutions and phenolic acid esters used in the food and pharmaceutical industries (Valladares Diestra et al., 2022). In this sense, CPH can be pretreated to separate its components, such as lignin from cellulose and hemicellulose, and to obtain molecules from xylo-oligosaccharides. (Rogoski et al., 2023). There are several types of pre-treatment, including thermal, chemical, physical, biological, and combinations of these (Fan et al., 2024; Rahandi Lubis et al., 2024). Microwave-assisted hydrothermal pretreatment (MA-HTP) is a pretreatment that combines microwave methodology with chemical reagents and is known to be more efficient than conventional heating (Muharja et al., 2021). Several studies have been reported on the extraction of pectin from CPH with MA-HTP in combination with strong acids, organic acids, enzymatic methods, etc. (Jarrín Chacón et al., 2023; Pangestu et al., 2020b; Priyangini et al., 2018; Saelee

et al., 2022), the results highlight the impact on time savings, which is why it is considered a green technology (Hennessey Ramos et al., 2021; Valladares Diestra et al., 2022). MAE has also been combined with solid acids to delignify lignocellulosic materials with maximum lignin removal (Muharja et al., 2021). The efficacy of MA-HTP can be assessed using a mathematical model that predicts the statistical significance of the dependent variables and their interactions, providing optimal conditions using tools that reduce the number of experiments (Ferreira et al., 2007), these are response surface analyses (RSA), which allow the variables of an experiment to be optimized to optimize a response (Jarrín Chacón et al., 2023). An example is the Box Behnken design (BBD), a mathematical model with first and second-order coefficients, a three-level incomplete factorial design for three factors (Aycan et al., 2023). The BBD is slightly more efficient than the central composite design and much more efficient than the three-level full factorial designs (Ferreira et al., 2007; Nassef et al., 2024). Therefore, MA-HTP, in combination with RSA, is a technique used to extract active compounds from plant materials, where the relationship between solvent, extraction time, and irradiation power is studied. (Chan et al., 2014). On the other hand, NMR has been successfully used as a quantitative method for natural products, as all components resonate at deficient concentrations, just above the detection threshold (5-10 µM) (del Campo et al., 2016); it is also over 98% accurate and is therefore considered a reliable technique for quantitative estimation. This has been established through validation procedures for precision, accuracy, linearity, reproducibility, robustness, selectivity, and specificity. However, this is only true if the sampling and processing parameters are well-known (Bharti & Roy, 2012). ^{1}H NMR spectra of carbohydrates have constant shifts concerning a reference because they are not affected by pH or ionic strength due to the absence of ionizable groups. Therefore, each sample's acquisition parameters remain constant, and a list of frequencies can be constructed (Schievano et al., 2017).

To date, no publication proposes the extraction of mono- and disaccharides from CPH by MA-HTP using only water as an extraction medium, so it is feasible to

seek the optimization of the process parameters, reducing time and cost using a statistical model. Therefore, the present study aims to optimize the conditions for the extraction of carbohydrates from CPH and HMC-CPH by MA-HTP, using DBB response surface analysis and ^{1}H NMR quantification, to make a green pretreatment for this lignocellulosic material more efficient.

4.2. Materials and Methods

4.2.1. Study material

In this study, CPH and HMC-CPH were used as raw materials. The species studied was *Theobroma cacao* L. variety Carmelo, collected at the Rancheria Rio Seco, municipality of Cunduacan [latitude: 18° 7'55.90" N, longitude: 93° 18'4.49" W] and at the farm Jesús María, municipality of Comalcalco [latitude: 18° 11'0.22" N, longitude: 93° 14'28.02" W], state of Tabasco, Mexico. At altitudes of 10 and 13 meters above sea level (MASL). The CPH was dried under ambient conditions with indirect sun exposure (average maximum temperature 21°C, average minimum temperature 6 °C, and relative humidity less than 10%). The CPH was then mechanically ground in a Thomas Wiley Model 4 laboratory mill (TP4274E70520A, Thomas Scientific, Swedesboro, New Jersey, USA), and the particle size was graded using a US STD 100 laboratory sieve (W. S. Tyler, Ohio, USA). The particle that remained on the sieve was classified as more significant than 150 µm, and the particle that passed through the sieve was classified as smaller than 150 µm, the latter used for this study. Finally, this material was oven-dried at 102°C ± 3°C to constant weight. Hemicellulose was obtained from a fraction of this material according to the methodology reported by Peng & She (Peng & She, 2014). Finally, CPH and HMC-CPH were reserved for further processing.

4.2.2. Reagents

The standards and solvents used for the analysis of the pretreatments were D-fructose ≥ 99% (CAS no. 57-48-7), sucrose ≥ 99.5% (CAS no. 57-50-1), and D-

glucose ≥ 99.5% (CAS no. 50-99-7), deuterium oxide 99.9% (CAS no. 7789-20-0) and 3-(trimethylsilyl)propionic acid-2,2,2,3,3-d4-acid sodium salt 98% (CAS no. 24493-21-8), all from Sigma Aldrich (Merck KGaA, Saint Louis Missouri, USA).

4.2.3. Microwave-Assisted Hydrothermal Pretreatment of CPH and HMC-CPH

A MARS 6TM microwave digestion system (CEM Corporation, Mecklenburg, North Carolina, USA) was used to extract CPH and HMC-CPH. 1 g of CPH fines less than 150 µm and 1 g of HMC-CPH were weighed and placed separately in lidded silicone cups (Xpress Plus, CEM Corporation) to which 10 mL of distilled water was added to obtain a 1:10 ratio (CPH: distilled H_2O). The CPH and HMC-CPH cups with their respective triplicates were placed in the carousel of the microwave oven programmed with a temperature, power, and time previously defined by the BBD. From this pre-treatment, a heterogeneous biphasic mixture with a solid and an aqueous phase was obtained, the solid phase corresponding to the fine particles of the study sample that did not undergo visible changes after the treatment and the aqueous phase to the product of the hydrothermal extraction; this mixture was poured into a conical tube (50 mL Eppendorf) and centrifuged at 3700 rpm for 40 minutes in a Centra CL2 centrifuge (Cat. No. 426, Thermo Scientific, Needham, Massachusetts, USA). The aqueous fraction was collected and freeze-dried at -50°C under reduced pressure of 0.045 bar in a 2.5 L Freezone Legacy freeze-dryer (Labconco Corporation, Kansas City, Missouri, USA), and the freeze-dried CPH and HMC extract was then analyzed by nuclear magnetic resonance (NMR).

4.2.4. Acquisition of ^{1}H NMR spectra

NMR spectroscopy was performed on a VARIAN 600 MHz (14.1 T) Premium COMPAC spectrometer (Agilent Technologies Inc., California, USA). For the analysis of CPH and HMC-CPH extracts, analytes were prepared by dissolving 30 mg of each lyophilized extract in 600 µL of a five mM solution of TSP (3-(trimethylsilyl) propionic acid-2,2,2,3,3-d4-acid sodium salt) and D2O (deuterium

oxide). For this purpose, 5 mm NMR tubes were used and sonicated for 20 minutes to dilute the sample completely. The ^{1}H NMR spectra were acquired with 64 scans of 32 k complex points at 25°C, a spectral width of 16 ppm, an acquisition time of 4 s, a relaxation time of 2 s, an angle of 90°, and an acquisition time of 4 min. The water suppression scheme used was PRESAT; the presaturation during the relaxation delay was performed at the minimum power for complete water suppression, and the receiver gain was set to 30 and held constant for all spectra. Spectrum acquisition and data processing were performed according to the methodology of del Campo et al. and Hernández Bolio et al. (del Campo et al., 2016; Hernández Bolio et al., 2021). The phase correction, baseline correction, and integration of the spectra obtained on the signals of interest were performed in MNova software (Mestrelab Research SL, Santiago de Compostela). The integral values were entered into Microsoft Excel spreadsheets for the quantification process. Quantification was carried out using the formula:

$$P_x = \frac{I_x}{I_{std}} * \frac{N_{std}}{N_x} * \frac{M_x}{M_{std}} * \frac{W_{std}}{W_x} * P_{std}$$

Where: *I, N, M, W,* and P, are the area of the integral, the number of nuclei in the molecule, the molar mass, the gravimetric weight, and the purity of the analyte (x) and standard (std), respectively (Bharti & Roy, 2012; Hou et al., 2014).

4.2.5. Análisis de superficie de respuesta, diseño Box Behnken

The RSA methodology was used to optimize the microwave-assisted hydrothermal pretreatment for carbohydrate extraction from CPH and HMC-CPH, using the Box Behnken design to model and optimize the experimental conditions using the second-order polynomial equation (Equation 1):

$$y = \beta o + \sum_{i=1}^{k} \beta i Xi + \sum_{i=1}^{k} \beta ii X^2 i + \sum_{i=1}^{k} \sum_{j=1}^{k} \beta ij Xi Xj + \varepsilon \qquad \text{Equation 1}$$

Where: y is the dependent variable; k is the sample number, ij are the index numbers of the samples; βo is the lag term; βi is the first-order linear effect of the input factor (Xi); βii is the quadratic effect of the input factor (squared) (Xi) and

βij is the linear-linear interaction effect between the input factors $Xi - Xj$. (Aycan et al., 2023; Chan et al., 2014; Jarrín Chacón et al., 2023; Muharja et al., 2021). In the Box Behnken design, there is a relationship between the uncoded and coded independent variables, shown in Equation 2.

$$X_i = (x_i - x_0)/\,\Delta_{xi} \qquad \text{Equation 2}$$

Where X_i, x_i, x_0 are coded value, natural value, and natural value at the midpoint (of the i-th independent variable) and Δ_{xi} is the change value of an independent variable (Aycan et al., 2023).

Three independent variables with three levels each (-1, 0, and 1) were used; the independent variables were temperature (T=100, 150, and 200°C), power (P= 100, 200, and 300 W), and time (t= 5, 10 and 15 min), giving a total of 15 experiments performed in triplicate, the runs were conducted in an ordered fashion according to the BBD, the response variable was the quantification of carbohydrates determined by NMR. Analyses were performed in R software (RStudio Inc. Version 4.2.3, Boston, Massachusetts, USA).

4.3. Results and discussion

4.3.1. ^{1}H NMR spectra elucidation

As a result, two monosaccharides (glucose and fructose) and one disaccharide (sucrose) could be identified by proton nuclear magnetic resonance from the freeze-dried aqueous extracts obtained from microwave-assisted hydrothermal treatments of cocoa pod husk, as well as from the hemicellulose extracted from the same. Table 1 shows the chemical shifts and the diversity of signals identified in the experimental ^{1}H NMR spectra of MA-HTP of CPH and HMC-CPH, as well as in the spectra of the D-fructose, sucrose, and D-glucose standards.

Table 9 Characteristics of the ^{1}H NMR signals observed in MA-HTP of CPH and HMC-CPH.

D-Fructose, Sucrose, and D-Glucose Standards			CPH and HMC-CPH MA-HTP Extracts		
Assignment	Chemical shifts used for identity check		Chemical shift		
	δ (ppm)	Multiplicity	identity check δ (ppm)	quantify δ (ppm)	Multiplicity
Fructose	4.11	d		4.11	d
Fructose	4.01	t		4.01	t
Fructose	3.89	dd	3.90		m
Sucrose	5.40	d		5.40	d
Sucrose	4.21	d		4.20	d
Sucrose	4.05	t	4.05		t
Sucrose	3.76	t	3.76		t
Sucrose	3.67	s	3.67		s
Sucrose	3.55	dd	n/d		n/d
Sucrose	3.47	t	3.47		t
Glucose	5.24	d		5.24	d
Glucose	4.65	d		4.65	d
Glucose	3.89	dd	n/d		n/d
Glucose	3.53	dd	3.53		dd
Glucose	3.25	t	3.25		t

s, singlet; d, doublet; t, triplet; m, multiplet; dd, doublet–doublet

As seen in Table 1, the identification of the monosaccharides and disaccharides was supported by the standard spectra and data from the literature. Three signals were identified for fructose elucidation, and only two for quantification. In contrast, six signals were identified for sucrose elucidation, two of them for quantification, and two of the four signals were identified for glucose quantification. In this

respect, del Campo et al. (del Campo et al. 2016) report for fructose from honey samples two signals identified as multiplets at the same chemical shifts reported in this study (4.00 ppm and 4.10 ppm), but with different multiplicity. On the other hand, for sucrose, they report one signal with a doublet multiplicity at the chemical shift of 5.42 ppm, which coincides with the same signal reported in this study. Finally, they reported five signals for glucose; for β-glucose, one doublet and two doublets of doublets were identified at chemical shifts of 4.65 ppm, 3.40 ppm, and 3.25 ppm, respectively. In contrast, for α-glucose, two signals, one doublet and one doublet of doublets, were identified at chemical shifts of 5.22 ppm and 3.42 ppm, respectively. Four of the five signals identified in this study, two for β-glucose and two for α-glucose, were coincident, with one doublet of doublets missing at the chemical shift of 3.40 ppm corresponding to β-glucose. On the other hand, Al-Mekhlafi et al. (Al-Mekhlafi et al., 2021) reported six, five, and two signals respectively at chemical shifts 4.08 ppm, 4.00 ppm, 3.93 ppm, 3.79-3.85 ppm, 3.69 ppm, and 3.51-3.57 ppm for fructose, 5.39 ppm, 4.16 ppm, 4.07 ppm, 3.87-3.67 ppm and 3.53-3.42 ppm for sucrose and 5.19 ppm and 4.59 ppm for glucose, with differences of about 0.04-0.12 ppm for most chemical shifts. Finally, agreement was also found with the data of Spiteri et al. (Spiteri et al., 2015). They identified a signal for fructose with a chemical shift of 4.1 ppm, a signal for sucrose at 4.22 ppm, and two signals for glucose at 5.23 ppm and 3.24 ppm, corresponding to α-glucose and β-glucose respectively, but did not report signal multiplicity data. Identification of sugars by ^{1}H NMR is limited because oligosaccharides incorporated into an isolated spin system per monomer unit are separated from each other by glycosidic bonds and, considering that oligosaccharides can be formed from identical monomers and in the same sequences, they will overlap in the spectra so that only a few sugars can be identified (Schievano et al., 2017).

According to the PubChem databases (CID: 5984), fructose is a monosaccharide found in fruits and honey, soluble in water, ether, and alcohol, and is mainly used as a preservative and as an intravenous infusion in parenteral nutrition. Sucrose (PubChem CID: 5988), a glycosyl-glycoside composed of glucose and fructose

units, is used as an osmolyte, food sweetener, human metabolite, etc., and is mainly obtained from sugar cane and sugar beet. Finally, according to the PubChem databases (CID: 5793), D-glucose or dextrose, the most abundant isomer of glucose in nature, is produced by photosynthesis in plants and hepatic gluconeogenesis in humans. On the other hand, a study by Perwitasari et al. (Perwitasari et al., 2021) hydrolyzed CPH into monosaccharides, which were then metabolized by *A. niger* via the tricarboxylic acid cycle to citric acid, which could also be used in the food industry.

4.3.2. Optimization of CPH and HMC-CPH MA-HTP conditions through RSA-BBD

The ^{1}H NMR Qu results were used to optimize the MA-HTP of CPH and HMC-CPH using BBD response surface analysis. The BBD indicates the optimal points of temperature, power, and time in the MA-HTP of CPH and HMC-CPH. The three factors were coded as -1, 0, and 1 for low, medium, and high levels. This study presents the optimal level for each dependent variable, quadratic polynomial equations, and three-dimensional graphical representations. Table 2 shows the ^{1}H NMR Qu of carbohydrates, glucose, sucrose, and fructose obtained from 15 MA-HTP of CPH and 15 MA-HTP of HMC-CPH.

Table 10 BBD model for MA-HTP optimization of CPH and HMC-CPH by ^{1}H NMR Qu.

Independent variables			**Dependent variable Y**					
Coded -1,0,1			**Concentration**			**Concentration**		
[Uncoded]			**CPH**			**HMC-CPH**		
A (T °C)	B (P W)	C (t min)	Glucose (%)	Sucrose (%)	Fructose (%)	Glucose (%)	Sucrose (%)	Fructose (%)
-1 [100]	**-1** [100]	**0** [10]	1.0	51.10	29.90	1.2	7.70	16.91
1 [200]	**-1** [100]	**0** [10]	1.0	43.17	27.80	0.4	12.00	25.16
-1 [100]	**1** [300]	**0** [10]	9.4	68.02	70.26	2.2	26.23	44.78
1 [200]	**1** [300]	**0** [10]	1.0	39.20	27.07	0.6	8.67	23.87

-1 [100]	**0** [200]	**-1** [5]	1.2	50.39	29.56	0.5	8.15	20.20
1 [200]	**0** [200]	**-1** [5]	0.8	43.38	26.87	1.3	17.19	27.49
-1 [100]	**0** [200]	**1** [15]	0.7	39.41	23.51	1.1	20.90	35.43
1 [200]	**0** [200]	**1** [15]	1.1	37.91	28.08	1.7	26.92	49.61
0 [150]	**-1** [100]	**-1** [5]	0.5	34.63	20.94	1.4	13.98	21.58
0 [150]	**1** [300]	**-1** [5]	1.1	36.78	29.23	2.3	25.76	47.07
0 [150]	**-1** [100]	**1** [15]	7.5	52.79	57.48	2.8	16.42	41.23
0 [150]	**1** [300]	**1** [15]	14.0	61.78	88.57	1.8	22.96	44.31
0 [150]	**0** [200]	**0** [10]	1.0	42.15	27.91	7.8	35.80	75.56
0 [150]	**0** [200]	**0** [10]	0.8	35.13	23.10	2.3	28.57	51.86
0 [150]	**0** [200]	**0** [10]	1.4	43.38	31.83	2.1	17.45	34.10

4.3.3. RSA - BBD ANOVA

The concentration of glucose, sucrose, and fructose extracted from CPH and HMC-CPH by MA-HTP were determined in percent by ^{1}H NMR Qu for each treatment designed by BBD. Table 2 shows an apparent difference in mono- and disaccharide concentrations between the CPH and HMC-CPH whole matrix; however, by analysis of variance, only a statistically significant difference ($p \leq 0.05$) was found between the sucrose concentrations of CPH and HMC-CPH, with the highest concentration being the sucrose content extracted from cocoa pod husk.

On the other hand, a significant statistical difference ($p \leq 0.05$) was found between the concentrations of sucrose, fructose, and glucose extracted from CPH, given the variables of time and power. With 95% confidence, two groups of means were identified for sugar concentrations, two groups for time and two groups for power, finding sucrose and fructose from CPH in the same group with a mean of 45.62% and 37.88%, respectively, being the sugars with the highest concentration in cocoa pod husk. In contrast, a power of 300 W and time of 15 min is identified as the highest extraction of sugars, 37.69% and 36.08%, respectively. Similarly, a statistically significant difference ($p \leq 0.05$) was obtained between sucrose, fructose, and glucose concentrations of HMC-CPH; for sugars extracted from hemicellulose, temperature and power were significant. Using analysis (Tukey),

two groups were identified for temperature and two groups for power, with the temperature of 150°C giving the highest concentration of sugars with a mean of 23.37% and the power of 200 to 300 W with mean concentrations of 20.78% and 20.77% of sugars, while for the concentration of mono- and disaccharides three groups were obtained, the highest concentration being fructose, followed by sucrose and finally glucose with mean values of 35.37%, 17.34% and 0.93% respectively. Minimal glucose content was found in both CPH and HMC-CPH matrices, which could indicate cellulose's absence or minimal depolymerization. The higher glucose content can be attributed to free sugars reported in the literature as an extractable fraction of the lignocellulosic matrix.(Quiceno Suarez et al., 2024b; Vasquez et al., 2019). In summary, the hemicellulose fraction extracted from the cocoa pod husk is predominantly fructose, while the whole cocoa pod husk matrix is rich in sucrose and fructose. Table 3 shows the optimal treatment results predicted by the Box Behnken response surface design for glucose, sucrose, and fructose extracted from cocoa pod husk and the hemicellulose fraction of the same.

Table 11 Optimum values provided by RSA-BBD for glucose, sucrose, and fructose from CPH and HMC-CPH

Source	**Carbohydrate**	**Threshold (0.01)**		
		Temperature (°C)	**Power (Watt)**	**Time (min)**
CPH	*Glucose*	135.4	180.6	5.8
CPH	*Sacarose*	154.3	256.3	20.2
CPH	*Fructose*	129.5	173.8	5.27
HMC-CPH	*Glucose*	142.2	204.4	10.5
HMC-CPH	*Sacarose*	148.8	215.6	14.3
HMC-CPH	*Fructose*	151.6	231.6	13

As can be seen, the predicted optimum temperatures, power, and time for the extraction of glucose, sucrose, and fructose from CPH range between 129.5°C

and 154.3°C, between 173.8 W and 256.3 W, and between 5.27 min and 20.2 min, with higher conditions for sucrose and lower for fructose. Contrary to the conditions proposed for the extraction of glucose, sucrose and fructose extracted from HMC-CPH ranged between 142.2°C and 151.6°C, power between 204.4 and 231.6 W, and times between 10.5 min and 13 min, with central values prevailing in all cases.

The coded factor equations can predict the response to determine the levels of each factor, with high levels coded as 1, low levels -1, and intermediate levels coded as 0. Positive represents the synergistic relationship, while negative represents the antagonistic relationship between the variables (Ouattara et al., 2023). Uncoded quadratic polynomial equations of the independent variables to explain the efficacy of MA-HTP, constructed by multiple regression analysis of the BBD matrix for all dependent variables (equations 2-7), are presented.

$$\textit{CPH glucose} = -3.8 + 0.223\,T - 0.0843\,P - 1.24\,t - 0.000558\,T^*T + 0.000343\,P^*P + 0.0512\,t^*t - 0.000420\,T^*P + 0.00080\,T^*t + 0.00295\,P^*t \quad \text{Equation 3}$$

$$\textit{HMC-CPH glucose} = -25.2 + 0.241\,T + 0.0577\,P + 1.04\,t - 0.000778\,T^*T - 0.000102\,P^*P - 0.0388\,t^*t - 0.000040\,T^*P - 0.00020\,T^*t - 0.00095\,P^*t \quad \text{Equation 4}$$

$$\textit{CPH sucrose} = 134.8 - 0.765\,T - 0.265\,P - 3.11\,t + 0.00269\,T^*T + 0.001044\,P^*P + 0.113\,t^*t - 0.00104\,T^*P + 0.0055\,T^*t + 0.0034\,P^*t \quad \text{Equation 5}$$

$$\textit{HMC-CPH sucrose} = -127.9 + 1.160\,T + 0.475\,P + 2.67\,t - 0.00302\,T^*T - 0.000607\,P^*P - 0.057\,t^*t - 0.001093\,T^*P - 0.0030\,T^*t - 0.00262\,P^*t \quad \text{Equation 6}$$

$$\textit{CPH fructose} = 10 + 0.88\,T - 0.371\,P - 4.97\,t - 0.00218\,T^*T + 0.001660\,P^*P + 0.194\,t^*t - 0.00205\,T^*P + 0.0073\,T^*t + 0.0114\,P^*t \quad \text{Equation 7}$$

$$\textit{HMC-CPH fructose} = -236 + 2.14\,T + 0.816\,P + 6.48\,t - 0.00631\,T^*T - 0.001040\,P^*P - 0.196\,t^*t - 0.00146\,T^*P + 0.0069\,T^*t - 0.0112\,P^*t \quad \text{Equation 8}$$

Where: T, temperature; P, power; t, time.

The model is significant ($p \leq 0.05$) for sucrose and fructose from CPH; the response surface curves are shown in Figure 1 D/E/F and G/H/I, respectively. The model is significant ($p \leq 0.05$) for fructose from HMC-CPH; the response surface curves are shown in Figure 2 P/Q/R.

Regarding the calculated coefficient of determination, the CPH models for glucose, sucrose, and fructose and HMC-CPH for sucrose were found to be well-fitting ($p > 0.05$).

Figure 1 shows the three-dimensional valley response plots of the effect of the variables temperature, power, and time of CPH MA-HTP; these follow a minimum stationary point pattern, indicating that the response (carbohydrate yield) increases with central values of temperature and power and a low level of time for CPH glucose concentration. Meanwhile, CPH sucrose's yield is maximized at high power, time, and a central temperature value. For CPH fructose, the yield is maximized with central values of temperature and control and a low level of time.

Optimal extraction temperatures of 129.5°C and 154.3°C correlate with power ratings of 173.8 W and 256.3 W for glucose, sucrose, and fructose CPH, respectively. Extraction times range from 5.27 min to 20.2 min for average concentrations of 4.57% for glucose (Figure 1 A, B, and C), 45.62% for sucrose (1 D, E and F), and 37.88% for fructose (G, H and I). The extraction concentration is given on the y-axis.

The three-dimensional response surface plots in Figure 2 follow a maximum stationary point type pattern, with optimum yields of mono- and disaccharides increasing with a central value of temperature, power, and time for HMC-CPH glucose, and central values of temperature and control with a high level of time for maximum yields in HMC-CPH sucrose, and a central value of temperature and high levels of power and time required for optimum values of HMC-CPH fructose.

The optimum extraction temperatures for HMC-CPH carbohydrates range from 142.2°C to 151.6°C, and the optimum power for these carbohydrates ranges from

204.4 W to 231.6 W for HMC-CPH glucose, sucrose, and fructose. Extraction times range from 10.5 min to 14.3 min for average concentrations of 0.931% for glucose (Figure 2 J, K and L), 17.34% for sucrose (Figure 2 M, N and O), and 35.37% for fructose (Figure P, Q and R).

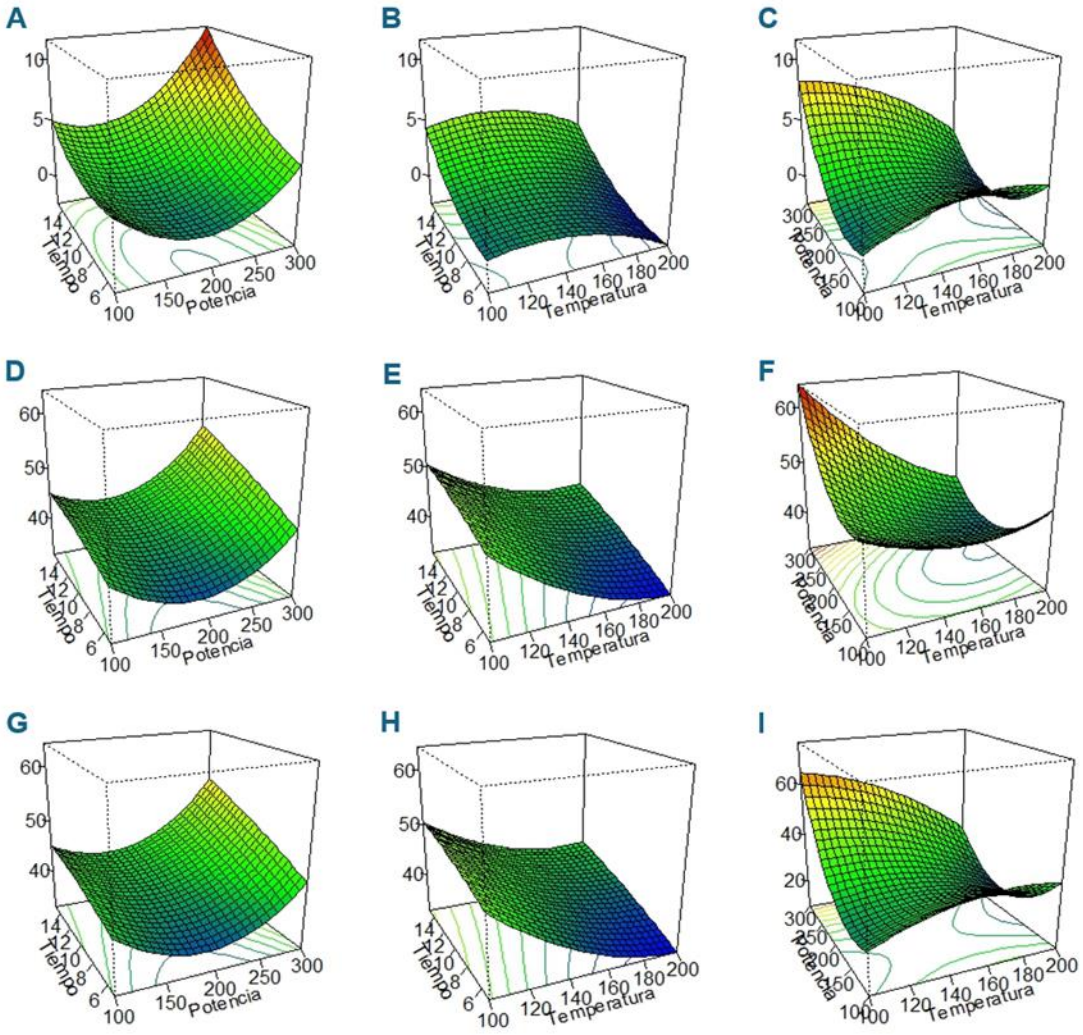


Figure 16 Effect of temperature, power, and time ratio in microwave-assisted hydrothermal pretreatment on total extraction yield of fructose, sucrose and glucose from CPH: (A) temperature plot, (B) power plot and (C) time plot

corresponding to glucose from CPH; (D) temperature plot, (E) power plot and (F) time plot corresponding to sucrose from CPH; and (G) temperature plot, (H) power plot and (I) time plot corresponding to fructose from CPH.

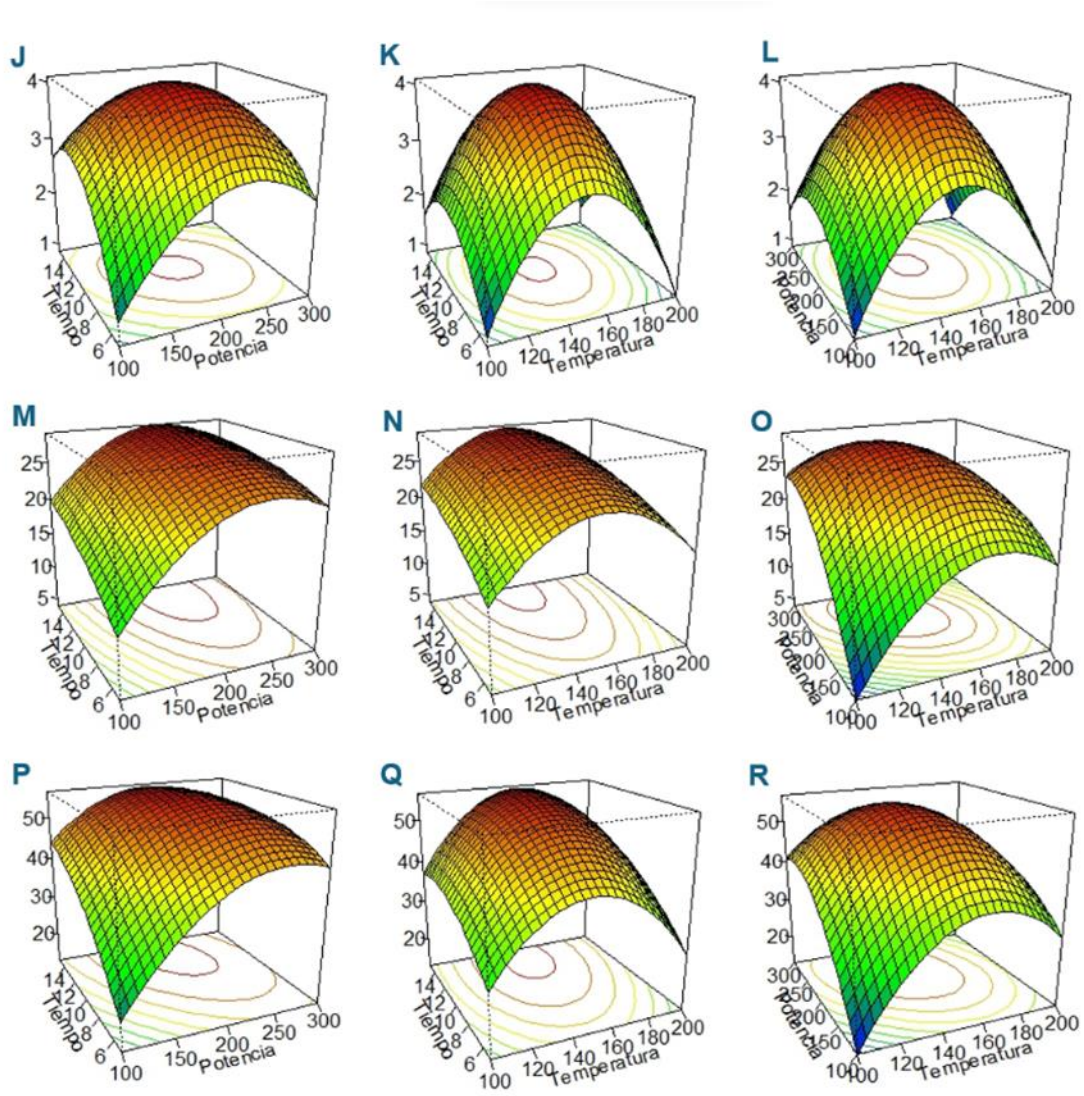


Figure 17 Effect of temperature, power, and time relationship in microwave-assisted hydrothermal pretreatment on total extraction yield of fructose, sucrose and glucose from HMC-CPH: (J) temperature plot, (K) power plot and (L) time plot corresponding to glucose from HMC-CPH; (M) temperature plot, (N) power plot and (O) time plot corresponding to sucrose from HMC-CPH; and (P) temperature plot, (Q) power plot and (R) time plot corresponding to fructose from HMC-CPH.

4.4. Conclusions

The conditions for microwave-assisted hydrothermal pretreatment of cocoa pod husk and cocoa pod husk of hemicellulose were optimized for glucose, sucrose, and fructose and predicted by Box Behnken response surface design. For the optimization of mono- and disaccharides from cocoa pod husk, the expected values of optimum temperature, power, and time were 1) for glucose 135.4 °C, 180.6 W, and 5.8 min, 2) for sucrose 154.3 °C, 256.3 W and 20.2 min and 3) for fructose 129.5 °C, 173.8 W and 5.27 min. For the optimization of mono- and disaccharides from cocoa pod husk of hemicellulose, the predicted optimal temperature, power, and time values were 4) for glucose 142.2°C, 204.4 W, and 10.5 min, 5) for sucrose 148.8°C, 215.6 W and 14.3 min and 6) for fructose 151.6°C, 231.6 W and 13 min. The proposed optimization of microwave-assisted hydrothermal pretreatment for the extraction of glucose, sucrose, and fructose can be a green protocol for utilizing cocoa pod husk.

Acknowledgments

The authors thank the Consejo Nacional de Humanidades, Ciencia y Tecnología (CONAHCYT) for the Postgraduate Grant no. 763652, the Laboratorio de Productos Naturales del Departamento de Química of the Universidad Autónoma Chapingo, the Laboratorios de Geoquímica Marina and LANNBIO of CINVESTAV-IPN, Unidad Mérida, the Instituto Nacional de Investigaciones Forestales Agrícolas y Pecuarias (INIFAP) and C. Carlos Hernández Echeverría for providing the raw material for the study (cocoa pod husk of the Carmelo variety).

4.5. References

Al-Mekhlafi, N. A., Mediani, A., Ismail, N. H., Abas, F., Dymerski, T., Lubinska-Szczygeł, M., Vearasilp, S., & Gorinstein, S. (2021). Metabolomic and antioxidant properties of different varieties and origins of Dragon fruit. *Microchemical Journal*, *160*(105687), 1-12. https://doi.org/https://doi.org/10.1016/j.microc.2020.105687

Aycan, D., Nurdan , Aydın, T., Fulya, & Gamze Turan, N. (2023). Role of different natural materials in reducing nitrogen loss during industrial sludge composting: Modelling and optimization. *Bioresource Technology 385*(129464), 1-10. https://doi.org/https://doi.org/10.1016/j.biortech.2023.129464

Bharti, S. K., & Roy, R. (2012). Quantitative ^{1}H NMR spectroscopy. *TrAC Trends in Analytical Chemistry*, *35*, 5-26. https://doi.org/10.1016/j.trac.2012.02.007

Chan, C. H., Yusoff, R., & Ngoh, G. C. (2014). Optimization of microwave-assisted extraction based on absorbed microwave power and energy. *Chemical Engineering Science*, *111*, 41-47. https://doi.org/10.1016/j.ces.2014.02.011

del Campo, G., Zuriarrain, J., Zuriarrain, A., & Berregi, I. (2016). Quantitative determination of carboxylic acids, amino acids, carbohydrates, ethanol, and hydroxymethylfurfural in honey by (1)H NMR. *Food Chemistry*, *196*, 1031-1039. https://doi.org/10.1016/j.foodchem.2015.10.036

Fan, X., Ren, M., Zhou, C., Kong, F., Hua, C., Fakayode, O. A., Okonkwo, C. E., Li, H., Liang, J., & Wang, X. (2024). Total utilization of lignocellulosic biomass with xylooligosaccharides production priority: A review. *Biomass and Bioenergy*, *181*. https://doi.org/10.1016/j.biombioe.2023.107038

Ferreira, S. L., Bruns, R. E., Ferreira, H. S., Matos, G. D., David, J. M., Brandao, G. C., da Silva, E. G., Portugal, L. A., dos Reis, P. S., Souza, A. S., & dos Santos, W. N. (2007). Box-Behnken design: an alternative for the optimization of analytical methods. *Analytica Chimica Acta*, *597*(2), 179-186. https://doi.org/10.1016/j.aca.2007.07.011

Hennessey Ramos, L., Murillo Arango, W., Vasco Correa, J., & Paz Astudillo, I. C. (2021). Enzymatic Extraction and Characterization of Pectin from Cocoa Pod Husks (*Theobroma cacao* L.) Using Celluclast((R)) 1.5 L. *Molecules*, *26*(5). https://doi.org/10.3390/molecules26051473

Hernández Bolio, G. I., Fagundo Mollineda, A., Caamal Fuentes, E. E., Robledo, D., Freile Pelegrin, Y., & Hernández Núñez, E. (2021). NMR Metabolic Profiling of Sargassum Species Under Different Stabilization/Extraction Processes. *Journal of Phycology*, *57(2)*, 655–663. https://doi.org/doi: 10.1111/jpy.13117

Hou, Z., Liang, X., Du, L., Su, F., & Su, W. (2014). Quantitative determination and validation of avermectin B1a in commercial products using quantitative

nuclear magnetic resonance spectroscopy. *Magn. Reson. Chem.*, *52*, 480-485. https://doi.org/DOI 10.1002/mrc.4098

Jarrín Chacón, J. P., Núñez Pérez, J., Espín Valladares, R. D. C., Manosalvas Quiroz, L. A., Rodríguez Cabrera, H. M., & Pais Chanfrau, J. M. (2023). Pectin Extraction from Residues of the Cocoa Fruit (*Theobroma cacao* L.) by Different Organic Acids: A Comparative Study. *Foods*, *12*(3). https://doi.org/10.3390/foods12030590

Lu, F., Rodriguez Garcia, J., Van Damme, I., Westwood, N. J., Shaw, L., Robinson, J. S., Warren, G., Chatzifragkou, A., McQueen Mason, S., Gomez, L., Faas, L., Balcombe, K., Srinivasan, C., Picchioni, F., Hadley, P., & Charalampopoulos, D. (2018). Valorization strategies for cocoa pod husk and its fractions. *Current Opinion in Green and Sustainable Chemistry*, *14*, 80-88. https://doi.org/10.1016/j.cogsc.2018.07.007

Muharja, M., Darmayanti, R. F., Palupi, B., Rahmawati, I., Fachri, B. A., Setiawan, F. A., Amini, H. W., Rizkiana, M. F., Rahmawati, A., Susanti, A., & Putri, D. K. Y. (2021). Optimization of Microwave-Assisted Alkali Pretreatment for Enhancement of Delignification Process of Cocoa Pod Husk. *Bulletin of Chemical Reaction Engineering & Catalysis*, *16*(1), 31-43. https://doi.org/10.9767/bcrec.16.1.8872.31-43

Muñoz Almagro, N., Valadez Carmona, L., Mendiola, J. A., Ibanez, E., & Villamiel, M. (2019). Structural characterization of pectin obtained from cacao pod husk. Comparison of conventional and subcritical water extraction. *Carbohydr Polym*, *217*, 69-78. https://doi.org/10.1016/j.carbpol.2019.04.040

Nassef, H. M., Al Hazmi, G. A. A. M., Alayyafi, A. A., El Desouky, M. G., & El Bindary, A. A. (2024). Synthesis and characterization of new composite sponge combining of metal-organic framework and chitosan for the elimination of Pb (II), Cu (II), and Cd (II) ions from aqueous solutions: Batch adsorption and optimization using Box-Behnken design. *Journal of Molecular Liquids*, *394*. https://doi.org/10.1016/j.molliq.2023.123741

Ouattara, L. Y., Appiah Kouassi, E. K., Soro, D., Kouassi, B. Y., Didier Fanou, G., Drogui, A. P., & Rajeshwar Tyagi, D. (2023). Optimization of thermochemical hydrolysis of potassium hydroxide-delignified Cocoa (*Theobroma cacao* L) pod husks under low combined severity factors (CSF) conditions. *Scientific African*, *22*(e01908), 1-14. https://doi.org/https://doi.org/10.1016/j.sciaf.2023.e01908

Pangestu, R., Amanah, S., Juanssilfero, A. B., Yopi, & Perwitasari, U. (2020). Response surface methodology for microwave-assisted extraction of pectin from cocoa pod husk (*Theobroma cacao*) mediated by oxalic acid. *Journal of Food Measurement and Characterization*, *14*, 2126–2133. https://doi.org/https://doi.org/10.1007/s11694-020-00459-4

Peng, P., & She, D. (2014). Isolation, structural characterization, and potential applications of hemicelluloses from bamboo: A review. *Carbohydrate Polymers*, *112*, 701-720. https://doi.org/10.1016/j.carbpol.2014.06.068

Perwitasari, U., Agustina, N. T., Pangestu, R., Amanah, S., Saputra, H., Andriani, A., Fahrurrozi, Juanssilfero, A. B., Thontowi, A., Widyaningsih, T. D., Eris, D. D., Amaniyah, M., Yopi, & Habibi, M. S. (2021). Using the response surface method, Cacao pod husk is used for citric acid production under solid-state fermentation. *Biomass Conversion and Biorefinery*, *13*(8), 7165-7173. https://doi.org/10.1007/s13399-021-01690-9

Priyangini, F., Walde, S. G., & Chidambaram, R. (2018). Extraction optimization of pectin from cocoa pod husks (*Theobroma cacao* L.) with ascorbic acid using response surface methodology. *Carbohydr Polym*, *202*, 497-503. https://doi.org/10.1016/j.carbpol.2018.08.103

Quiceno Suarez, A., Cadena Chamorro, E. M., Ciro Velásquez, H. J., & Arango Tobón, J. C. (2024). By-products of the cocoa agribusiness: high value-added materials based on their bromatological and chemical characterization. *Rev. Facultad Nacional de Agronomia*, *77(1)*, 10585-10599. https://doi.org/https://doi.org/10.15446/rfnam.v77n1.107738

Rahandi Lubis, M. A., Seng Hua, L., Efri, M., Souvia, R., Petar, A., Robi, A., Luboš, K. á., & Bambang, N. (2024). *Biomass Conversion and Sustainable Biorefinery* (M. A. R. L. S. H. L. E. M. S. R. P. A. R. A. L. u. K. B. Nurhadi, Ed.). Springer Nature. https://doi.org/10.1007/978-981-99-7769-7

Ramos Escudero, F., Casimiro Gonzales, S., Cadiz Gurrea, M. L., Cancino Chavez, K., Basilio Atencio, J., Ordonez, E. S., Munoz, A. M., & Segura Carretero, A. (2023). Optimizing vacuum drying process of polyphenols, flavanols, and DPPH radical scavenging assay in pod husk and bean shell cocoa. *Sci Rep*, *13*(1), 13900. https://doi.org/10.1038/s41598-023-40815-0

Rogoski, W., Pereira, G. N., Cesca, K., de Oliveira, D., & de Andrade, C. J. (2023). An Overview on Pretreatments for the Production of Cassava Peels-based Xyloligosaccharides: State of Art and Challenges. *Waste and Biomass Valorization*, *14*(7), 2115-2131. https://doi.org/10.1007/s12649-023-02044-4

Saelee, M., Sundaram Sivamaruthi, B., Tansrisook, C., Duangsri, S., Chaiyasut, K., Kesika, P., Peerajan, S., & Chaiyasut, C. (2022). Response Surface Methodological Approach for Optimizing *Theobroma cacao* L. Oil Extraction. *Applied Sciences*, *12* (5482), 1-15. https://doi.org/ https://doi.org/10.3390/app12115482

Schievano, E., Tonoli, M., & Rastrelli, F. (2017). NMR Quantification of Carbohydrates in Complex Mixtures. A Challenge on Honey. *Anal Chem*, *89*(24), 13405-13414. https://doi.org/10.1021/acs.analchem.7b03656

Soares, T. F., & Oliveira, M. (2022). Cocoa By-Products: Characterization of Bioactive Compounds and Beneficial Health Effects. *Molecules*, *27*(5). https://doi.org/10.3390/molecules27051625

Spiteri, M., Jamin, E., Thomas, F., Rebours, A., Lees, M., Rogers, K. M., & Rutledge, D. N. (2015). Fast and global authenticity screening of honey using ^{1}H-NMR profiling. *Food Chemistry*, *189*, 60-66. https://doi.org/http://dx.doi.org/10.1016/j.foodchem.2014.11.099

Valladares Diestra, K. K., Porto de Souza Vandenberghe, L., & Soccol, C. R. (2022). A biorefinery approach for pectin extraction and second-generation bioethanol production from cocoa pod husk. *Bioresour Technol*, *346*, 126635. https://doi.org/10.1016/j.biortech.2021.126635

Vasquez, Z. S., de Carvalho Neto, D. P., Pereira, G. V. M., Vandenberghe, L. P. S., de Oliveira, P. Z., Tiburcio, P. B., Rogez, H. L. G., Goes Neto, A., & Soccol, C. R. (2019). Biotechnological approaches for cocoa waste management: A review. Waste Manag, 90, 72-83. https://doi.org/10.1016/j.wasman.2019.04.030

5. GENERAL CONCLUSION

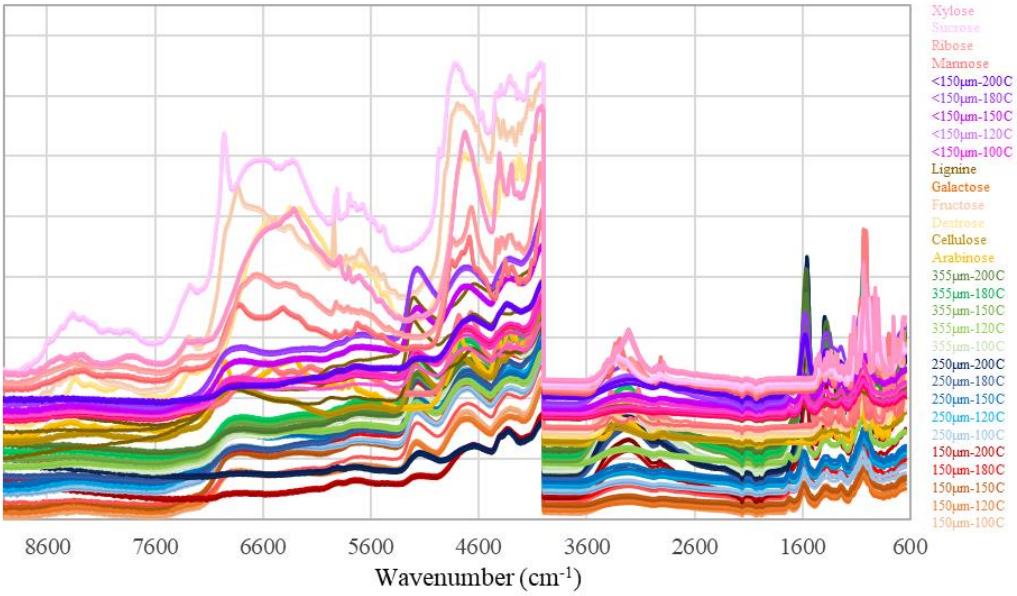
A vital information contribution was made by the characterization of the cocoa pod husk of the *Theobroma cacao* L. variety Carmelo concerning the content of major chemical constituents, its antioxidant capacity, and mineral profile. No significance was found compared with data reported in the literature on cocoa shells of other varieties. It was found that CPH could be a potential source of lignin, equivalent to woods with higher lignin production. Potential carbohydrate content was determined for obtaining biomolecules that could be used to obtain biofuels, as other lignocellulosic sources; however, the high polyphenols and its antioxidant capacity make it a potential material mainly for the food and pharmaceutical industry.

The present study provided the opportunity to apply a new method for discriminating between simple and complex cellulose, hemicellulose, and lignin carbohydrates and partial degradation molecules by supervised and unsupervised chemometric methods such as Principal Component Analysis (PCA), Partial Least Squares Discriminant Analysis (PLS-DA), reduced partial (SPLS-DA) and orthogonal (OPLS-DA) with the first derivative absorbance matrix of MIR-NIR NMR.

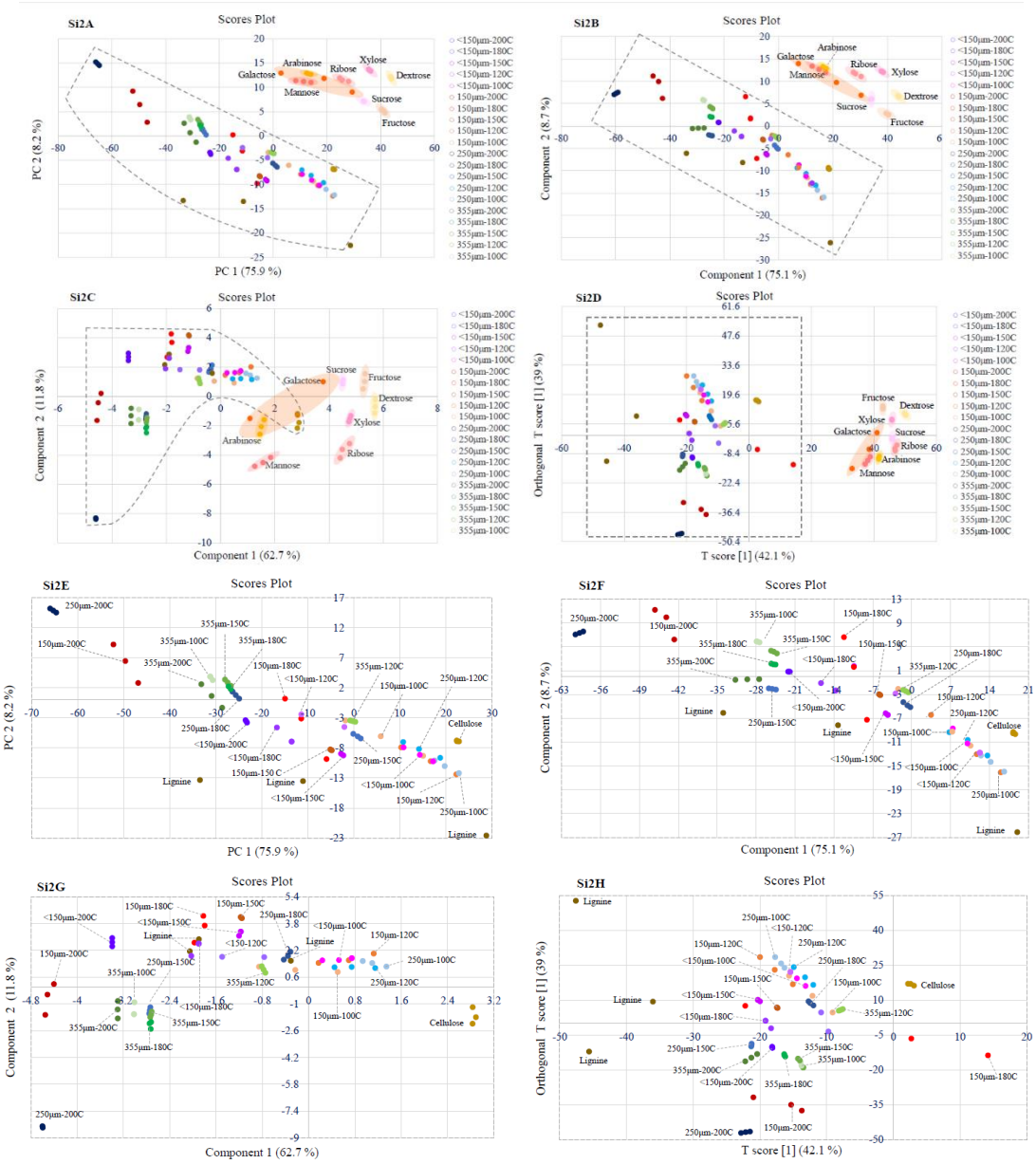
Likewise, the work provided an innovative method with NMR-driven analysis by rigorous matching between HSQC, TOCSY, and J-resolved entries concerning the Human Metabolome Database (HMDB), identifying at least 20 discriminating metabolites.

Future studies could use the present methodology to evaluate other dependencies of microwave-assisted hydrothermal extraction of cocoa pod husks, such as microwave frequencies, pressure dependence, other ramps, temperature increments, and smaller screening processes.

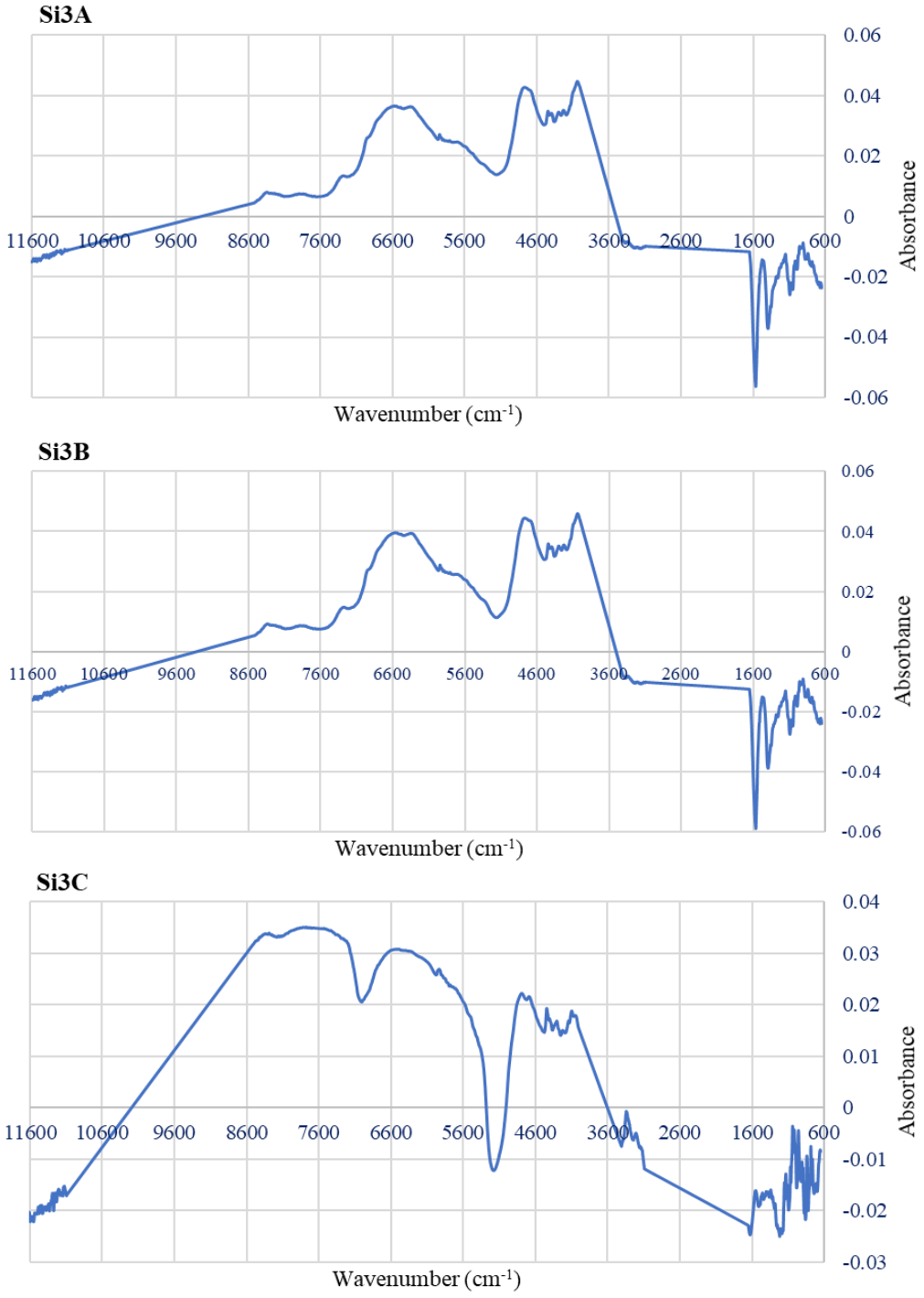
6. APPENDIXES



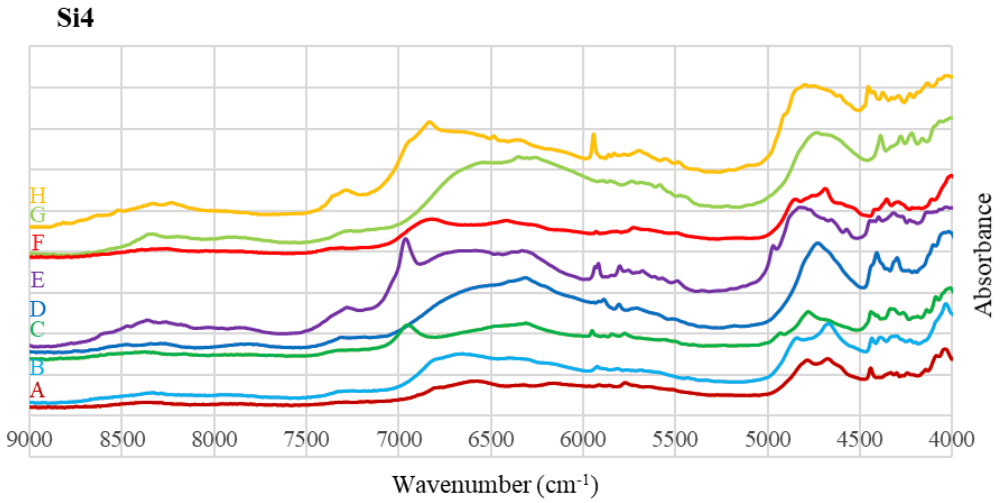
Appendix 2 Stacked NIR-MIR spectra of cocoa pod husk samples after applying hydrothermal pretreatments, pentose, hexose, and sucrose standards, and lignin and cellulose polymers.



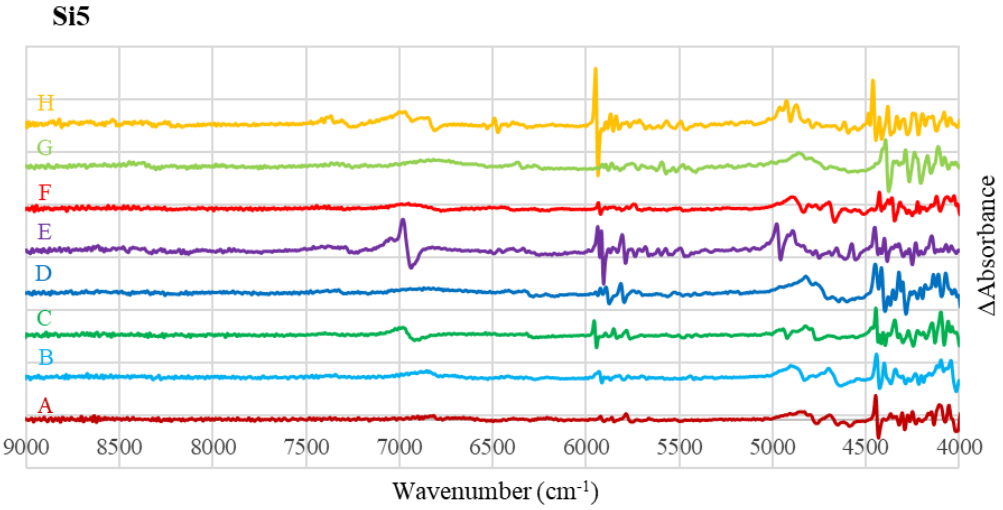
Appendix 3 Stacked NIR-MIR spectra of cocoa pod husk samples after applying hydrothermal pretreatments, pentose, hexose and sucrose standards and lignin and cellulose polymers; 2A) score plots of the unsupervised principal component analysis (PCA), 2B) supervised standard partial least squares (PLS-DA), 2D) scattered partial least squares (sPLS-DA) and 2C) orthogonal projections to latent structures discriminant analysis (OPLS-DA).



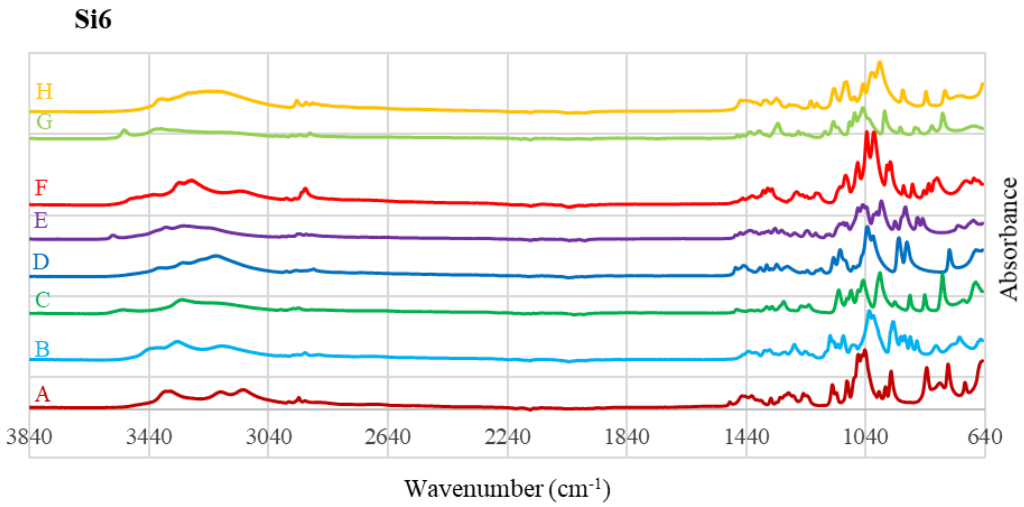
Appendix 4 MIR-NIR loading plots comprising from Principal component analysis (Si3A), partial least squares discriminant analysis (Si3B) and Orthogonal PLS-DA permutation (Si3C).



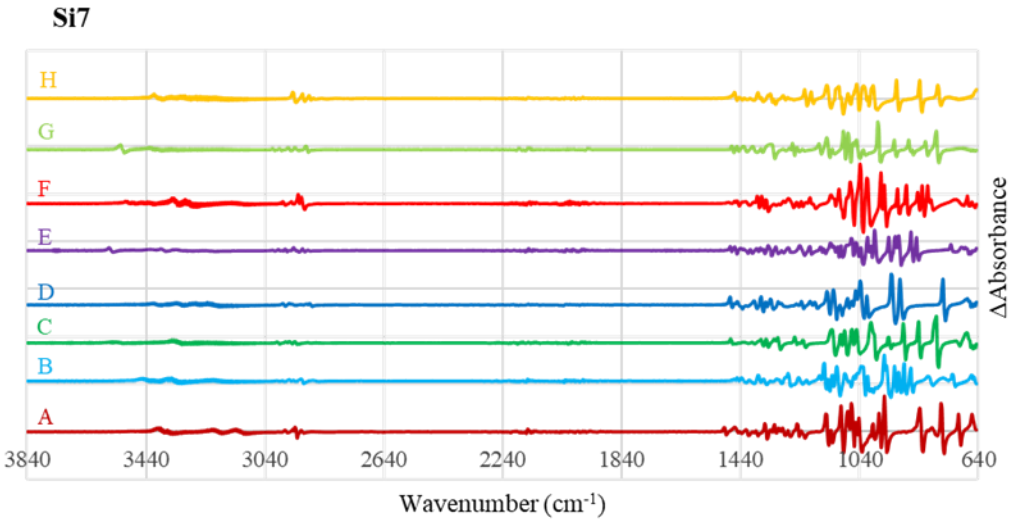
Appendix 5 Stacked NIR spectra of galactose (A), ribose (B), arabinose (C), xylose (D), sucrose €, mannose (F), fructose (G) and dextrose (H) standards.



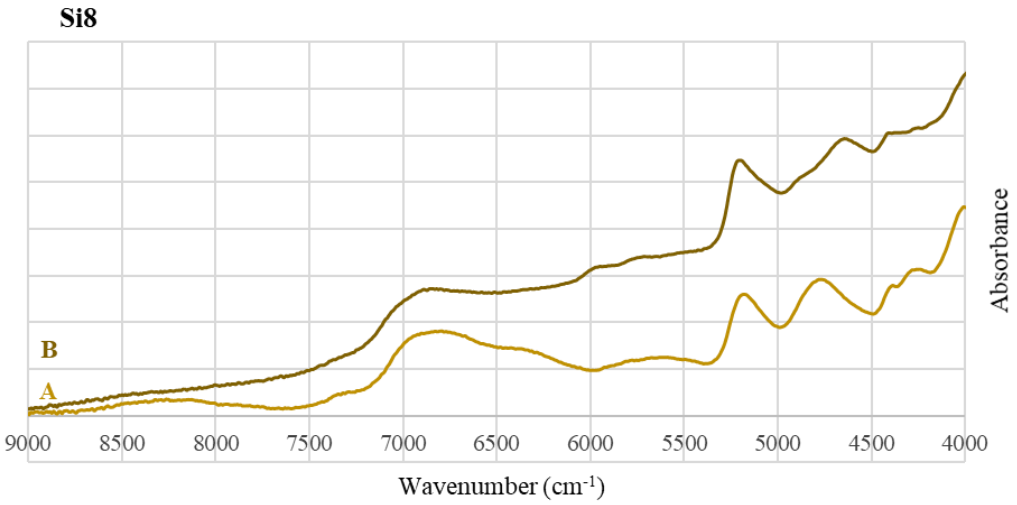
Appendix 6 Stacked NIR spectra of galactose (A), ribose (B), arabinose (C), xylose (D), sucrose €, mannose (F), fructose (G), and dextrose (H) standards after applying preprocessing.



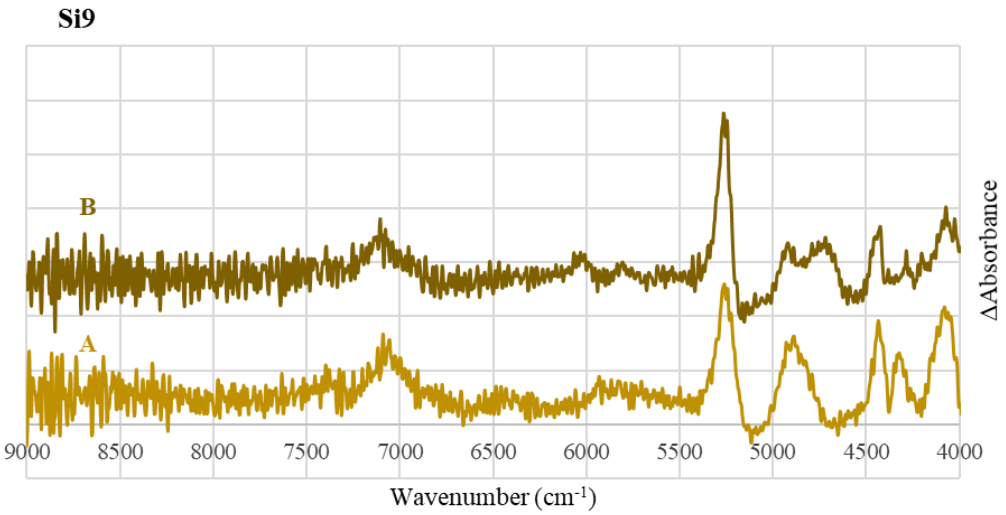
Appendix 7 Stacked MIR spectra of galactose (A), ribose (B), arabinose (C), xylose (D), sucrose (E), mannose (F), fructose (G) and dextrose (H) standards.



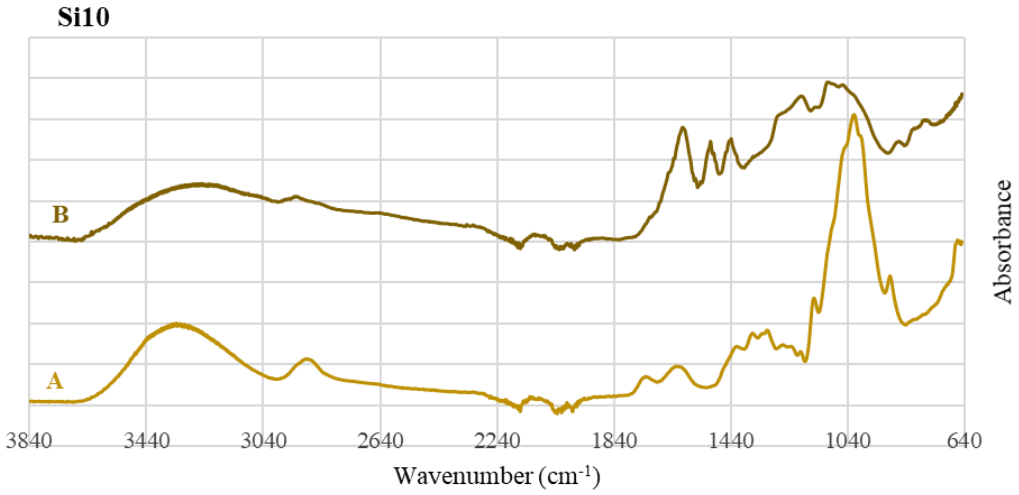
Appendix 8 Stacked MIR spectra of galactose (A), ribose (B), arabinose (C), xylose (D), sucrose (E), mannose (F), fructose (G), and dextrose (H) standards after applying preprocessing.



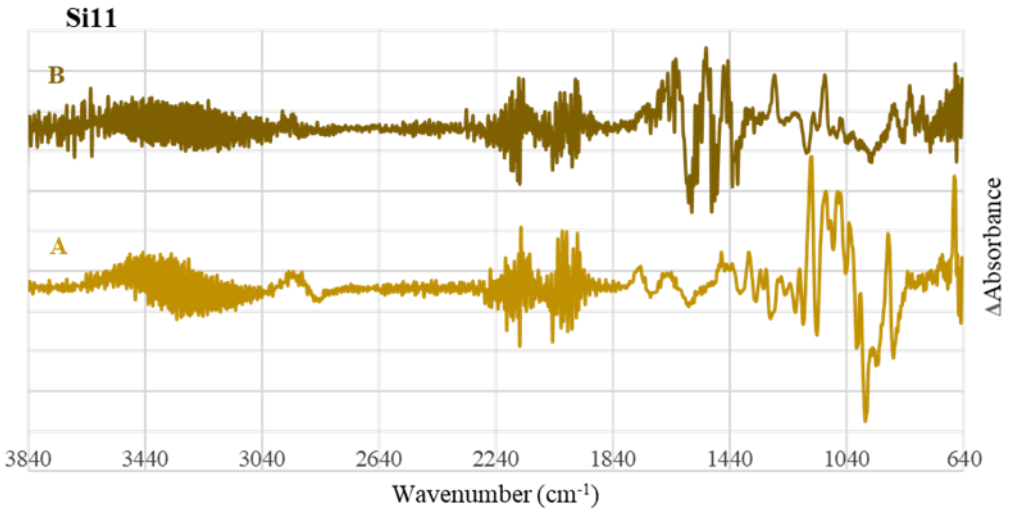
Appendix 9 Stacked NIR cellulose (A) and lignin (B) polymers.



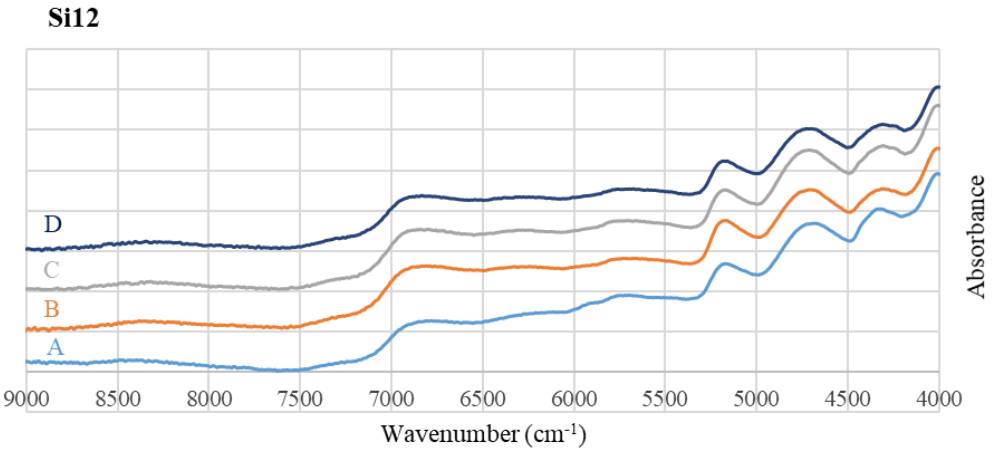
Appendix 10 Stacked NIR cellulose (A) and lignin (B) polymers after applying preprocessing.



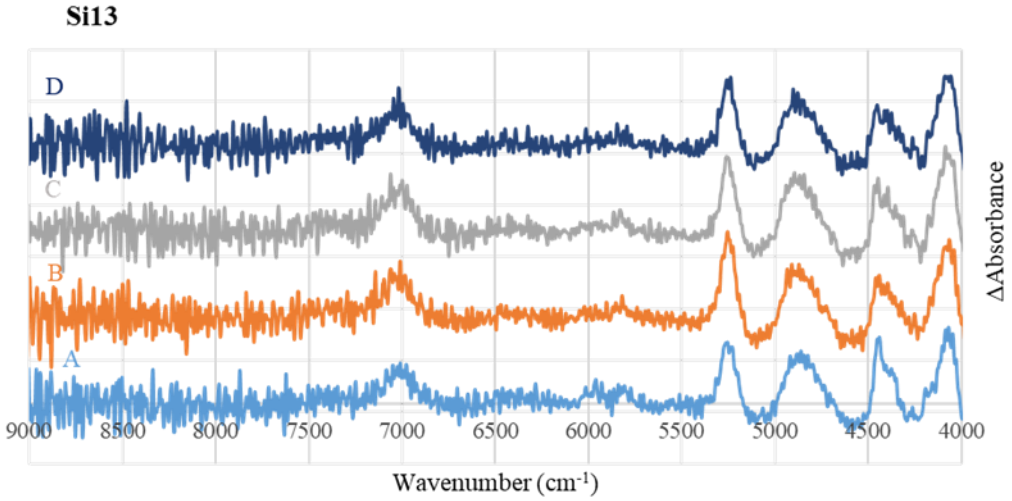
Appendix 11 Stacked MIR cellulose (A) and lignin (B) polymers.



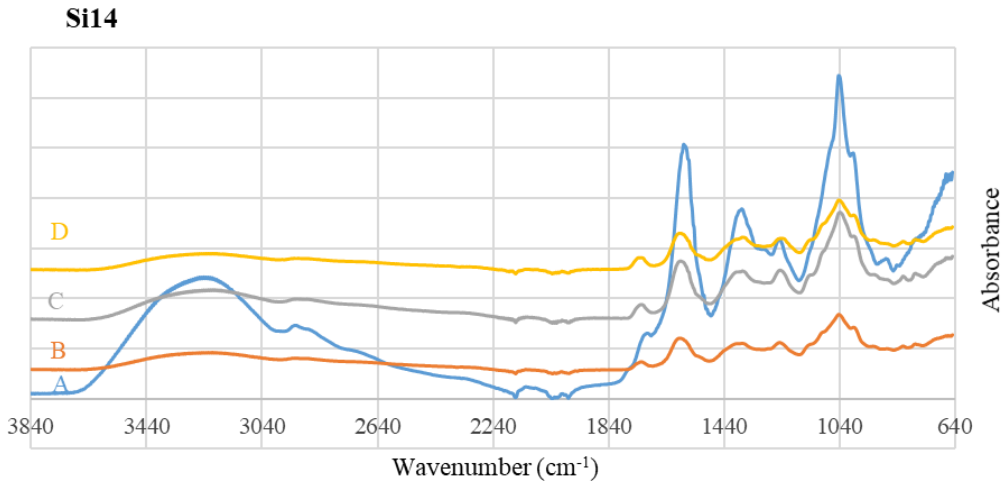
Appendix 12 Stacked MIR cellulose (A) and lignin (B) polymers after applying preprocessing.



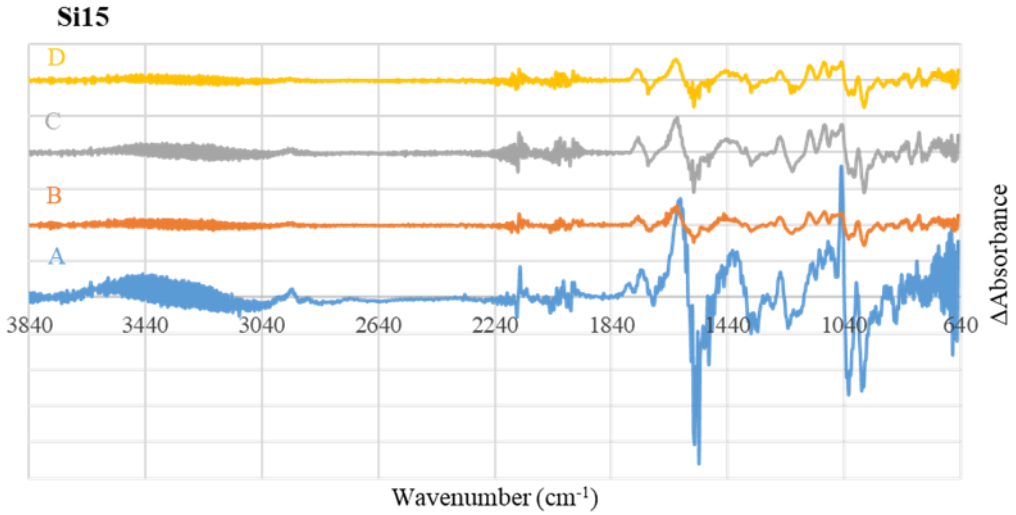
Appendix 13 Stacked NIR 355µm (A), 250 µm (B), 150 µm (C), and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 100°C.



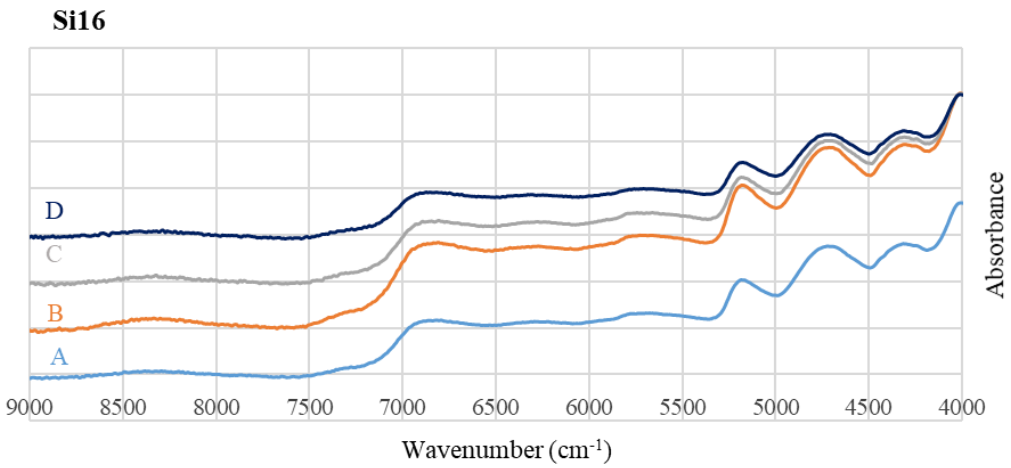
Appendix 14 Stacked NIR 355µm (A), 250 µm (B), 150 µm (C), and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 100°C after preprocessing.



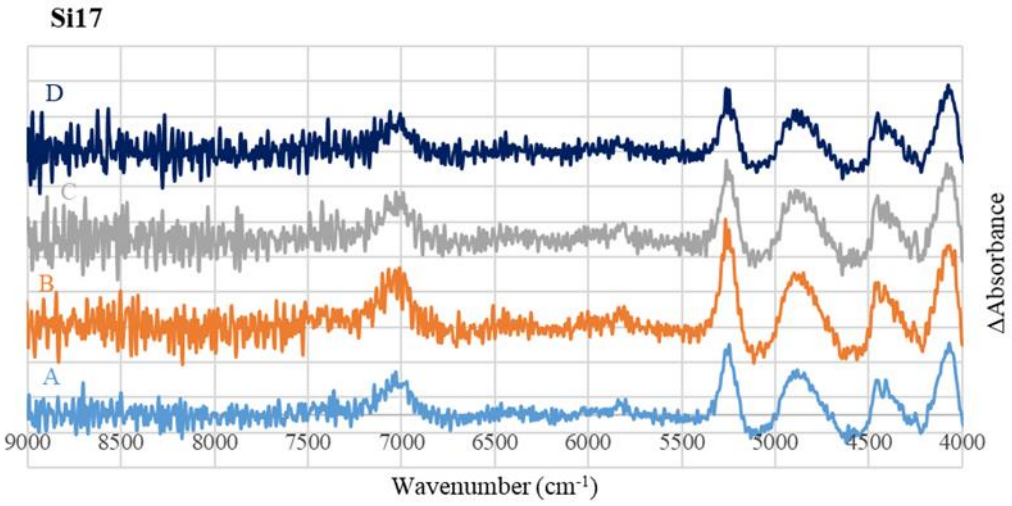
Appendix 16 Stacked MIR 355µm (A), 250 µm (B), 150 µm (C) and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 100°C.



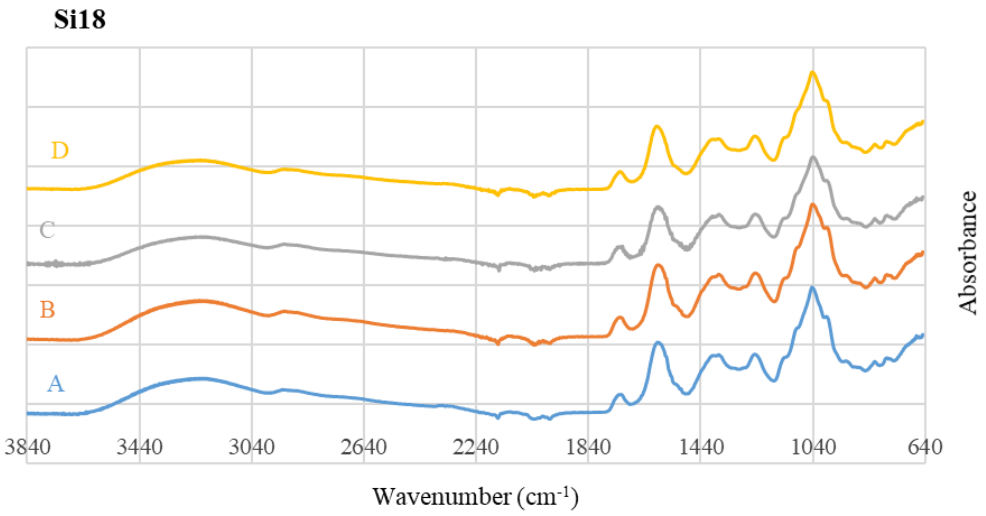
Appendix 15 Stacked MIR 355µm (A), 250 µm (B), 150 µm (C) and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 100°C after preprocessing.



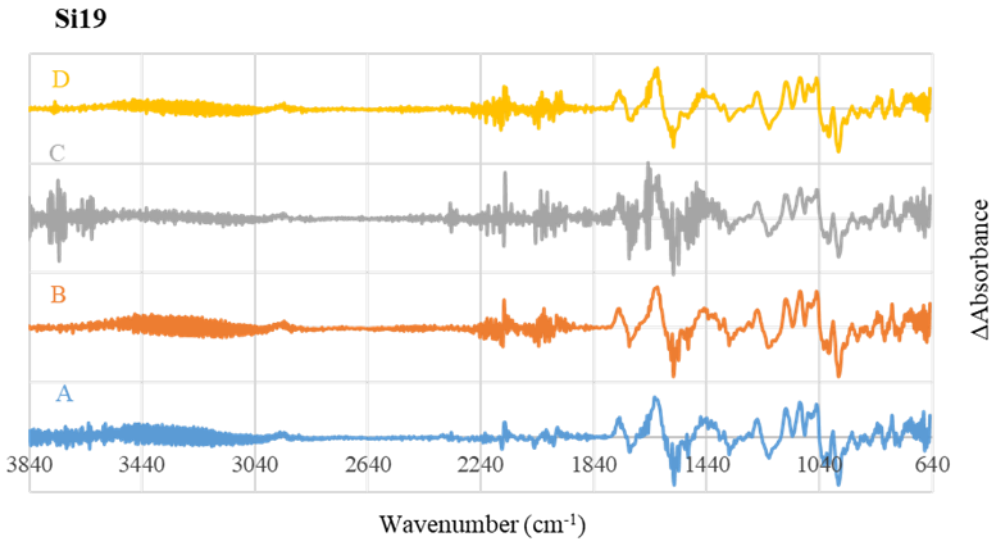
Appendix 17 Stacked NIR 355µm (A), 250 µm (B), 150 µm (C), and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 120°C.



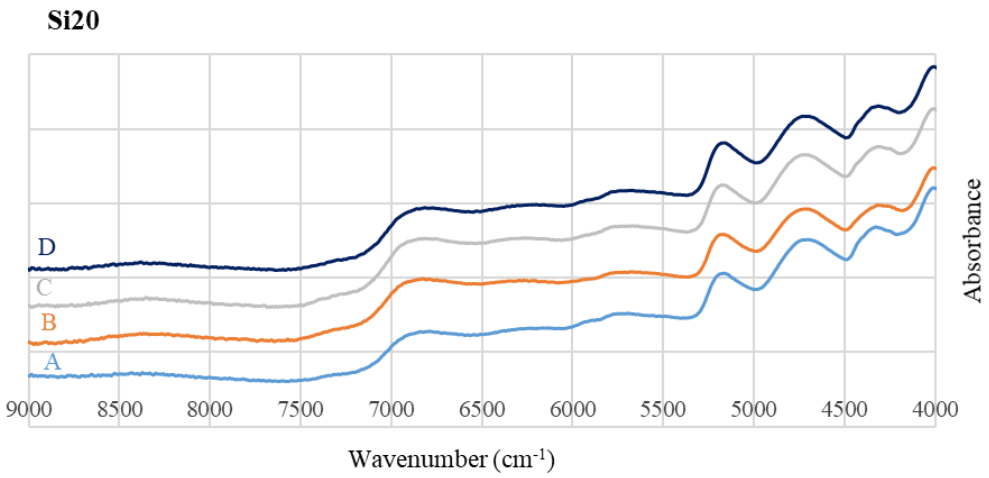
Appendix 18 Stacked NIR 355µm (A), 250 µm (B), 150 µm (C) and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 120°C after preprocessing.



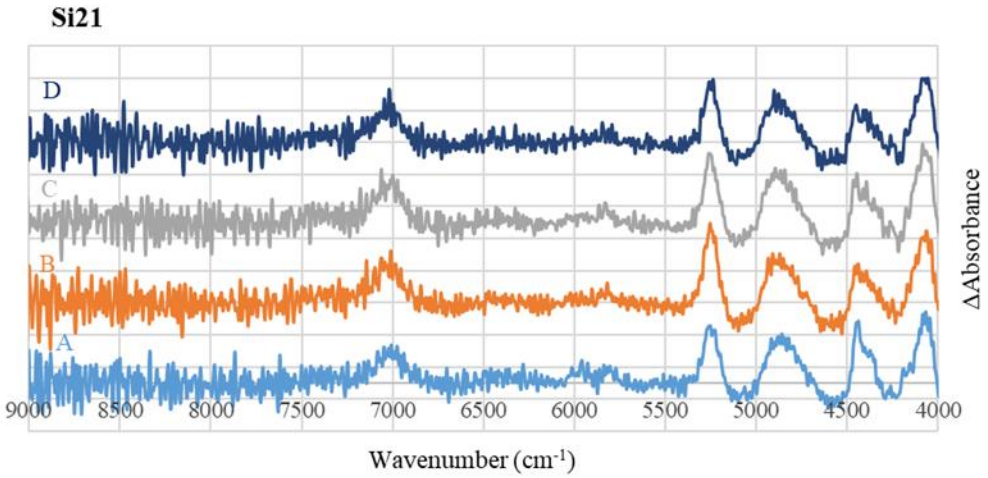
Appendix 19 Stacked MIR 355µm (A), 250 µm (B), 150 µm (C), and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pre-treatments at 120°C after preprocessing.



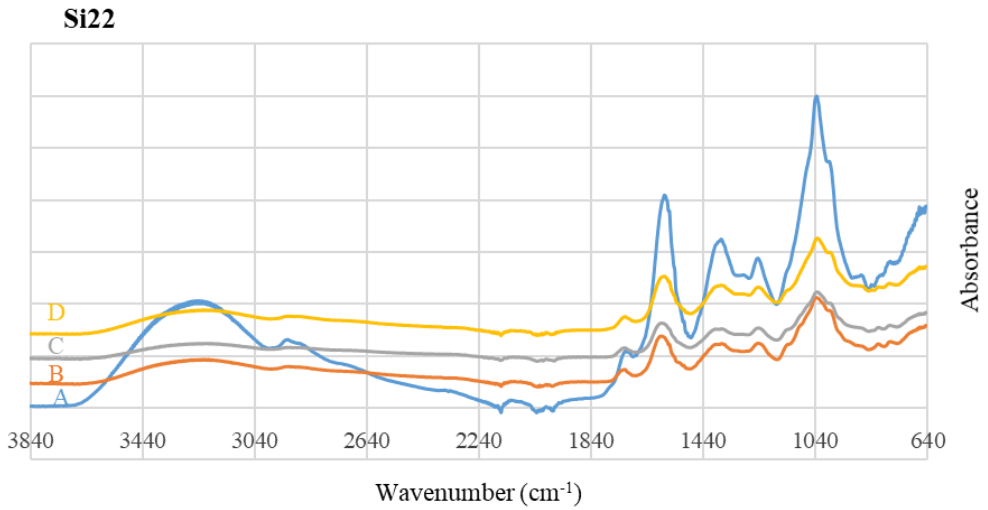
Appendix 20 Stacked MIR 355µm (A), 250 µm (B), 150 µm (C), and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 120°C.



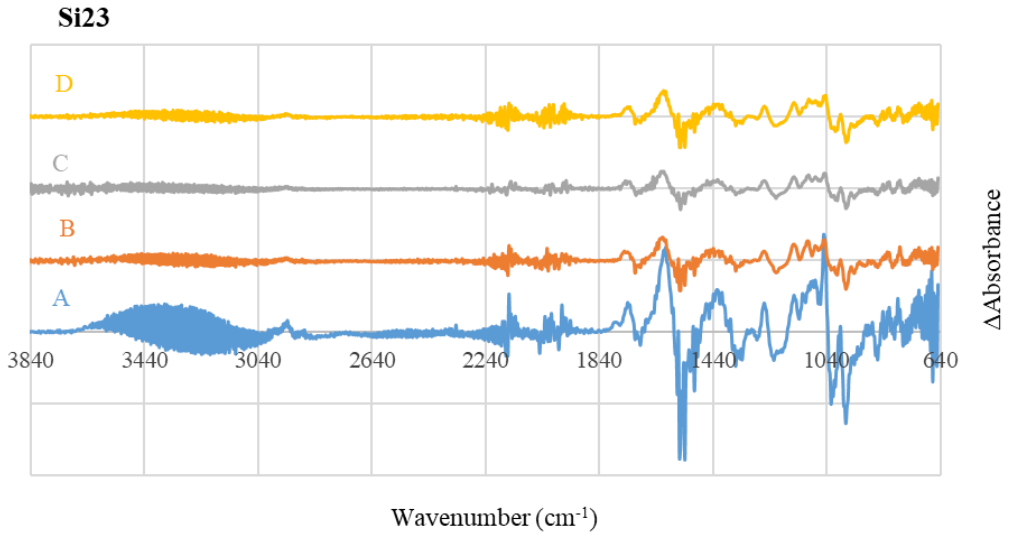
Appendix 21 Stacked NIR 355µm (A), 250 µm (B), 150 µm (C), and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 150°C.



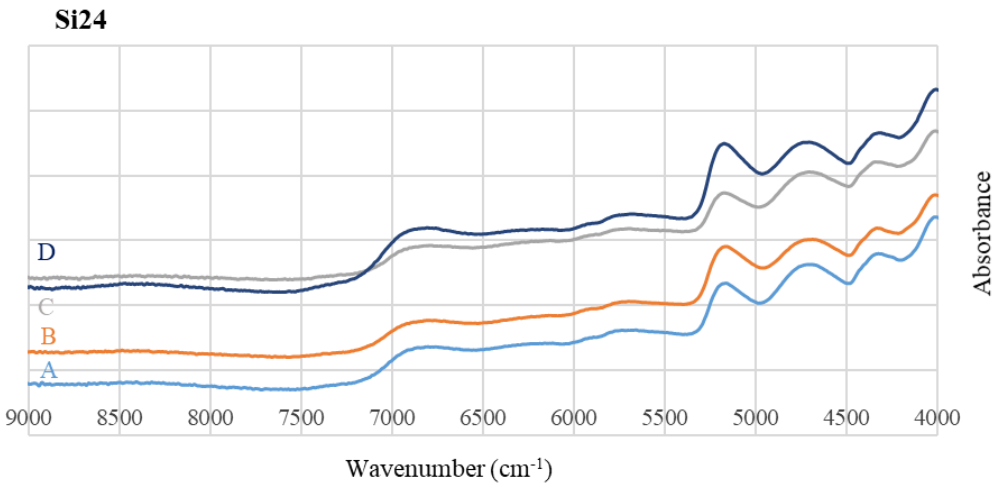
Appendix 22 Stacked NIR 355µm (A), 250 µm (B), 150 µm (C), and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 150°C after pre-processing.



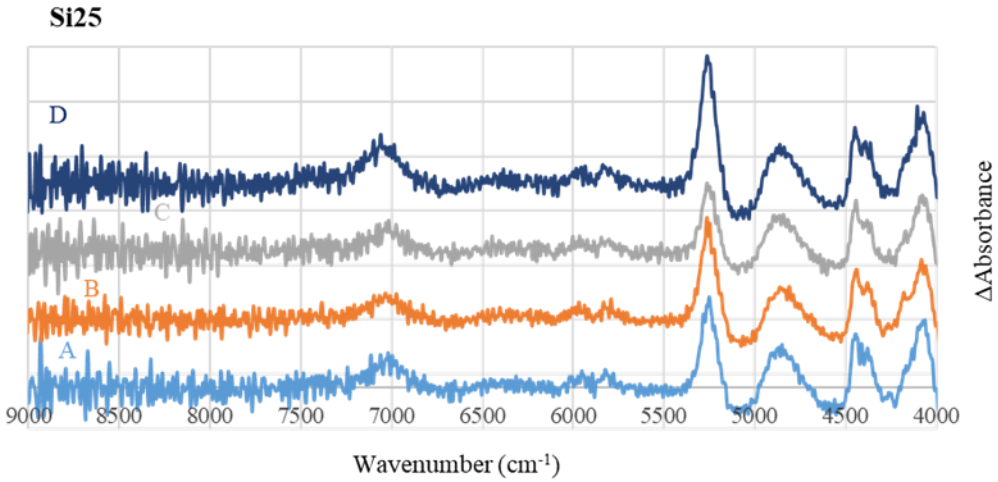
Appendix 23 Figure 38 Stacked MIR 355µm (A), 250 µm (B), 150 µm (C), and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 150°C.



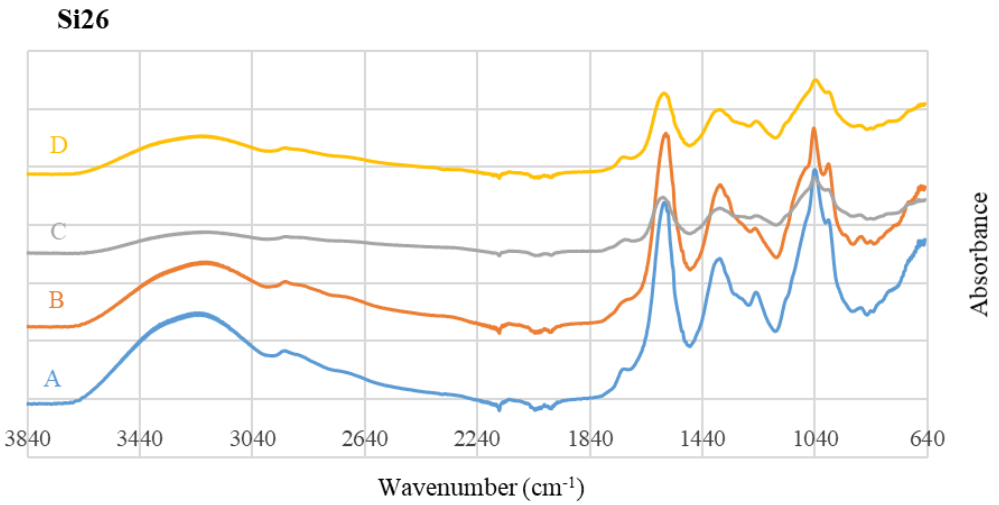
Appendix 24 Stacked MIR 355µm (A), 250 µm (B), 150 µm (C) and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 150°C after preprocessing.



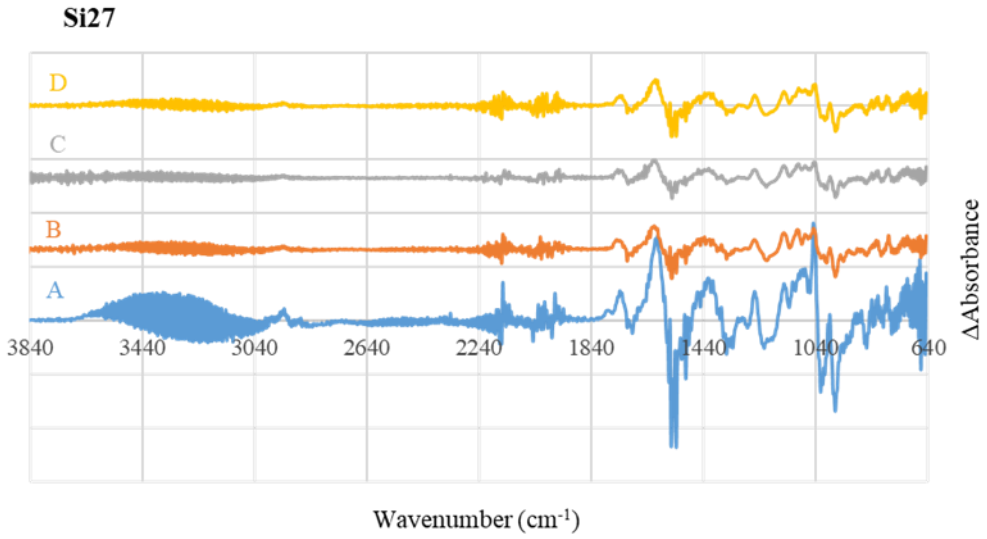
Appendix 25 Stacked NIR 355µm (A), 250 µm (B), 150 µm (C), and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 180°C after preprocessing.



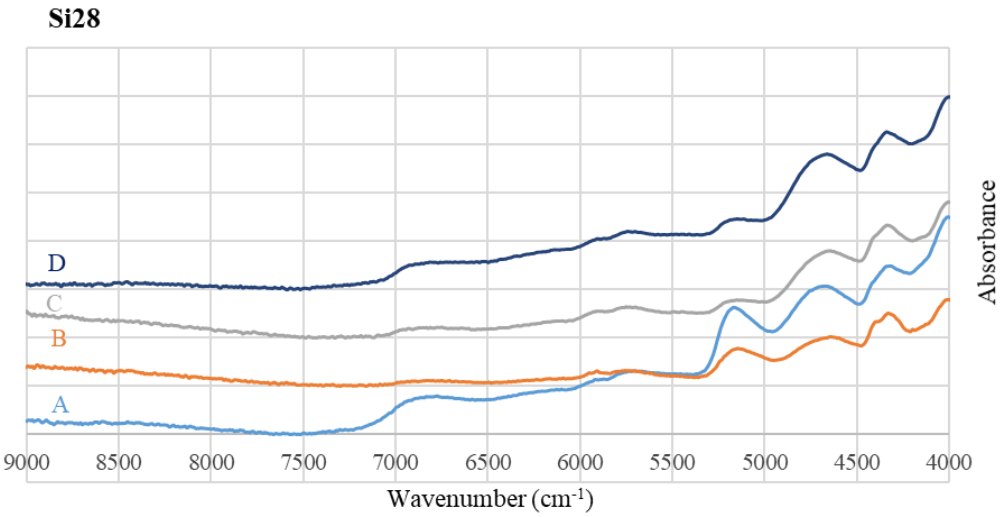
Appendix 26 Stacked NIR 355µm (A), 250 µm (B), 150 µm (C), and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 180°C.



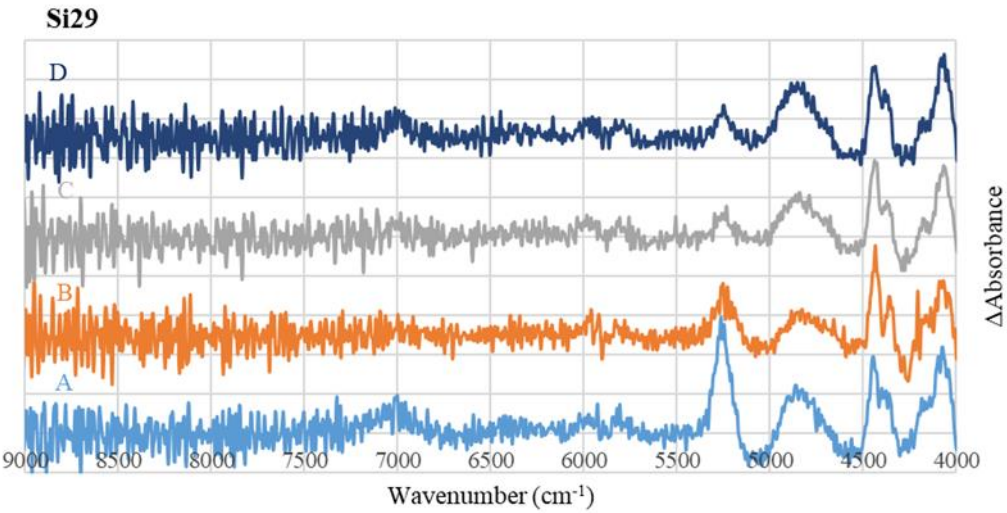
Appendix 27 Stacked MIR 355µm (A), 250 µm (B), 150 µm (C) and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 180°C.



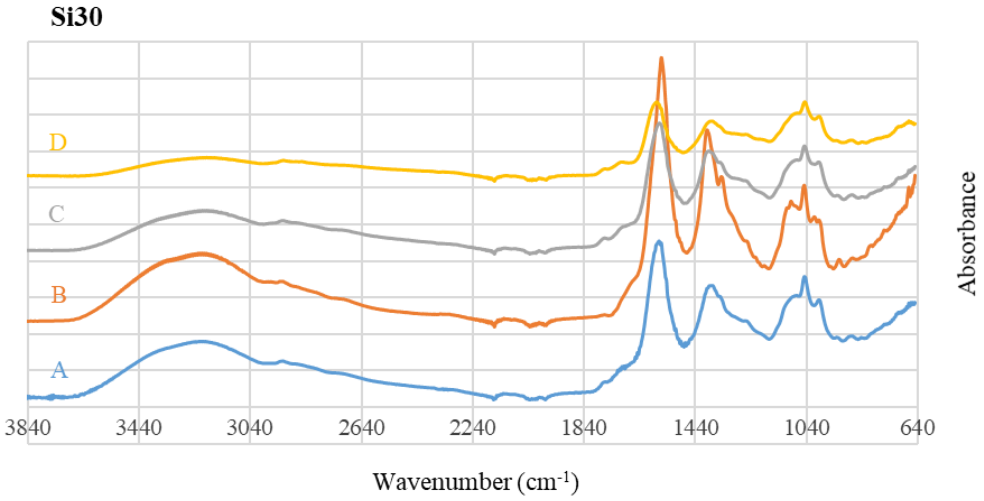
Appendix 28 Stacked MIR355µm (A), 250 µm (B), 150 µm (C), and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 180°C after preprocessing.



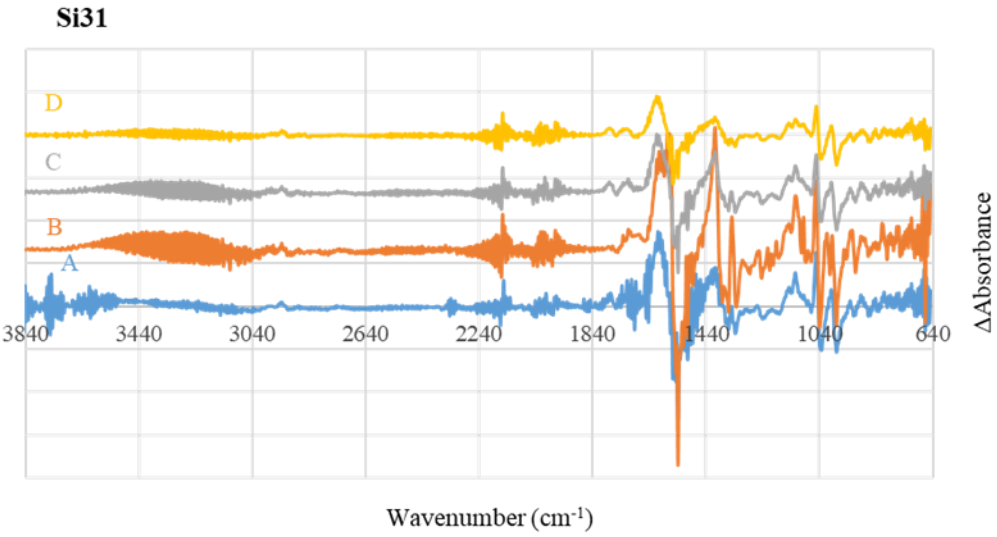
Appendix 29 Stacked NIR 355µm (A), 250 µm (B), 150 µm (C), and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 200°C.



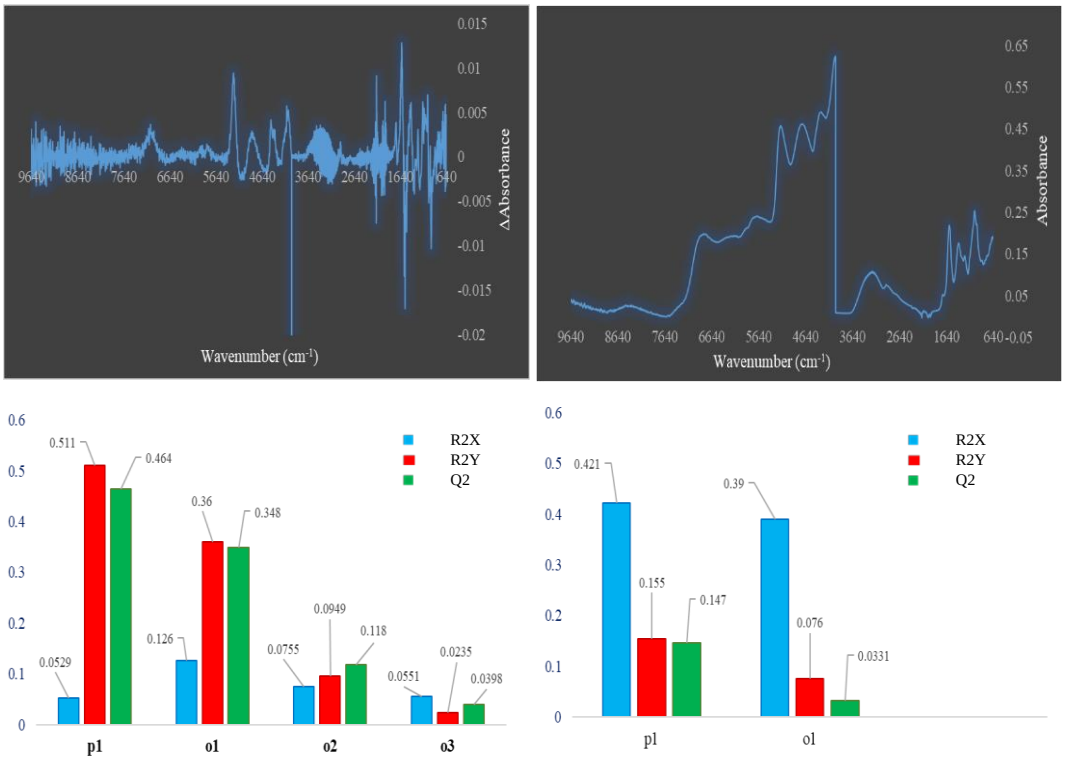
Appendix 30 Stacked NIR 355µm (A), 250 µm (B), 150 µm (C), and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 200°C after preprocessing.



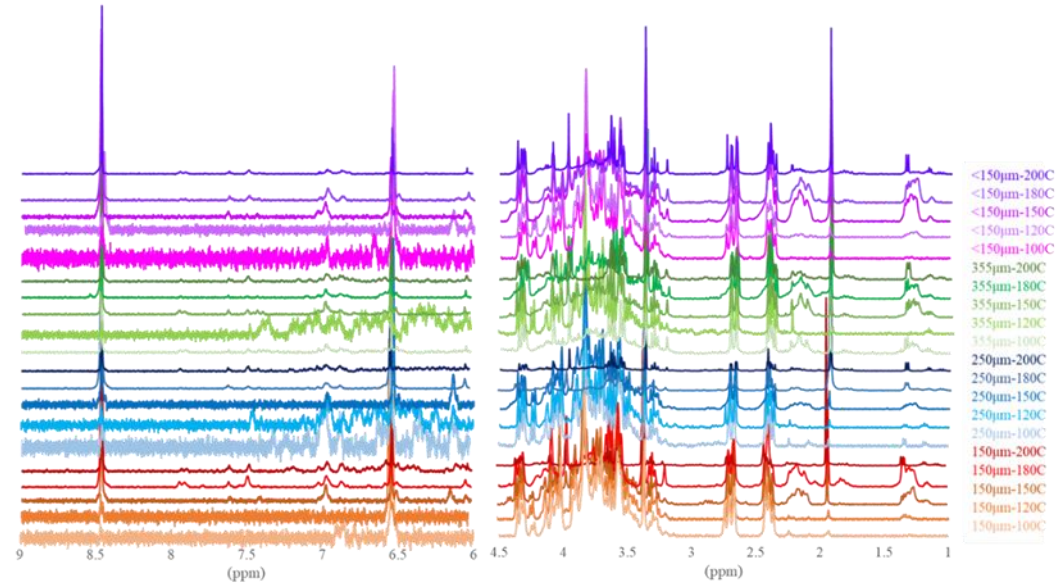
Appendix 31 Stacked MIR 355µm (A), 250 µm (B), 150 µm (C), and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 200°C.



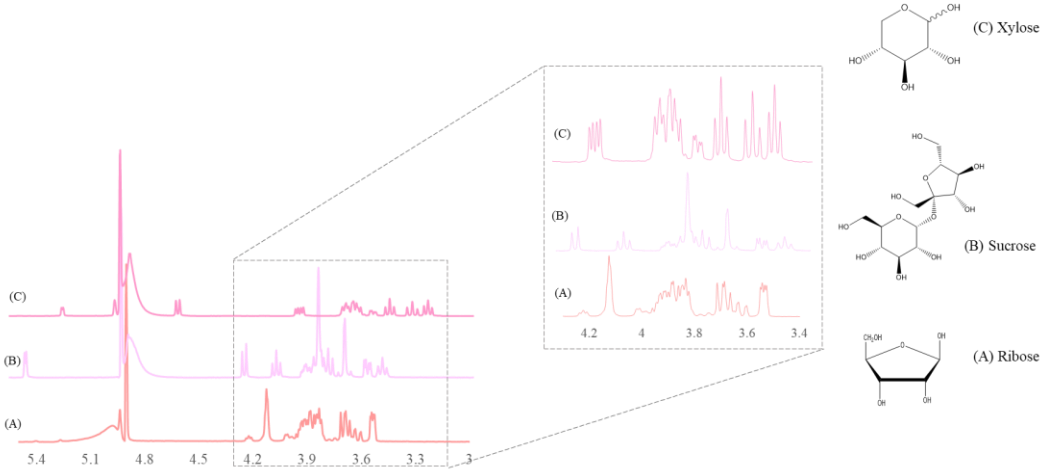
Appendix 32 Stacked MIR 355µm (A), 250 µm (B), 150 µm (C), and <150 µm (D) spectra of cocoa pod husk samples after hydrothermal pretreatments at 200°C after preprocessing.



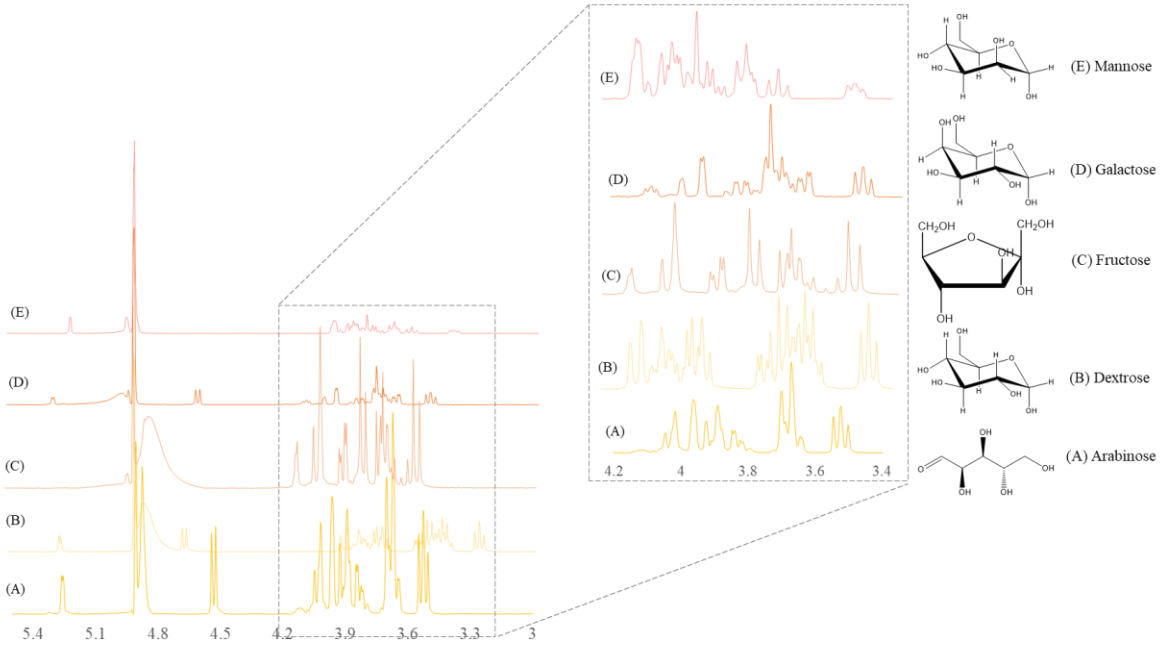
Appendix 33 The permutation analysis between one predictive (p1) and three orthogonal (o1, o2, and o3) components produced the observed and cross-validated R2X, R2Y, and Q2 coefficients.



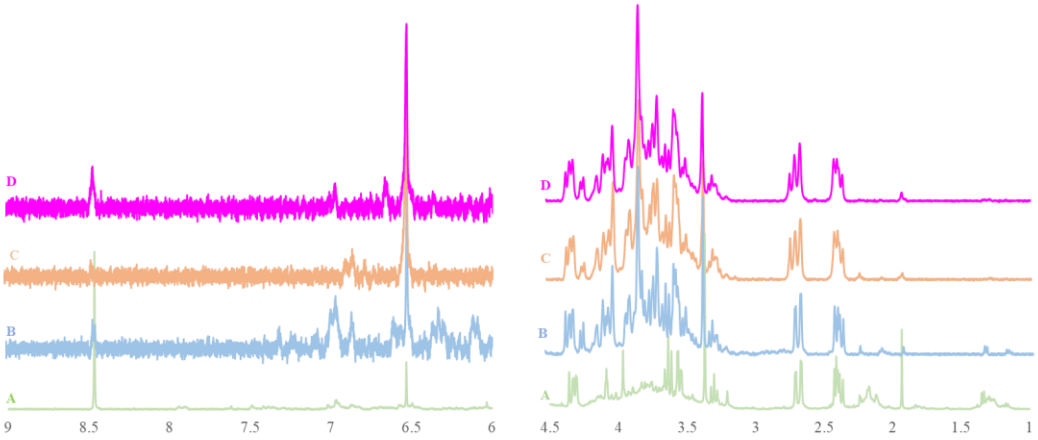
Appendix 34 Stacked one-dimensional ^{1}H water_presat NMR spectra of cocoa pod husk samples after applying hydrothermal pretreatments.



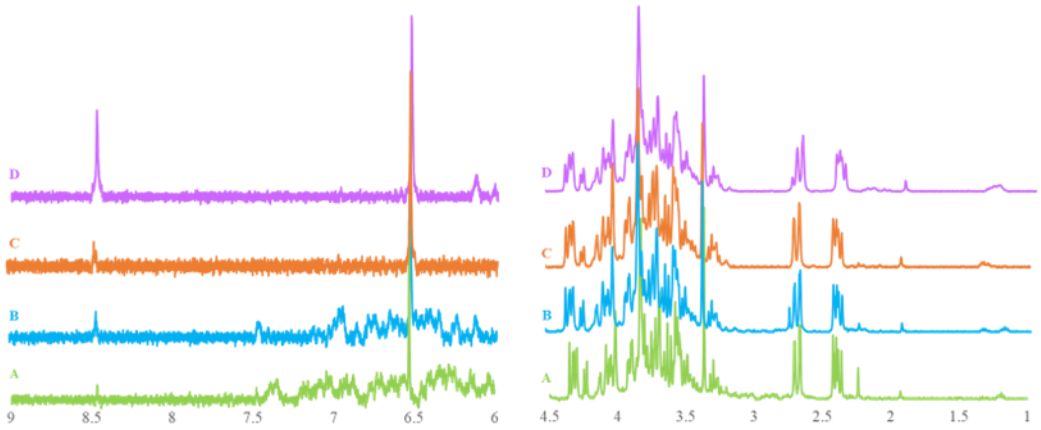
Appendix 35 Stacked one dimensional ^{1}H water_presat NMR spectra of ribose (A), sucrose (B) y xylose (C) standards.



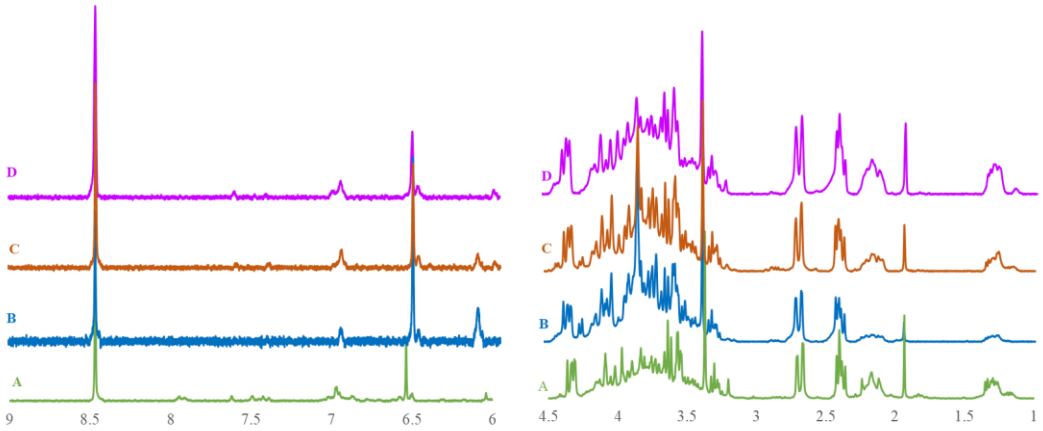
Appendix 36 Stacked one dimensional ^{1}H water_presat NMR spectra of arabinose (A), dextrose (B), fructose (C), galactose (D) y mannose (E) standards.



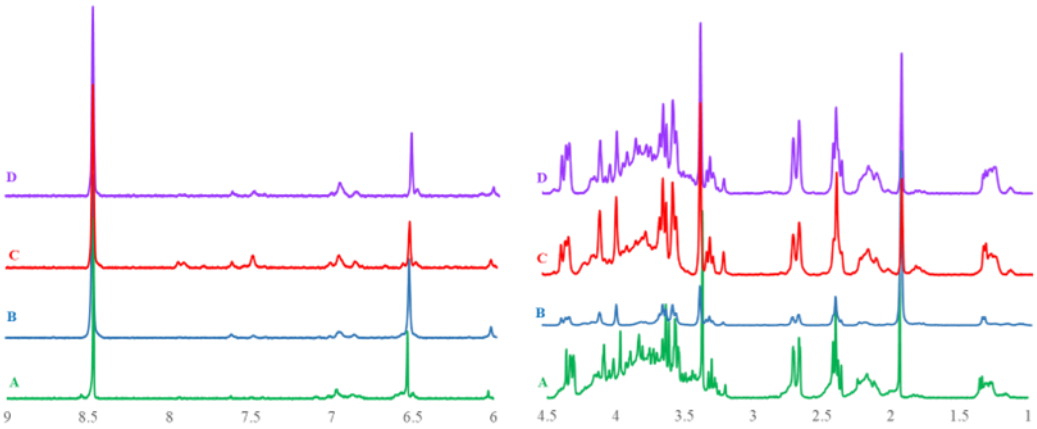
Appendix 37 Stacked one-dimensional ^{1}H water_presat NMR spectra of cocoa pod husk samples after applying hydrothermal pretreatments 355µm-100°C (A), 250µm-100°C (B), 150µm-100°C (C) y <150µm-100°C (D).



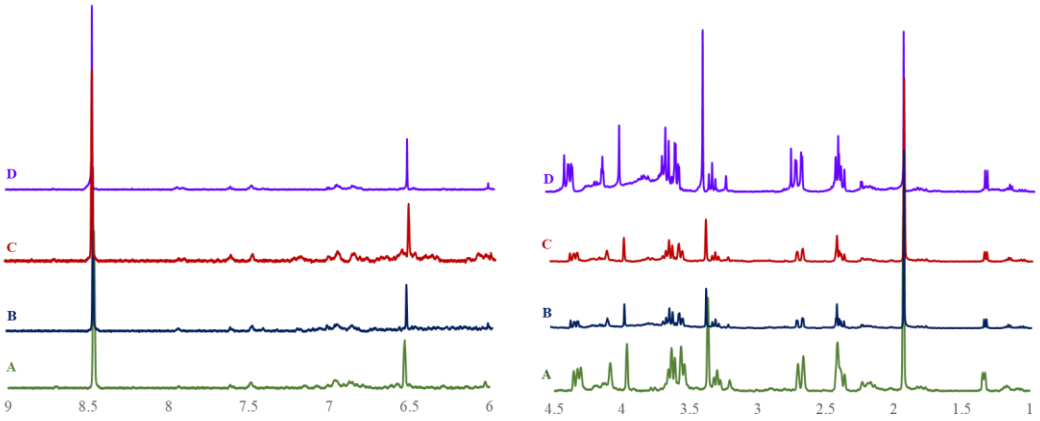
Appendix 38 Stacked one-dimensional ^{1}H water_presat NMR spectra of cocoa pod husk samples after applying hydrothermal pretreatments 355µm-120°C (A), 250µm-120°C (B), 150µm-120°C (C) y <150µm-120°C (D).



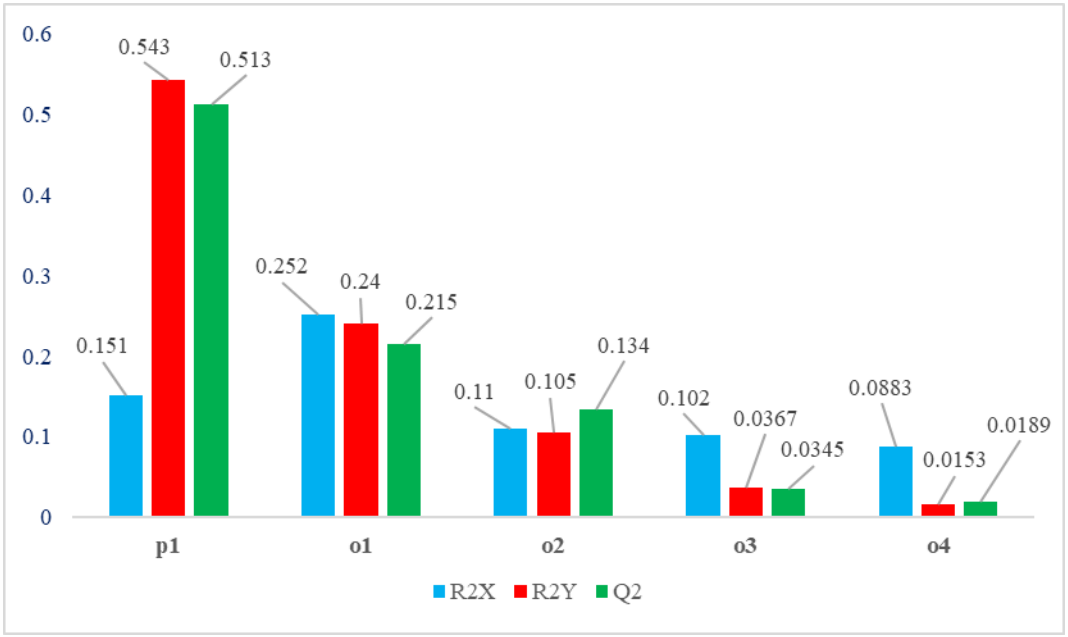
Appendix 39 Stacked one-dimensional ^{1}H water_presat NMR spectra of cocoa pod husk samples after applying hydrothermal pretreatments 355µm-150°C (A), 250µm-150°C (B), 150µm-150°C (C) y <150µm-150°C (D).



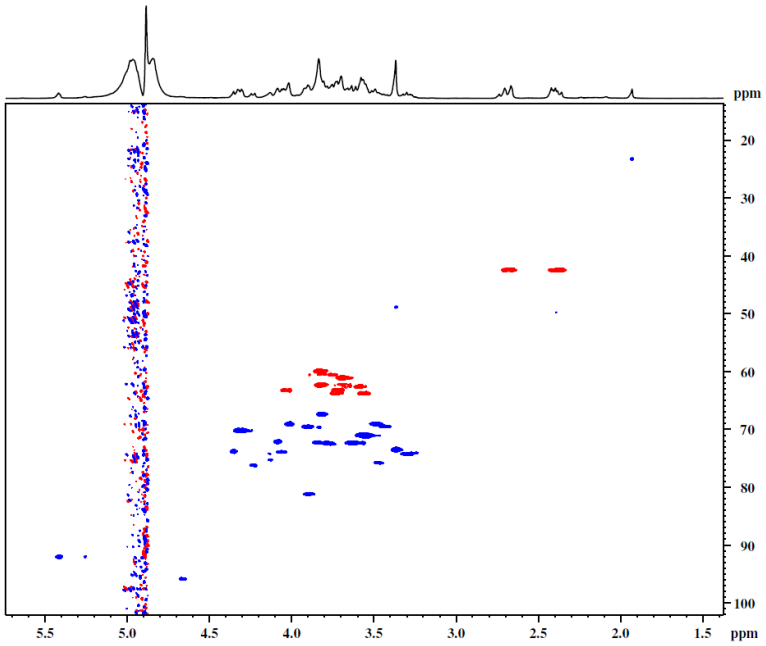
Appendix 40 Stacked one-dimensional ^{1}H water_presat NMR spectra of cocoa pod husk samples after applying hydrothermal pretreatments 355µm-180°C (A), 250µm-180°C (B), 150µm-180°C (C) y <150µm-180°C (D).



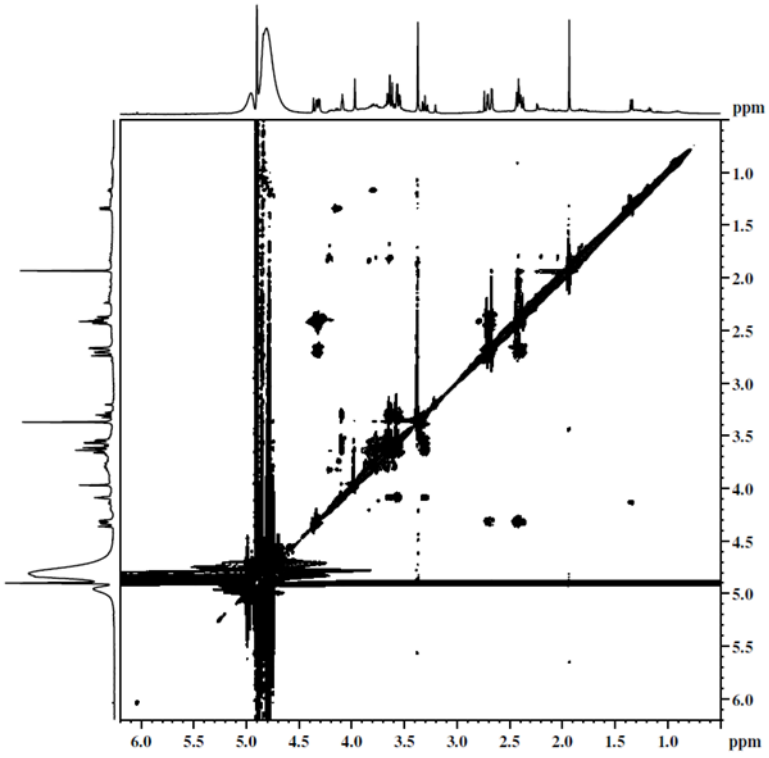
Appendix 41 Stacked one dimensional ^{1}H water_presat NMR spectra of cocoa pod husk samples after applying hydrothermal pretreatments 355µm-200°C (A), 250µm-200°C (B), 150µm-200°C (C) y <150µm-200°C (D).



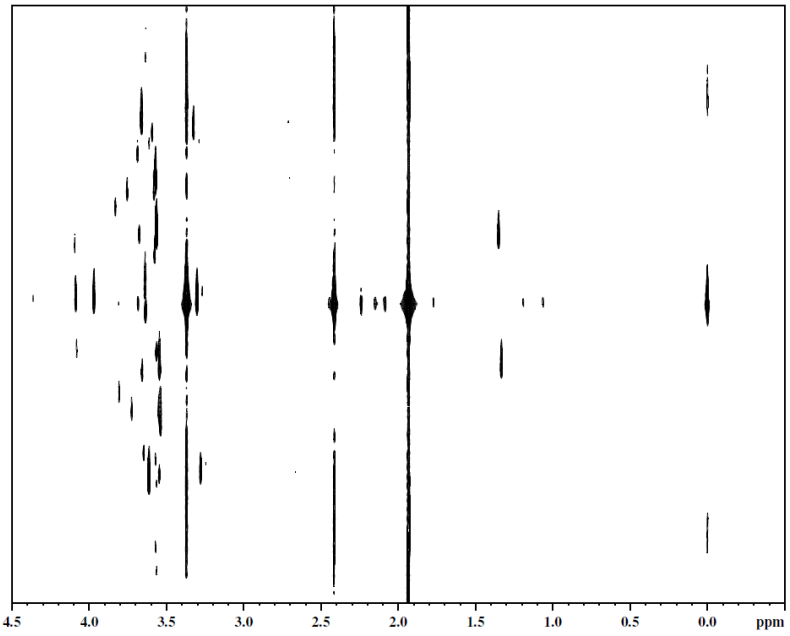
Appendix 42 The permutation analysis between one predictive (p1) and four orthogonal (o1, o2, o3, and o4) components produced the observed and cross-validated R2X, R2Y, and Q2 coefficients.



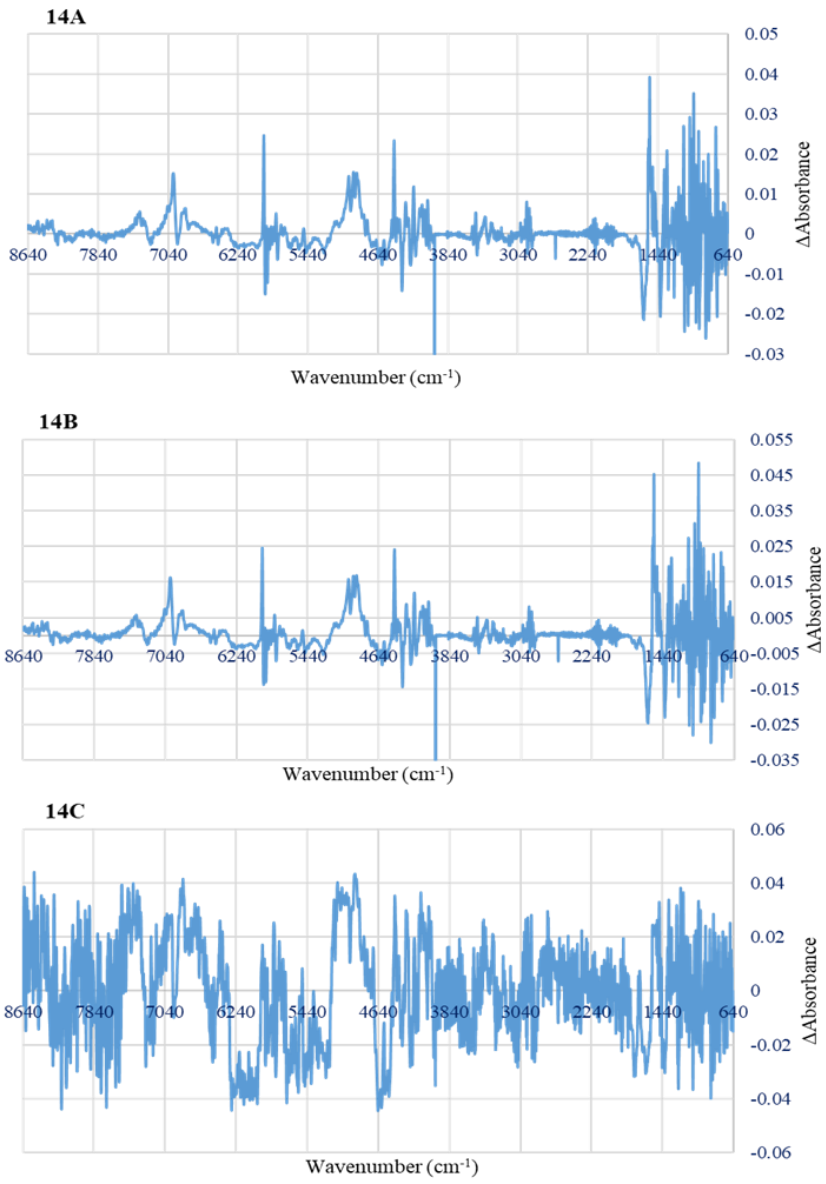
Appendix 43 HSQC150-1 NMR spectra



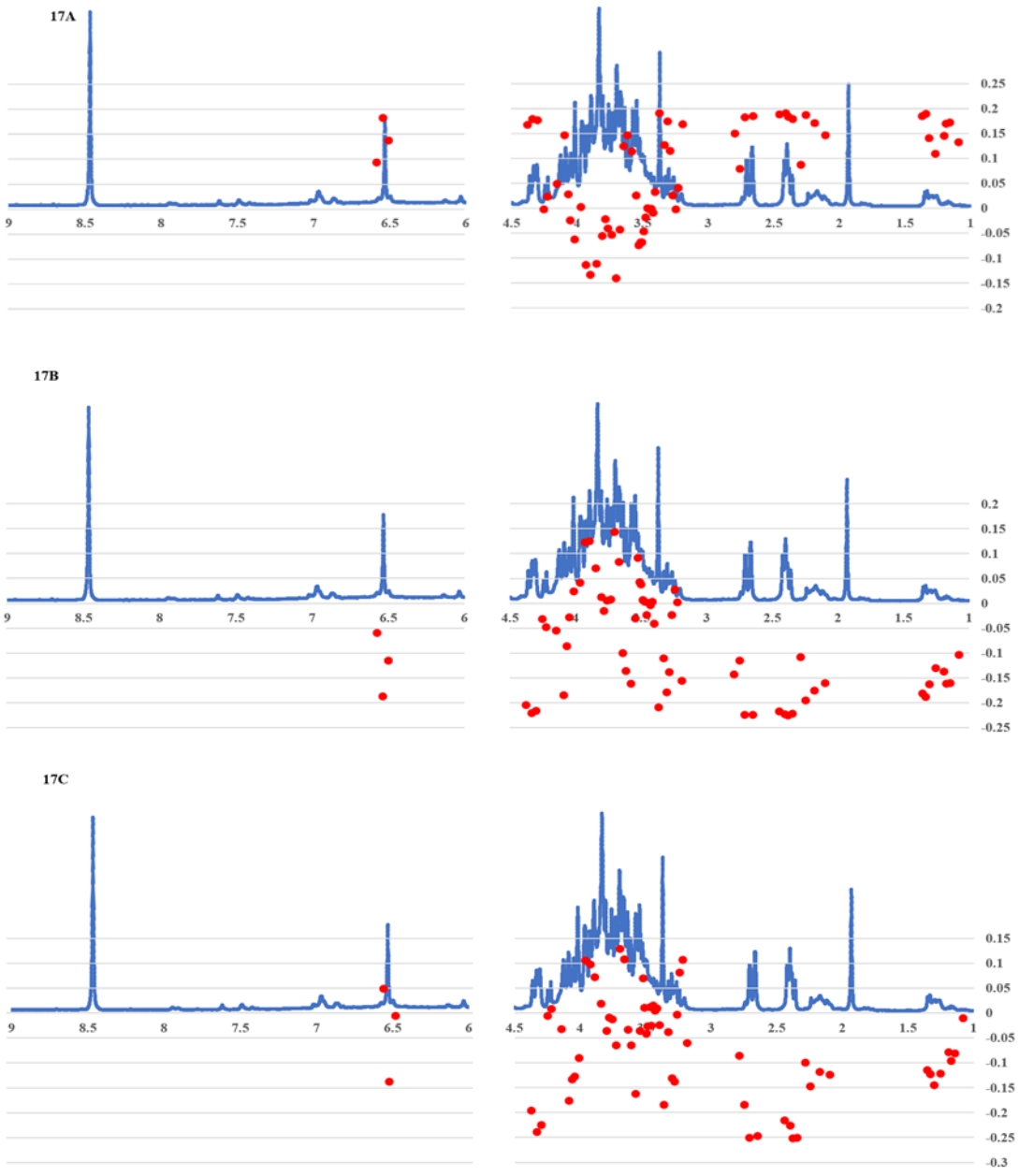
Appendix 44 TOCSY150-5 NMR spectra



Appendix 45 JRES150-4 NMR spectra



Appendix 46 1st derivative MIR-NIR loading plots comprising from Principal component analysis (14A), partial least squares discriminant analysis (14B), and Orthogonal PLS-DA permutation (14C).



Appendix 47 NMR loading plots from PCA Loadings (17A), PLS-DA Loadings (17B), and OPLSDA Loadings (17C).