Food Research International 162. Part B (2022): 112117

DOI: https://doi.org/https://doi.org/10.1016/j.foodres.2022.112117

Journal Pre-proofs

Gastrointestinal fate of phenolic compounds and amino derivatives from the cocoa shell: An *in vitro* and *in silico* approach

Silvia Cañas, Miguel Rebollo-Hernanz, Cheyenne Braojos, Vanesa Benítez, Rebeca Ferreras-Charro, Montserrat Dueñas, Yolanda Aguilera, María A. Martín-Cabrejas

PII: S0963-9969(22)01175-9
DOI: https://doi.org/10.1016/j.foodres.2022.112117
Reference: FRIN 112117

To appear in: *Food Research International*

Received Date: 8 August 2022
Revised Date: 20 October 2022
Accepted Date: 6 November 2022

Please cite this article as: Cañas, S., Rebollo-Hernanz, M., Braojos, C., Benítez, V., Ferreras-Charro, R., Dueñas, M., Aguilera, Y., Martín-Cabrejas, M.A., Gastrointestinal fate of phenolic compounds and amino derivatives from the cocoa shell: An *in vitro* and *in silico* approach, *Food Research International* (2022), doi: https://doi.org/10.1016/j.foodres.2022.112117

Gastrointestinal fate of phenolic compounds and amino derivatives from the cocoa shell: An *in vitro* and *in silico* approach

Silvia Cañas[1,2], Miguel Rebollo-Hernanz[1,2], Cheyenne Braojos[1,2], Vanesa Benítez[1,2], Rebeca Ferreras-Charro[3], Montserrat Dueñas[3], Yolanda Aguilera[1,2], María A. Martín-Cabrejas*[1,2]

[1] Department of Agricultural Chemistry and Food Science, Faculty of Science, C/ Francisco Tomás y Valiente, 7. Universidad Autónoma de Madrid, 28049, Madrid, Spain.

[2] Institute of Food Science Research (CIAL, UAM-CSIC). C/ Nicolás Cabrera, 9. Universidad Autónoma de Madrid, 28049, Madrid, Spain.

[3] Grupo de Investigación en Polifenoles, Unidad de Nutrición y Bromatología, Facultad de Farmacia, Universidad de Salamanca, Campus Miguel de Unamuno, 37007 Salamanca, Spain.

*Corresponding author.

E-mail address: maria.martin@uam.es

Abstract

The objective of this study was to assess how *in vitro* gastrointestinal digestion influenced the bioaccessibility and potential bioavailability of phenolic compounds and methylxanthines in the cocoa shell (CS) in the form of flour (CSF) and aqueous extract (CSE). To comprehend how these phytochemicals behaved during gastrointestinal digestion, we also modeled *in silico* the colonic microbial biotransformation of the phenolic compounds in the CS. Different groups of phenolic compounds (mainly gallic and protocatechuic acids, and catechin) and methylxanthines (theobromine and caffeine) could be found in the CS. Methylxanthines and phenolic compounds were released differently during gastrointestinal digestion. Whereas digestion triggered the release of hydroxybenzoic acids (67–73%) and flavan-3-ols (73–88%) during the intestinal phase, it also caused the degradation of flavonols and flavones. Besides, the release of phytochemicals was significantly influenced by the CS matrix type. Phenolic compounds were protected by the CSF matrix. Phenolic acids from CSF were more bioaccessible in the intestinal (1.2-fold, $p < 0.05$) and colonic (1.3-fold, $p < 0.05$) phases than those from the CSE. Methylxanthines were also more bioaccessible in the intestinal (1.8-fold, $p < 0.01$) and colonic phases (1.3-fold, $p < 0.001$) and bioavailable (1.8-fold, $p < 0.001$) in the CSF. Colonic metabolism demonstrated that the gut microbiota could biotransform non-absorbed phenolic compounds into other lower molecular weight and more bioavailable metabolites. These findings support the CS's potential as a source of bioaccessible, bioavailable, and active phytochemicals.

Keywords: cocoa shell; cocoa by-products; phenolic compounds; *N*-phenylpropenoyl-L-amino acids; methylxanthines; *in vitro* digestion; bioaccessibility; bioavailability

1. Introduction

The cocoa shell (CS) is the main by-product of cocoa processing. Specifically, the CS is separated from the cotyledons during the roasting stage, being considered a waste (Panak Balentić et al., 2018). Currently, 4.7M tons of cocoa are produced annually, and consequently, over 900k tons of cocoa shells are generated since the CS accounts for 20% of the cocoa bean (Boeckx et al., 2020). The large amount of CS discarded during cocoa processing may pose economic problems by damaging the environment (Panak Balentić et al., 2018). Although the CS is considered an industrial by-product, its nutritional composition is similar to that of the cocoa bean, except for its lower fat and higher fiber contents. The CS is composed of proteins (10−27%), carbohydrates (8−70%), dietary fiber (39−66%), and lipids (1−8%) with a similar fatty acid composition as the cocoa butter (Cinar et al., 2021). Likewise, the CS contains a plethora of bioactive compounds, which can confer health benefits. The main phytochemicals in the CS include methylxanthines, such as theobromine and caffeine, and phenolic compounds (Rojo-Poveda et al., 2021). The phenolic composition of the CS is comparable to that of the cocoa beans, primarily comprising phenolic acids and flaval-3-ols. Furthermore, the CS may contain *N*-phenylpropenoyl-L-amino acids (NPAs), a family of phenolic/amino acid conjugates exclusive from cocoa (Lechtenberg et al., 2012). Phenolic compounds are among the most frequent plant secondary metabolites. Considered natural antioxidants, they elicit diverse beneficial effects in humans. We previously evidenced the *in vitro* effects of the methylxanthines and phenolic compounds found in the CS on ameliorating biomarkers of metabolic syndrome in both liver (Rebollo-Hernanz et al., 2022) and adipose tissue (Rebollo-Hernanz et al., 2019). Our studies have also shown the vasoactive properties of a CS aqueous extract, counteracting age-related endothelial dysfunction *in vivo* (Rodríguez-Rodríguez et al., 2022).

Nonetheless, the bioactivity of these compounds depends on how much they are released from the matrix, how they change during gastrointestinal digestion, and their metabolism in the gut, liver, and target tissues. Bioaccessibility measures the proportion of a compound available for intestinal absorption after being released from the food matrix into the gastrointestinal tract and biotransformed into forms that could be more bioavailable (including host and gut microbiota reactions) (Rodrigues et al., 2022). Thus, evaluating compounds' bioaccessibility is critical in analyzing their bioactivity (Cosme et al., 2020). Several digestion models have been described to study the phytochemicals' bioaccessibility. Among them, static models simulating gastrointestinal digestion can be useful in predicting the results of *in vivo* digestion, as they present many advantages, such as simplicity, rapidity, and cheapness (Brodkorb et al., 2019). Successively, bioavailability refers to the concentration of compounds absorbed in the gastrointestinal tract that reaches the target tissues in the intact or biotransformed form to exert their bioactivity or to be stored (Rodrigues et al., 2022). The standard analyses of potential absorption and bioavailability are carried out using Caco-2 cells, which simulate the human intestinal epithelium. However, they present some disadvantages, such as long culture times and high costs. Hence, simulated models emerge as practical approaches to assess target compounds' potential intestinal permeability and mimic their gut microbiota metabolism rapidly and inexpensively (Wang et al., 2016). This study aimed to evaluate the impact of *in vitro* gastrointestinal digestion on the bioaccessibility and potential bioavailability of phenolic compounds and methylxanthines of the CS. We also modeled the colonic microbial biotransformation of the CS's phenolic compounds to understand these phytochemicals' behavior throughout gastrointestinal digestion.

2. Materials and Methods

2.1. Materials

Methanol, sodium hydroxide, formic acid, acetonitrile, hydrochloride acid, sodium carbonate, ferric chloride hexahydrate, potassium chloride, sodium bicarbonate, sodium chloride, and Folin-Ciocalteu reagent were supplied by Panreac Química SLU (Barcelona, Spain). Polyphenols standards (≥ 96%), including gallic and protocatechuic acids, and catechin, epicatechin, and quercetin were provided by Sigma Aldrich (Sigma-Aldrich, St. Louis, MO), and Extrasynthese (Genay, France). Alkaloids standards (theobromine and caffeine (≥ 98.5%)) were purchased from Sigma Aldrich (Sigma-Aldrich, St. Louis, MO). 6-hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid (Trolox), 2,2′-Azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS), potassium persulfate, 2,4,6-tris(2-pyridyl)-s-triazine (TPTZ, 99%), ammonium carbonate, calcium chloride dihydrate, magnesium chloride hexahydrate, potassium phosphate monobasic, porcine pepsin, porcine pancreatin, pronase E and Viscozyme were purchased from Sigma Aldrich (Sigma-Aldrich, St. Louis, MO).

2.2. Sample preparation

The CS was provided as a dry material by Chocolates Santocildes (Castrocontrigo, León, Spain) and stored at 4º C until use. The CS was milled using a laboratory grinder, obtaining the CS flour (CSF). The flour was stored in sealed flasks at −20 ºC until analysis. The CS aqueous extract (CSE) was produced according to extraction conditions previously optimized for the ground CS (Rebollo-Hernanz et al., 2021). Briefly, CSE (0.02 g mL^{-1} solid-to-solvent ratio) was added to boiling water (100 °C) and stirred for 90 min. After extraction, the aqueous extract was filtered, frozen at −20 °C, freeze-dried, and stored at −20 °C until further usage.

2.3. In vitro simulated digestion

The simulated gastrointestinal digestion was performed according to the INFOGEST *in vitro* digestion protocol with slight modifications (Brodkorb et al., 2019). Briefly, 1 g of CSF or 100 mg of CSE were mixed with the simulated salivary fluid for simulating the oral phase, and the mixture was maintained for 2 min at 37 °C under agitation. The gastric phase was performed by combining the oral phase with simulated gastric fluid and adding porcine pepsin solution (2000 U mL^{-1} of digestion), and the samples were incubated at 37 °C for 2 h under stirring. The intestinal phase was simulated by mixing the gastric phase with simulated intestinal fluid containing pancreatin (100 U trypsin activity mL^{-1} of digest). The mixture was incubated at 37 °C while stirring for 2 h. Colonic digestion was simulated as described by Papillo et al. (2014). Pronase E (5 mL, 1 mg mL^{-1}) was added to the intestinal phase, and the pH was adjusted to 8.0. The samples were incubated at 37 ºC for 1 h under stirring. After incubation, the samples were adjusted to pH 4, and 150 µL of Viscozyme was added. The samples were incubated at 37°C for 16 h under agitation. A digestion blank was prepared for each digestive phase containing the simulated digestion fluids. The supernatants and residues obtained from each digestive stage were freeze-dried and stored at −20 ºC until utilization.

2.4. Extraction of free and bound phenolic compounds

Free and bound phenolic compounds were extracted according to Rebollo-Hernanz et al. (2020). For free phenolic compounds extraction, 1 g of CSF and the residues obtained from each digestive phase of CSF digestions were combined with 50 mL of methanol: HCl (1‰)-water 80:20 (v/v). The samples were sonicated for 30 min and incubated at 40 ºC for 16 h under agitation. After incubation, the samples were centrifuged at 4000 × *g* for 15 min at room temperature, collecting the supernatants obtained. The extraction process was repeated twice. The supernatants were combined and then concentrated at 40 ºC under vacuum. Alkaline extraction of bound phenolic compounds was carried out by adding 5 mL of 4 mol L^{-1} NaOH

to the residue obtained after free phenolic compounds extraction. The samples were then agitated for 1 h under a nitrogen atmosphere. The samples were acidified to pH 2.0 and centrifuged at 4000 × *g* for 15 min at room temperature, collecting the supernatant. Afterward, extraction with 1:1 ethyl ether: ethyl acetate was performed. Three more extractions were performed using methanol: HCl (1‰)-water 80:20 (v/v). Finally, the supernatants recovered were combined and concentrated at 40 ºC under vacuum.

2.5. Spectrophotometric analysis of total phenolic content and antioxidant capacity

2.5.1. Total phenolic content (TPC)

Total phenolic compounds were assessed by the Folin-Ciocalteu method, according to a previously adapted protocol (Rebollo-Hernanz et al., 2021). Briefly, 10 µL of the sample were mixed with 150 µL of diluted Folin-Ciocalteu reagent (1:14, v/v in Milli-Q water) in a 96-well plate and incubated at room temperature for 3 min. Then, 50 µL of Na_2CO_3 20% were added to the mixture, and the plate was incubated for 2 h at room temperature in the dark. After incubation, absorbance was recorded at 750 nm. A standard curve of gallic acid (0.01–0.2 mg mL^{-1}) was done to estimate the concentration of total phenolic compounds. The results were expressed as mg gallic acid equivalents per gram (mg GAE g^{-1}).

2.5.2. ABTS radical scavenging capacity

Radical scavenging capacity was measured by the $ABTS^{\cdot+}$ assay as previously described (Rebollo-Hernanz et al., 2021). To obtain 2.2′-Azino-bis(3-ethylbenzothiazoline-6-sulfonic) acid radical cations ($ABTS^{\cdot+}$), a solution of ABTS (7 mmol L^{-1}) was mixed with potassium persulfate (2.45 mmol L^{-1}) and shaken in the dark for 16 h. The $ABTS^{\cdot+}$ solution was diluted in PBS 5 mmol L^{-1}, pH 7.4, to adjust the absorbance to 0.70 at 734 nm. The assay was

performed by mixing 30 µL of the sample and 270 µL of the diluted $ABTS^{\cdot+}$ solution in each well in a 96-well plate. The plate was incubated for 10 min, and the absorbance was measured at 734 nm in a microplate reader. A calibration curve was prepared using Trolox reagent as a standard solution (0–0.06 mg mL^{-1}). The values were expressed as mg Trolox equivalent per gram (mg TE g^{-1}).

2.5.3. Ferric Reducing Antioxidant Power (FRAP)

The antioxidant capacity was measured through the FRAP assay as previously described (Rebollo-Hernanz et al., 2020). Samples (10 µL) were mixed with 300 µL of a working FRAP reagent (acetate buffer 0.3 mol L^{-1} pH 3.6, 10 mmol L^{-1} tripyridyl s-triazine, 40 mmol L^{-1} HCl, 20 mmol L^{-1} $FeCl_3 \cdot 6H_2O$ (10:1:1) (v/v/v)) to each well in a 96-well plate. After incubating the plate for 10 min at 37°C, the absorbance was read at 593 nm. Trolox was used as a standard solution (25–800 µmol L^{-1}). FRAP results were calculated and expressed as mmol Trolox equivalent per gram (mmol TE g^{-1}).

2.6. HPLC-DAD-ESI/MS^n qualitative and quantitative analyses of phenolic compounds and methylxanthines

Samples were analyzed using a Hewlett–Packard-1100 HPLC-diode array detector (DAD) chromatograph (Agilent Technologies, Palo Alto, CA). The solvents used were 0.1% formic acid in water (solvent A) and 100% acetonitrile (solvent B). The elution gradient established was isocratic 15% B for 5 min, 15–20% B for 5 min, 20–25% B for 10 min, 25–35% B for 10 min, 35–50% B for 10 min, and re-equilibration of the column. The separation of phytochemicals was performed in a Spherisorb S3 ODS-2 C8 column (Waters, Millford, USA) (3 µm, 150 mm × 4.6 mm i.d.) at 35 °C, with a flow rate of 0.5 mL min^{-1}. DAD-detection was carried out at 280 nm and 370 nm as preferred wavelengths. The mass spectrometer (MS) was connected to the HPLC system via the DAD cell outlet, and detection was performed in an API-

3200 Qtrap (Applied Biosystems, Darmstadt, Germany) equipped with an ESI source and triple quadrupole-ion trap mass analyzer. The phenolic compounds and methylxanthines were characterized by their retention times, UV and mass spectra, and fragmentation patterns compared to authentic standards when available. For quantitative analysis, amino derivatives of caffeic and *p*-coumaric acid were quantified by calibration curves of the corresponding free acid. Apigenin 6-*C*-glucoside was used for *C*-glycosides flavones derived from apigenin; flavonols, derivatives of quercetin by the curve of quercetin-3-*O*-glucoside. Theobromine and caffeine were quantified by the calibration curves of their respective standard. The results were expressed as μg g^{-1} of sample.

2.7. Retention index and bioaccessibility calculation

The retention index and bioaccessibility of the phenolic compounds and methylxanthines from the CS, expressed as a percentage, were determined as follows:

$$\text{Retention Index or Bioaccessibility (\%)} = \frac{C_{Digested\ fraction}}{C_{\text{Non - digested fraction}}} \times 100$$

where $C_{Digested\ fraction}$ corresponds to the concentration of compounds in the soluble fraction obtained after *in vitro* digestion, and $C_{Non\text{-}digested\ fraction}$ is the concentration of compounds in the sample before *in vitro* digestion. The retention index was calculated for oral and gastric stages, whereas bioaccessibility was calculated for intestinal and colonic stages.

2.8. Simulated intestinal absorption and bioavailability calculation

The potential absorption of the bioactive compounds found in the CSF and CSE was evaluated *in silico*. Predictions of Caco-2 Absorption (C2A) and Human Intestinal Absorption (HIA) were calculated using canonical SMILES sequences obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov/, accessed on April 19th, 2022), using pkCSM-

pharmacokinetics (http://biosig.unimelb.edu.au/pkcsm/, accessed on April 19th, 2022) and ADMETlab (https://admet.scbdd.com/, accessed on April 19th, 2022) cheminformatics free software. The potential bioavailability of the bioactive compound was calculated as follows:

$$\text{Bioavailability (\%)} = \frac{C_{Intestinal\ fraction} \times Absorption}{C_{\text{Non - digested}\ fraction}} \times 100$$

where $C_{Intestinal\ fraction}$ corresponds to the concentration of compounds in the soluble intestinal fraction obtained after *in vitro* digestion, *Absorption* corresponds to the percentage of absorption estimated *in silico* for each compound, and $C_{Non\text{-}digested\ fraction}$ is the concentration of compounds in the sample before *in vitro* digestion.

2.9. Simulated colonic gut biotransformation

The human gut metabolism of the tentatively identified compounds was predicted *in silico* using Biotransformer (Djoumbou-Feunang et al., 2019). The metabolism prediction was carried out using the Human Gut Microbial Transformations option. This software predicts small-molecule metabolism by gut microbial enzymes after entering each compound's canonical SMILES sequences. For the compounds containing unknown glycosides, i.e., hexosides and pentosides, glucosides and arabinosides were considered, respectively, when selecting the SMILE. Only reactions in the phenolic moiety of the compounds were considered (excluding potential changes in the conjugated amino acid). Only catabolic reactions (hydrolysis, reduction, dehydroxylation, oxidation, and C-ring fission) were considered, while conjugation reactions (methylation, sulphation, or glucuronidation) were ignored.

2.10. Statistical analysis

Results are expressed as the mean ± standard deviation (SD) of at least three independent experiments ($n = 3$). Data were analyzed by one-way analysis of variance (ANOVA) and post

hoc Tukey test for comparisons among digestive phases. *T*-test comparisons were performed between the retention indexes and the bioaccessibility of CSF and CSE. Differences were considered significant at $p < 0.05$.

3. Results

3.1. The cocoa shell is a source of different phenolic compounds and methylxanthines

HPLC-DAD-MS^n analysis in the CS revealed the presence of 13 phenolic compounds and 2 methylxanthines (**Table 1**). Among the phenolic compounds studied, two major groups were identified and quantified: phenolic acids (9) and flavonoids (4). Phenolic acids were further classified into hydroxybenzoic acids and NPAs. As hydroxybenzoic acids, gallic (compound 1) and protocatechuic (compound 3) acids were identified by comparing their retention times and UV spectra with commercial standards. Furthermore, based on UV–vis and HPLC-DAD-MS^n analyses and previous reports (Oracz et al., 2019), six NPAs were tentatively identified as *N*-caffeoyl-L-aspartate, *N*-coumaroyl-L-aspartate *cis*, *N*-coumaroyl-L-aspartate *trans*, *N*-caffeoyl-L-DOPA *cis*, *N*-caffeoyl-L-DOPA *trans,* and *N*-coumaroyl-tyrosine. These hydroxycinnamic acid/amino acid conjugates (compounds 4, 5, 7, 10, 12, and 14) were identified in the different samples. Compound 4 produced a pseudomolecular ion $[M-H]^-$ at *m/z* 294, yielding a major MS^2 fragment ion at *m/z* 179, owing to caffeic acid excision (Oracz et al., 2019). This molecule was putatively identified as *N*-caffeoyl-L-aspartate by contrasting the MS^2 fragmentation patterns in previous reports (Cádiz-Gurrea et al., 2014). Compounds 5 and 7 exhibited a $[M-H]^-$ ion at *m/z* 278 and generated MS^2 fragmentation ions at *m/z* 235 and 163, indicative of the fragmentation of *N*-coumaroyl-L-aspartate (Cádiz-Gurrea et al., 2014). Compounds 10 and 12 exhibited the pseudomolecular ion $[M-H]^-$ at *m/z* 358 and fragment ions at *m/z* 179, identified

as clovamide (*N*-caffeoy-L-DOPA). Compound 14 showed the $[M-H]^-$ ion at *m/z* 326 and fragment ions at *m/z* 282, 163, and 146 corresponding to *N*-coumaroyl-L-tyrosine, based on the MS^2 fragmentation patterns retrieved from published data (Oracz et al., 2019).

Flavonoids were classified into three groups: flavan-3-ols, flavonols, and flavones. Concerning flavan-3-ols, compounds 6 and 8 showed UV spectra like (+)-catechin and (−)-epicatechin, respectively. In the case of flavonols, the deprotonated ions of the compounds (*m/z* 463) yielded a fragment ion at *m/z* 301, equivalent to quercetin 3-*O*-hexoside (compound 13) and quercetin 3-*O*-pentoside (compound 15). They were identified according to their retention time, mass, and UV-vis characteristics by comparison with commercial standards. Regarding flavones, compound 9 (*m/z* 593) was identified as apigenin 6,8-di-*C*-glucoside by comparing its feature with the standard. Its fragmentations in MS^2 analyses were characteristic of *C*-glycosidic flavones (Martini et al., 2018).

Compounds 2 and 11 had a maximum absorbance at 272 and 273 nm, respectively, the characteristic maximum of methylxanthines, and did not produce ions in the MS negative mode. They were identified as theobromine and caffeine compared with the pure standard retention time and spectra.

The composition of phenolic compounds studied in the CSF revealed that hydroxybenzoic acids constituted the primary phenolic fraction (52.0%); gallic acid accounted for 36.9% and protocatechuic acid represented 15.1% of the total phenolics (**Table 2**). The NPAs detected in CSF (17.0%) included *N*-caffeoyl-L-aspartate (11.8%), *N*-caffeoyl-L-DOPA *cis* (2.1%), *N*-coumaroyl-L-aspartate *cis* (2.0%) and *N*-coumaroyl-L-aspartate *trans* (1.1%). Catechin constituted the major flavonoid (26.5% of the concentration of total phenolics), while epicatechin (3.1%), quercetin 3-*O*-glucoside (0.7%), and quercetin 3-*O*-pentoside (0.8%) were found in minor concentrations. Methylxanthines were the main fraction among the identified

compounds, present in 16.0-fold higher concentration than phenolic compounds. Among methylxanthines, theobromine (75.6%) was the primary compound, followed by caffeine (24.4%). The phenolic profile of the CSE followed the same pattern as the CSF (**Table 2**). Correspondingly, hydroxybenzoic acids (53.8%) were the major phenolic fraction, whereas gallic and protocatechuic acids were two of the main phenolic compounds, representing 36.6 and 17.2% of the total phenolics, respectively. Among the NPAs found in CSE, the content of *N*-caffeoyl-L-aspartate (9.5%) stood out among others, followed by *N*-coumaroyl-L-aspartate *cis* (3.1%), *N*-coumaroyl-L-aspartate *trans* (2.2%), *N*-caffeoyl-L-DOPA *cis* (1.9%), and *N*-coumaroyl-L-tyrosine (0.4%). Among flavonoids, catechin constituted 22.8% of the total phenolic compounds in the CSE. Likewise, as in the flour, epicatechin (1.7%), quercetin 3-*O*-glucoside (0.7%), and quercetin 3-*O*-pentoside (0.6%) represented a smaller proportion of the total phenolics. In addition, two new phenolics which had not been detected in CSF were identified in the CSE (*N*-caffeoyl-L-DOPA *trans* (0.5%) and apigenin 6,8-di-*C*-glucoside (1.4%). Methylxanthines also represented a large fraction of the total identified compounds. Theobromine showed a 5.0-fold higher concentration in the CSE than in the CSF, while caffeine was 80% lower in the CSE.

3.2. Phenolic compounds and methylxanthines were released throughout the digestion

The *in vitro* simulated digestion elicited an increase in the TPC from 46.3 to 58.1 mg GAE g^{-1} in the digested CSE (D-CSE) (**Figure 1A**). In the digested CSF (D-CSF), the TPC reached 6.6 mg GAE g^{-1}, while a significant decrease (from 34.2 to 24.8 mg GAE g^{-1}) was observed in the remained total non-digested phenolics (TND-CSF). Consequently, this reduction was also noticed in the free (FND-CSF) and bound (BND-CSF) phenolic compounds from the non-digested phase. However, a significant increase (from 10.5 to 14.7 mg GAE g^{-1}) was detected

from the intestinal phase to the colonic phase in the case of the FND-CSF. Then, the release of bound phenolic compounds from the non-digested CSF residue (BND-CSF) during the colonic stage favored the subsequent rise in the FND-CSF. The antioxidant capacity determined by the ABTS method increased significantly (7.4-fold) throughout gastrointestinal digestion in the CSE's bioaccessible fraction (D-CSE), undergoing a slight decrease (5.0%) in the colonic phase (**Figure 1B**). In the case of the CSF's digested fraction (D-CSF), the antioxidant capacity increased in the final stages of digestion due to the release of phenolic compounds from the matrix, which consequently decreased the antioxidant capacity (from 73.4 to 59.7 mg TE g^{-1}) in the non-digested CSF residue (TND-CSF). Similarly, the antioxidant capacity in the FND-CSF suffered a slight decrease at the first stages of digestion, increasing in the colonic phase caused by the phenolic compounds' liberation from the matrix to the digested fraction. Conversely, the digestive process reduced by 49.0% the antioxidant capacity in the D-CSE measured by the FRAP method, while in the D-CSF, it caused a gradual increase in the antioxidant capacity until the intestinal phase (**Figure 1C**). In the remained non-digested TND-CSF, the antioxidant capacity decreased by 20.8% throughout the *in vitro* digestion. The antioxidant capacity relative to the fractions of free and bound phenolics showed a similar pattern observed in ABTS and TPC, underlining the increase in the antioxidant capacity of the FND-CSF and their decrease in the BND-CSF during the colonic phase.

The concentration of phenolic acids (hydroxybenzoic and NPAs) in the CSF did not experience any significant change at the end of digestion compared to the non-digested CSF (**Figure 1D**). Phenolic acids from the CSE were ultimately released during the oral phase, experiencing a slight decrease (25.6%) at the end of the colonic stage of digestion. The same behavior was also noticed for the hydroxybenzoic acids of both matrices since they represented the most important fraction among the phenolic acids studied. However, the concentration of NPAs decreased by 42.2% from the intestinal to the colonic phase in the CSF, and by 84.6%, during the intestinal

phase, in the case of the CSE. Regarding the CSF, flavonoids (flavan-3-ols and flavonols) increased 2.9-fold from the oral to the colonic phase, and the same behavior was observed for flavan-3-ols, which constituted their main fraction (**Figure 1E**). Although a substantial fraction of flavonols was released during the oral and gastric phases, they were not detected during the intestinal and colonic stages. Flavonoids remained stable in the CSE throughout the digestion process but during the colonic phase. The flavonoid content significantly increased 1.9-fold in the colonic phase compared to the intestinal phase. Flavan-3-ols revealed the same behavior. Contrariwise, flavonols in the CSE were released entirely during the oral and gastric phases (92.6%), but they were not detected during the intestinal and colonic phases. Total phenolics in the CSF were significantly released during the oral phase and increased (2.6-fold) throughout digestion, being wholly released during the colonic phase (**Figure 1F**). In contrast, total phenolics from the CSE were fully released during the oral phase. Although the digestive process resulted in a slight reduction in the concentration of total phenolics, especially from the gastric to the intestinal phase (23.9%), the concentration increased from the intestinal to the colonic phase (1.5-fold). Methylxanthines were sequentially released from the CSF throughout gastrointestinal digestion, whereas they were partially degraded in the CSE (**Figure 1F**).

3.3. Phenolic compounds' bioaccessibility depended on the matrix type

The bioaccessibility of total phenolics increased during digestion in the CSF while remaining unchanged in the CSE, where there was a slight decrease during the intestinal phase (**Figure 2A**). Phenolic compounds in the CSE reached a significantly higher retention index than CSF ones during the oral and gastric phases but not in the intestinal and colonic phases. Phenolic acids presented similar behavior as total phenolics (**Figure 2B**). In the CSF, phenolic acids were released from the matrix during the digestive process, showing the highest bioaccessibility in the colonic phase (96.3%). Phenolic acids in the CSE presented a higher retention index than

in the CSF in the oral and gastric phases, unlike the intestinal and colonic phases. Flavonoid retention index and bioaccessibility increased in the CSF and the CSE throughout gastrointestinal digestion (**Figure 2C**). The retention index and bioaccessibility of flavonoids were higher in the CSE than in the CSF, excluding the colonic stage. Methylxanthines increased their retention index and bioaccessibility during digestion in the CSF (**Figure 2D**). CSF methylxanthines' bioaccessibility increased from the oral (42.0%) to the colonic phase (92.6%). Contrarily, the CSE methylxanthines' bioaccessibility did not significantly change at the end of *in vitro* digestion, although it decreased during the intestinal phase. Overall, we observed two different trends: phenolic acids experienced a sequential release from the oral to the colonic phases, exhibiting lower (40%, $p < 0.001$) retention indexes in the CSF than in the CSE during the oral phase but higher during the intestinal (1.2-fold, $p < 0.05$) and colonic (1.3-fold, $p < 0.05$) phases. Methylxanthines exhibited the same trend, being more bioaccessible in the intestinal (1.8-fold, $p < 0.01$) and colonic phases (1.3-fold, $p < 0.001$) in the CSF than in the CSE. On the contrary, flavonoids tended to be more bioaccessible in the CSE than in the CSF, independently of the digestion stage.

3.4. Hydroxybenzoic acids and flavan-3-ols were released during digestion, but flavonols and flavones degraded

Phenolic compounds from the CSF generally showed a low release during the oral phase (**Tables 2–3**). Gallic acid was the primary hydroxybenzoic acid (72.0 μg g^{-1}), reaching a retention index of 45.0%. Protocatechuic acid was released by 40.6% (26.6 μg g^{-1}). Some NPAs, such as *N*-coumaroyl-L-aspartate (*cis* and *trans*), were utterly released. On the contrary, other compounds, such as *N*-caffeoyl-L-aspartate, were not detected. However, *N*-coumaroyl-L-tyrosine, not identified in the ND-CSF, was detected during the oral phase. A minor flavonoid fraction was released, except for epicatechin, which reached an oral retention index of 113.9%.

Theobromine and caffeine concentration was lower in the oral phase than before digestion (44.9 and 33.0%, respectively). In the CSE, phenolic compounds and methylxanthines were highly released during the oral phase, obtaining maximum retention indexes (82.7 to 127.7%) (**Table 3**). In general terms, oral retention indexes were significantly ($p < 0.001$) higher in the CSE than in the CSF (**Supplementary Table 1**).

In the CSF, the concentration of hydroxybenzoic acids and NPAs did not significantly change the gastric phase compared to the oral phase. There was no additional release of the above compounds, except for protocatechuic acid, which increased 2.0-fold (**Table 2**). However, some flavonoids, such as catechin and epicatechin, reached 86.8 and 153.8% retention indexes, respectively (**Table 3**). The concentration of methylxanthines increased during the gastric phase, achieving a retention index of 72.7% for theobromine and 59.6% for caffeine. On the contrary, in the CSE, gallic and protocatechuic acids decreased by 31.2–51.4 % compared to the oral phase. The concentration of catechin and epicatechin increased to achieve a gastric retention index of 121.7 and 173.4%, respectively. Flavan-3-ols exhibited a 1.3-fold higher ($p < 0.05$) gastric retention index in the CSE than in the CSF (**Supplementary Table 1**). Regarding methylxanthines, the concentration of theobromine decreased by 21.9%, and caffeine levels did not vary.

Gallic acid concentration in the CSF increased 1.7-fold in the intestinal phase, while the protocatechuic acid did not show significant changes (**Table 2**). NPAs exhibited different patterns. *N*-Coumaroyl-L-aspartate (*cis* and *trans*) reached a bioaccessibility of 123.9– 314.4%, whereas *N*-caffeoyl-L-DOPA *cis* decreased by 47.1% (**Table 3**). Catechin bioaccessibility decreased by 28.2% compared to the gastric phase. The flavonol fraction was not detected during the intestinal phase. The concentration of caffeine and theobromine increased during the intestinal phase (1.2–1.3-fold). In the CSE, phytochemicals underwent degradation during the intestinal phase, but hydroxybenzoic acids increased by 1.2-fold. Flavonols were not detected

in the CSE during the intestinal phase, and catechin and epicatechin decreased by 28.0% and 44.9%. Other NPAs, including *N*-coumaroyl-L-aspartate *cis*, *N*-caffeoyl-L-aspartate, and *N*-caffeoyl-L-DOPA *trans,* were also not detected during the intestinal phase. Likewise, the concentration of methylxanthines decreased (theobromine by 29.0% and caffeine by 20.8%) compared to the gastric phase. Therefore, their bioaccessibility declined to 48.0 and 58.7%, respectively.

During the colonic phase, the concentration of gallic (165.6 μg g^{-1}) and protocatechuic (107.5 μg g^{-1}) acids increased significantly in the CSF (1.4-fold and 2.3-fold higher than in the intestinal phase) (**Table 2**). These compounds reached a colonic bioaccessibility of 103.4 and 164.1%, respectively (**Table 3**). Among the NPAs, only *N*-caffeoyl-L-DOPA *cis* experienced an increase in its bioaccessibility. The catechin concentration increased during the colonic phase (142.9% of bioaccessibility). Theobromine concentration increased by 7.9% after the colonic phase (97.1% bioaccessibility), but no variations were observed in the caffeine concentration. In the CSE, the concentration of gallic acid increased significantly, displaying a high bioaccessibility (90.2%). *N*-coumaroyl-L-aspartate *trans* bioaccessibility increased to 125.0%. In opposition to the CSF, *N*-caffeoyl-L-DOPA *cis* decreased, reaching 64.2% colonic bioaccessibility. The concentration of flavan-3-ols experienced a sharp increase of 43.5% for catechin and 76.0% for epicatechin, achieving a bioaccessibility of 155.0 and 397.6%, respectively. Theobromine and caffeine concentration also increased during the colonic stage, exhibiting a noteworthy bioaccessibility of 73.7 and 73.0%, respectively.

3.5. Protocatechuic acid, epicatechin, and methylxanthines were the most bioavailable compounds

After the intestinal stage of simulated digestion of CSF and CSE, the potential bioavailability of phenolic compounds and methylxanthines was determined using two *in silico* models (C2A

and HIA) (**Tables 2–3** and **Supplementary Table 1**). The C2A model provided higher bioavailability values than the HIA model for hydroxybenzoic acids (1.2 to 1.4-fold), methylxanthines (1.1-fold), and most NPAs (1.2-fold). Conversely, the HIA model showed higher bioavailability values than the C2A model for *N*-caffeoyl-L-DOPA, *cis* (1.1-fold), and flavonoids (1.5 to 1.7-fold) (**Tables 2-3**).

In the CSF, the most bioavailable phenolic compounds were the hydroxybenzoic acids (35.7–46.3%) (**Supplementary Table 1**), highlighting protocatechuic acid (44.7–61.8%) (**Table 3**). Although, in general, NPAs did not show a high bioavailability (21.6–24.7%), some of them exhibited an extremely high bioavailability, such as *N*-coumaroyl-L-aspartate *cis* (60.7–74.3%) and *trans* (154.0–188.6%). Among flavonoids, epicatechin reached a bioavailability of 92.2%. Methylxanthines were highly bioavailable (76.8–79.5%). In the CSE, the most bioavailable phenolic compounds were also hydroxybenzoic acids (34.9–46.2%), especially protocatechuic acid, which could reach 83.7% of bioavailability. However, the NPAs presented low bioavailability after the intestinal phase of digestion (less than 6.3 %). The flavonoid fraction reached a bioavailability of 43.8%, and the methylxanthines showed a bioavailability of 42.8–43.5% (**Supplementary Table 1**). The results showed that NPAs (3.4–4.8-fold, $p < 0.001$) and methylxanthines (1.8-fold, $p < 0.001$) were more bioavailable in the CSF than in the CSE. From the results, hydroxybenzoic acids appeared to be more bioavailable than NPAs, and flavonoids were less bioavailable than phenolic acids (**Supplementary Table 1**).

3.6. The cocoa shell flour matrix protected phenolic compounds during digestion

The behavior of CSF and CSE throughout gastrointestinal digestion was visualized using principal component and hierarchical cluster analysis coupled to a heatmap (**Figure 3**). We observed 13 different factors explaining the variability among samples. The two principal

components could explain 84.4% of the whole variability (**Figure 3A**). The first component (64.9% variability) predominantly included the effects of theobromine, clovamide, vicenin, *N*-coumaroyl- and *N*-caffeoyl-L-aspartate, and gallic and protocatechuic acids. The second component (19.5% variability) included ABTS and epicatechin. Phenolic compounds and methylxanthines in the CSF were sequentially released from the matrix from the oral to the colonic phase. The antioxidant capacity of CSF-released compounds also increased over digestion. Inversely, phytochemicals in the CSE were fully bioaccessible in the oral phase and were gradually degraded from the gastric to the colonic phase. Hence, the poor absorption of phenolic compounds from both matrices yielded similar bioavailable fractions. We can perceive that the bioavailable fractions of CSF and CSE were similar, notwithstanding their initial composition. Complementarily, the hierarchical cluster analysis depicted the grouping of the CS digested fractions according to their composition (**Figure 3B**). The first group included non-digested CSE and oral and gastric fractions. The second group clustered the colonic and the bioavailable phases of CSF and CSE, and the third group contained CSF samples from the non-digested to the intestinal phase. Altogether, multivariate analysis demonstrated that gastrointestinal digestion transformed the CS composition matrix dependently.

3.7. Colonic biotransformation of non-absorbed phenolic compounds could generate smaller metabolites

Phenolic compound colonic metabolism by the gut microbiota was studied *in silico,* and the main routes involved in the biotransformation of the phenolics from the CS were summarized in **Figure 4**. The flavan-3-ols catechin and epicatechin are transformed into 5-(3,4-dihydroxyphenyl)-γ-valerolactone by C-ring fission. 5-(3,4-Dihydroxyphenyl)-γ-valerolactone can be dehydroxylated to 5-(3'-hydroxyphenyl)-gamma-valerolactone and 5-(4'-hydroxyphenyl)-gamma-valerolactone, and hydrolyzed to 4-hydroxy-5-(3,4-

dihydroxyphenyl)valeric acid. After further transformations (hydrolysis, dehydroxylation, and oxidation), flavan-3-ols yielded protocatechuic, 5-phenylvaleric, 3-(3-hydroxyphenyl)propionic, and 3-(4-hydroxyphenyl)propionic acids. Protocatechuic acid, which may also come from the dehydroxylation of gallic acid, can be dehydroxylated, yielding hydrocinnamic, *p*-salicylic, and *m*-salicylic acids. 5-Phenylvaleric acid can be oxidized to hydrocinnamic acid. In turn, 3-(3-hydroxyphenyl)propionic, and 3-(4-hydroxyphenyl)propionic acids can be oxidized, generating *m*-salicylic and *p*-salicylic acids, respectively, or dehydroxylated, producing hydrocinnamic acid. The smallest metabolite produced from the colonic metabolism of flavan-3-ols might be benzoic acid. Vicenin-2 was transformed by hydrolysis into apigenin 6-*C*-glucoside and subsequently hydrolyzed into apigenin. Apigenin was transformed into naringenin and chrysin by reduction and dehydroxylation, respectively. After further transformations involving reduction, dehydroxylation, and C-ring fission reactions, flavones yielded different chalcones. *p*-Salicylic and hydrocinnamic acids were the primary metabolites obtained from the colonic metabolism of flavones, ultimately generating benzoic acid. NPAs (*N*-caffeoyl-L-DOPA, *N*-coumaroyl-L-tyrosine, and *N*-coumaroyl-L-aspartate) experienced reduction and dehydroxylation reactions, producing *N*-3,3- and *N*-3,4-hydroxypheylpropanoyl-amino acids, and *N*-cinnamoyl-amino acids, and, ultimately obtaining *N*-hydroxycinnamoyl-amino acids. Colonic biotransformation by the gut microbiota could favorably influence the absorption of some phenolic compounds from the CS due to their transformation into different compounds than those found in the intestinal phase of the simulated digestion. The phytochemical diversity of the CS is then influenced by the colonic biotransformation yielding new metabolites. Distinct routes generated small, low molecular weight metabolites with increased potential absorption in both C2A and HIA models (**Supplementary Figure 1**). Flavan-3-ols exhibited a potential intestinal absorption of less than 55%. Vicenin-2 could be absorbed in less than 19%. NPAs could be

absorbed around 15–60%. In contrast, some low molecular weight metabolites obtained after colonic biotransformation of the non-absorbed compounds (i.e., *p*-salicylic, *m*-salicylic, and benzoic acids) could be fully absorbed in the intestine (**Supplementary Table 2**). Consequently, the biotransformation produced during colonic metabolism could play an essential role in phenolic absorption since smaller and maybe more adsorbable phenolic metabolites could be produced by intestinal biotransformation of non-absorbed phenolic compounds.

4. Discussion

The CS is rich in dietary fiber and contains bioactive compounds such as phenolic compounds and methylxanthines (Rojo-Poveda et al., 2021). In this sense, the revalorization of the CS can represent a strategy for a transition toward a circular and environmentally sustainable production model (Garcia et al., 2022). This study represents an advance in the knowledge of the bioaccessibility and potential bioavailability of phenolic compounds and methylxanthines from the CS under simulated gastrointestinal conditions.

The CS exhibited a large concentration of methylxanthines (theobromine and caffeine) and phenolic compounds. The hydroxycinnamic acid fraction was composed of NPAs, mostly *N*-caffeoyl-L-aspartate, aligning with previous studies, where this amino derivative was found to be the most abundant NPA in cocoa beans (Rojo-Poveda et al., 2021). These hydroxycinnamate derivatives are found exclusively in cocoa products, identified in all parts of the fruit, and related to the astringent properties (Lechtenberg et al., 2012). The flavonoid fraction was notable for its high catechin content, while a minor fraction of flavonols and flavones was also detected. Previous studies support our data, although other authors also identified a fraction of proanthocyanidins (Rojo-Poveda et al., 2021).

The concentration of phenolic compounds and methylxanthines was higher in the CSE than in the CSF. However, they exhibited similar bioactive compound profiles. The heat-assisted extraction produced the solubilization of phytochemicals and, perhaps, chemical transformations bringing out new compounds (Antony & Farid, 2022). The high insoluble dietary fiber content in the CSF could be another reason for the absence of some compounds in the non-digested CSF. Phenolic compounds are frequently strongly linked to cell wall polysaccharides, compromising their bioaccessibility (Zhu, 2018). The CSF can slowly and continuously release phenolic compounds under the acidic, alkaline, and enzymatic conditions of the *in vitro* simulated digestion, which could maintain a higher phenolic concentration beneficial for human health. However, due to the absence of a food matrix, there were no impediments to releasing and detecting phenolic compounds from the CSE during each step of the simulated gastrointestinal digestion. Different phenolic compounds were released from the food matrix, formed, or transformed during gastrointestinal digestion. The digestive system could increase the antioxidant capacity of phenolic compounds by changing their molecular weight and chemical structure during simulated digestion (Seraglio et al., 2017; Zhu et al., 2021). Potential interactions between antioxidant compounds and dietary constituents can also lead to complexes formations and, consequently, modify phenolic compounds' bioaccessibility and antioxidant properties (Seraglio et al., 2017).

The individual phenolic compounds showed different stability in the digestion process. During the oral phase, phenolics and methylxanthines in the CSE were free, whereas in the CSF, they were mainly bound to the food matrix. Although saliva contains enzymes that can enhance the release of some compounds from its matrix, generally, no significant changes occur since the food remains in the oral cavity for a short time. During the gastric phase, the acid pH, the enzymatic hydrolysis by pepsin, the particle size reduction, and the residence time of the food in the stomach enable the degradation of food compounds and facilitate the release of bound

phenolic compounds. However, the acid pH in the stomach causes the degradation of flavonoid oligomers to smaller units (Shu et al., 2019). Phenolic compounds are susceptible to alkaline conditions in the small intestine. Losses in phenolic content after intestinal digestion may be associated with the transformation and degradation of phenolic compounds suffered under these conditions. Conversely, the increment of certain phenolic compounds during the intestinal phase can be the consequence of the conversion of some compounds, such as gallic acid, which could indicate its dehydroxylation for the generation of protocatechuic acid (Jara-Palacios et al., 2018). Furthermore, intestinal enzymes can act on the matrix, facilitating the release of phenolic compounds (Wojtunik-Kulesza et al., 2020). Then again, the food matrix may also play an essential role in the bioaccessibility of the compounds, as phenolic compounds may be bound to some components, such as dietary fiber, proteins and fats, which may protect them from possible degradative processes (Mandalari et al., 2016).

Although most phenolic compounds presented high bioaccessibility, low bioavailability was estimated *in silico*. The absorption of phenolic compounds occurs mainly in the small intestine by passive diffusion or transporters located in the enterocyte membrane. The molecular weight, lipophilicity, or stereochemistry of phenolic compounds may influence their absorption. Therefore, phenolic compounds with low molecular weight and high lipophilicity are more absorbed (Cosme et al., 2020). Gallic acid exhibited a high bioaccessibility but lower bioavailability in both CSF and CSE. Several human studies have shown that this acid is quickly absorbed in the small intestine, exhibiting excellent bioavailability (Kaliora et al., 2013). Protocatechuic acid and *trans*-*N*-coumaroyl-L-aspartate showed the highest bioavailability among the phenolic acids studied. Protocatechuic can be efficiently absorbed by intestinal epithelial cells (Song et al., 2020). In turn, NPAs contain an amide bond which is more stable than the ester bond and therefore avoids intestinal hydrolysis. They may be absorbed through the small intestinal epithelium reaching the bloodstream, although their bioavailability has not

been extensively studied (Oracz et al., 2020). The bioavailability of flavonoids, especially flavan3-ols, can be affected by many factors, including stereoisomerism. Several studies have also shown that the bioavailability of (−)-epicatechin is higher than (+)-catechin (Di Pede, Mena, Bresciani, Achour, et al., 2022). Methylxanthines stood out for their high bioavailability, especially in the CSF. Caffeine is characterized by its rapid and complete absorption in the gastrointestinal tract (Barcelos et al., 2020). Theobromine is also absorbed in the small intestine. Both methylxanthines diffuse passively through enterocytes, reaching the bloodstream (Ellam & Williamson, 2013).

The CS has been described as a high dietary fiber matrix (Panak Balentić et al., 2018). Then, the compounds tightly bound to the dietary fiber matrix, and not released during the previous stages, are susceptible to being liberated by the colonic microbiota action. Hydroxybenzoic acids were noted for their high bioaccessibility during the colonic phase. Several studies reported more concentration of protocatechuic acid at the end of the colonic phase than the concentration ingested. Protocatechuic acid is a colonic metabolite that can be generated by the action of colonic microbiota from other compounds, such as flavonoids, including anthocyanins and flavan-3-ols (Zheng et al., 2019). Both the host and the gut microbiota metabolize large molecules into smaller compounds that can be further digested or absorbed. Microorganisms predominantly use hydrolytic and reductive processes to metabolize phytochemicals, many of which are unique to the gut microbiota species. In sharp contrast, the host metabolism tends to be more oxidative and conjugative. Many reactions, however, occur concurrently between the host and microbiota-mediated biotransformation. As a result, the combined metabolisms of the host and microbiota generate metabolites that would not be synthesized by the host alone, which can significantly modify phytochemicals' bioactivities within the human body (Koppel et al., 2017). The structural transformations of NPAs by colonic microbiota have not been extensively studied. Although some authors report that NPAs could be transformed into phenolic acids

575 during microbial degradation, there is no evidence that NPAs are hydrolyzed, conjugated, or
576 enzymatically degraded by intestinal microbes (Oracz et al., 2020). However, flavonoids that
577 reach the colon can be metabolized by enzymes of the gut microbiota, which remove the
578 glycosides producing flavonoid aglycones (Al-Ishaq et al., 2021). Flavan-3-ols underwent
579 biotransformation by gut microbiota, leading to phenyl-γ-valerolactones and related
580 phenylvaleric acids (Di Pede, Mena, Bresciani, Almutairi, et al., 2022). Low molecular weight
581 phenolic metabolites of the colonic microbiota are the ones that can be absorbed and reach the
582 bloodstream and peripheral tissues eliciting bioactive effects (Carregosa et al., 2022).

583

584 ## 5. Conclusion

585 This research presented new knowledge into the gastrointestinal fate of phenolic compounds
586 and methylxanthines in the CS under simulated conditions. For the first time, we investigated
587 the bioaccessibility, potential bioavailability, and colonic biotransformation of phytochemicals
588 from the CS. The CS matrix type played a significant role in the release of phytochemicals. The
589 CSF matrix protected phenolic compounds, turning them more bioaccessible and bioavailable
590 than those from the CSE. The digestion prompted the release of hydroxybenzoic acids and
591 flavan-3-ols but triggered the degradation of flavonols and flavones. Protocatechuic acid,
592 epicatechin, and methylxanthines were the main compounds absorbed in the intestine after *in*
593 *vitro* gastrointestinal digestion in both samples. In addition, new insights gained on the non-
594 bioaccessible phenolic compounds by *in silico* estimation of their colonic metabolism revealed
595 that the gut microbiota could biotransform non-absorbed phenolic compounds into diverse
596 lower molecular weight compounds. The results further strengthen the CS as a source of
597 potentially bioaccessible, bioavailable, and active phytochemicals, which might exert potential
598 health-promoting properties in the organism.

Credit and authorship contribution statement

Silvia Cañas: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization; **Miguel Rebollo-Hernanz:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – review & editing, Visualization; **Cheyenne Braojos:** Investigation; **Vanesa Benítez:** Conceptualization; **Rebeca Ferreras-Charro:** Formal analysis, Investigation; **Montserrat Dueñas:** Formal analysis; **Yolanda Aguilera:** Writing – review & editing; **María A. Martín-Cabrejas:** Conceptualization, Writing – original draft, Writing – review & editing, Supervision, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by the COCARDIOLAC project from the Spanish Ministry of Science and Innovation (RTI 2018-097504-B-I00) and the Excellence Line for University Teaching Staff within the Multiannual Agreement between the Community of Madrid and the UAM (2019-2023). Escalera de Excelencia CLU-2018-04 cofunded by the P.O. FEDER of Castilla y León 2014-2020 Spain. M. Rebollo-Hernanz received funding from the FPU program

of the Ministry of Universities for his predoctoral fellowship (FPU15/04238) and Margarita Salas Contract (CA1/RSUE/2021-00656).

References

Al-Ishaq, R. K., Liskova, A., Kubatka, P., & Büsselberg, D. (2021). Enzymatic metabolism of flavonoids by gut microbiota and its impact on gastrointestinal cancer. *Cancers*, *13*(16). https://doi.org/10.3390/cancers13163934

Antony, A., & Farid, M. (2022). Effect of Temperatures on Polyphenols during Extraction. *Applied Sciences*, *12*(4). https://doi.org/10.3390/app12042107

Barcelos, R. P., Lima, F. D., Carvalho, N. R., Bresciani, G., & Royes, L. F. (2020). Caffeine effects on systemic metabolism, oxidative-inflammatory pathways, and exercise performance. *Nutrition Research*, *80*, 1–17. https://doi.org/10.1016/j.nutres.2020.05.005

Boeckx, P., Bauters, M., & Dewettinck, K. (2020). Poverty and climate change challenges for sustainable intensification of cocoa systems. *Current Opinion in Environmental Sustainability*, *47*, 106–111. https://doi.org/10.1016/J.COSUST.2020.10.012

Brodkorb, A., Egger, L., Alminger, M., Alvito, P., Assunção, R., Ballance, S., Bohn, T., Bourlieu-Lacanal, C., Boutrou, R., Carrière, F., Clemente, A., Corredig, M., Dupont, D., Dufour, C., Edwards, C., Golding, M., Karakaya, S., Kirkhus, B., Le Feunteun, S., … Recio, I. (2019). INFOGEST static in vitro simulation of gastrointestinal food digestion. *Nature Protocols*, *14*(4), 991–1014. https://doi.org/10.1038/s41596-018-0119-1

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

Carregosa, D., Pinto, C., Ávila-Gálvez, M. Á., Bastos, P., Berry, D., & Santos, C. N. (2022). A look beyond dietary (poly)phenols: The low molecular weight phenolic metabolites and their concentrations in human circulation. *Comprehensive Reviews in Food Science and Food Safety*, *21*(5), 3931–3962. https://doi.org/10.1111/1541-4337.13006

Cinar, Z. Ö., Atanassova, M., Tumer, T. B., Caruso, G., Antika, G., Sharma, S., Sharifi-Rad, J., & Pezzani, R. (2021). Cocoa and cocoa bean shells role in human health: An updated review. *Journal of Food Composition and Analysis*, *103*, 104115. https://doi.org/10.1016/J.JFCA.2021.104115

Cosme, P., Rodríguez, A. B., Espino, J., & Garrido, M. (2020). Plant phenolics: Bioavailability as a key determinant of their potential health-promoting applications. *Antioxidants*, *9*(12), 1–20. https://doi.org/10.3390/antiox9121263

Di Pede, G., Mena, P., Bresciani, L., Achour, M., Estruch, R., Landberg, R., Kulling, S. E., Lamuela-ravent, R. M., Wishart, D., Rodriguez-Mateos, A., Crozier, A., Manach, C., Del, D., Di Pede, G., Mena, P., Bresciani, L., Achour, M., Lamuela-Raventós, R. M., Estruch, R., … Del Rio, D. (2022). Molecular Aspects of Medicine Revisiting the bioavailability of flavan-3-ols in humans : A systematic review and comprehensive data analysis. *Molecular Aspects of Medicine*, *July*, 101146. https://doi.org/10.1016/j.mam.2022.101146

Di Pede, G., Mena, P., Bresciani, L., Almutairi, T. M., Del Rio, D., Clifford, M. N., & Crozier, A. (2022). Human colonic catabolism of dietary flavan-3-ol bioactives. *Molecular Aspects of Medicine*, 101107. https://doi.org/https://doi.org/10.1016/j.mam.2022.101107

Djoumbou-Feunang, Y., Fiamoncini, J., Gil-de-la-Fuente, A., Greiner, R., Manach, C., & Wishart, D. S. (2019). BioTransformer: A comprehensive computational tool for small molecule metabolism prediction and metabolite identification. *Journal of Cheminformatics*,

11(1), 1–25. https://doi.org/10.1186/S13321-018-0324-5/FIGURES/9

Ellam, S., & Williamson, G. (2013). Cocoa and human health. *Annual Review of Nutrition*, *33*, 105–128.

Garcia, I. G., Simal-Gandara, J., & Gullo, M. (2022). Advances in Food, Bioproducts and Natural Byproducts for a Sustainable Future: From Conventional to Innovative Processes. *Applied Sciences*, *12*(6). https://doi.org/10.3390/app12062893

Jara-Palacios, M. J., Gonçalves, S., Hernanz, D., Heredia, F. J., & Romano, A. (2018). Effects of in vitro gastrointestinal digestion on phenolic compounds and antioxidant activity of different white winemaking byproducts extracts. *Food Research International*, *109*, 433–439. https://doi.org/10.1016/j.foodres.2018.04.060

Kaliora, A. C., Kanellos, P. T., & Kalogeropoulos, N. (2013). Gallic acid bioavailability in humans. In *Handbook on Gallic Acid: Natural Occurrences, Antioxidant Properties and Health Implications* (pp. 301–312).

Koppel, N., Rekdal, V. M., & Balskus, E. P. (2017). Chemical transformation of xenobiotics by the human gut microbiota. *Science*, *356*(6344), 1246–1257. https://doi.org/10.1126/science.aag2770

Lechtenberg, M., Henschel, K., Liefländer-Wulf, U., Quandt, B., & Hensel, A. (2012). Fast determination of N-phenylpropenoyl-l-amino acids (NPA) in cocoa samples from different origins by ultra-performance liquid chromatography and capillary electrophoresis. *Food Chemistry*, *135*(3). https://doi.org/10.1016/j.foodchem.2012.06.006

Mandalari, G., Vardakou, M., Faulks, R., Bisignano, C., Martorana, M., Smeriglio, A., & Trombetta, D. (2016). Food matrix effects of polyphenol bioaccessibility from almond skin during simulated human digestion. *Nutrients*, *8*(9). https://doi.org/10.3390/nu8090568

690 Martini, S., Conte, A., & Tagliazucchi, D. (2018). Comprehensive evaluation of phenolic
691 profile in dark chocolate and dark chocolate enriched with Sakura green tea leaves or turmeric
692 powder. *Food Research International*, *112*, 1–16.
693 https://doi.org/10.1016/J.FOODRES.2018.06.020

694 Oracz, J., Nebesny, E., & Żyżelewicz, D. (2019). Identification and quantification of free and
695 bound phenolic compounds contained in the high-molecular weight melanoidin fractions
696 derived from two different types of cocoa beans by UHPLC-DAD-ESI-HR-MSn. *Food*
697 *Research International*, *115*, 135–149. https://doi.org/10.1016/j.foodres.2018.08.028

698 Oracz, J., Nebesny, E., Zyzelewicz, D., Budryn, G., & Luzak, B. (2020). Bioavailability and
699 metabolism of selected cocoa bioactive compounds: A comprehensive review. *Critical Reviews*
700 *in Food Science and Nutrition*, *60*(12), 1947–1985.
701 https://doi.org/10.1080/10408398.2019.1619160

702 Panak Balentić, J., Ačkar, Đ., Jokić, S., Jozinović, A., Babić, J., Miličević, B., Šubarić, D., &
703 Pavlović, N. (2018). Cocoa Shell: A By-Product with Great Potential for Wide Application.
704 *Molecules*, *23*(6), 1404. https://doi.org/10.3390/molecules23061404

705 Papillo, V. A., Vitaglione, P., Graziani, G., Gokmen, V., & Fogliano, V. (2014). Release of
706 antioxidant capacity from five plant foods during a multistep enzymatic digestion protocol.
707 *Journal of Agricultural and Food Chemistry*. https://doi.org/10.1021/jf500695a

708 Rebollo-Hernanz, M., Aguilera, Y., Martin-Cabrejas, M. A., & Gonzalez de Mejia, E. (2022).
709 Phytochemicals from the Cocoa Shell Modulate Mitochondrial Function, Lipid and Glucose
710 Metabolism in Hepatocytes via Activation of FGF21/ERK, AKT, and mTOR Pathways.
711 *Antioxidants*, *11*(1), 136. https://doi.org/10.3390/antiox11010136

712 Rebollo-Hernanz, M., Zhang, Q., Aguilera, Y., Martín-Cabrejas, M. A., & de Mejia, E. G.
713 (2019). Cocoa Shell Aqueous Phenolic Extract Preserves Mitochondrial Function and Insulin

714 Sensitivity by Attenuating Inflammation between Macrophages and Adipocytes In Vitro.
715 *Molecular Nutrition and Food Research*, *63*(10), 1801413.
716 https://doi.org/10.1002/mnfr.201801413

717 Rebollo-Hernanz, M., Aguilera, Y., Herrera, T., Cayuelas, L. T., Dueñas, M.,
718 Rodríguez-Rodríguez, P., Ramiro-Cortijo, D., Arribas, S. M., & Martín-Cabrejas, M. A. (2020).
719 Bioavailability of melatonin from lentil sprouts and its role in the plasmatic antioxidant status
720 in rats. *Foods*, *9*(3), 330. https://doi.org/10.3390/foods9030330

721 Rebollo-Hernanz, M., Cañas, S., Taladrid, D., Bartolomé, B., Aguilera, Y., & Martin-Cabrejas,
722 M. A. (2021). Extraction of phenolic compounds from cocoa shell: modeling using response
723 surface methodology and artificial neural networks. *Separation and Purification Technology*,
724 *270*, 118779. https://doi.org/10.1016/j.seppur.2021.118779

725 Rodrigues, D. B., Marques, M. C., Hacke, A., Loubet Filho, P. S., Cazarin, C. B. B., & Mariutti,
726 L. R. B. (2022). Trust your gut: Bioavailability and bioaccessibility of dietary compounds.
727 *Current Research in Food Science*, *5*, 228–233. https://doi.org/10.1016/J.CRFS.2022.01.002

728 Rodríguez-Rodríguez, P., Ragusky, K., Phuthong, S., Ruvira, S., Ramiro-Cortijo, D., Cañas, S.,
729 Rebollo-Hernanz, M., Morales, M. D., López de Pablo, Á. L., Martín-Cabrejas, M. A., &
730 Arribas, S. M. (2022). Vasoactive Properties of a Cocoa Shell Extract: Mechanism of Action
731 and Effect on Endothelial Dysfunction in Aged Rats. *Antioxidants*, *11*(2).
732 https://doi.org/10.3390/antiox11020429

733 Rojo-Poveda, O., Zeppa, G., Ferrocino, I., Stévigny, C., & Barbosa-Pereira, L. (2021).
734 Chemometric classification of cocoa bean shells based on their polyphenolic profile determined
735 by rp-hplc-pda analysis and spectrophotometric assays. *Antioxidants*, *10*(10), 1533.
736 https://doi.org/10.3390/antiox10101533

737 Seraglio, S. K. T., Valese, A. C., Daguer, H., Bergamo, G., Azevedo, M. S., Nehring, P.,

738 Gonzaga, L. V., Fett, R., & Costa, A. C. O. (2017). Effect of in vitro gastrointestinal digestion
739 on the bioaccessibility of phenolic compounds, minerals, and antioxidant capacity of Mimosa
740 scabrella Bentham honeydew honeys. *Food Research International*, *99*, 670–678.
741 https://doi.org/10.1016/j.foodres.2017.06.024

742 Shu, Y., Li, J., Yang, X., Dong, X., & Wang, X. (2019). Effect of particle size on the
743 bioaccessibility of polyphenols and polysaccharides in green tea powder and its antioxidant
744 activity after simulated human digestion. *Journal of Food Science and Technology*, *56*(3),
745 1127–1133. https://doi.org/10.1007/s13197-019-03573-4

746 Song, J., He, Y., Luo, C., Feng, B., Ran, F., Xu, H., Ci, Z., Xu, R., Han, L., & Zhang, D. (2020).
747 New progress in the pharmacology of protocatechuic acid: A compound ingested in daily foods
748 and herbs frequently and heavily. *Pharmacological Research*, *161*.
749 https://doi.org/10.1016/j.phrs.2020.105109

750 Wang, N. N., Dong, J., Deng, Y. H., Zhu, M. F., Wen, M., Yao, Z. J., Lu, A. P., Wang, J. B.,
751 & Cao, D. S. (2016). ADME Properties Evaluation in Drug Discovery: Prediction of Caco-2
752 Cell Permeability Using a Combination of NSGA-II and Boosting. *Journal of Chemical*
753 *Information and Modeling*, *56*(4), 763–773.
754 https://doi.org/10.1021/ACS.JCIM.5B00642/SUPPL_FILE/CI5B00642_SI_001.XLSX

755 Wojtunik-Kulesza, K., Oniszczuk, A., Oniszczuk, T., Combrzyński, M., Nowakowska, D., &
756 Matwijczuk, A. (2020). Influence of In Vitro Digestion on Composition, Bioaccessibility and
757 Antioxidant Activity of Food Polyphenols-A Non-Systematic Review. *Nutrients*, *12*(5).
758 https://doi.org/10.3390/nu12051401

759 Zheng, J., Xiong, H., Li, Q., He, L., Weng, H., Ling, W., & Wang, D. (2019). Protocatechuic
760 acid from chicory is bioavailable and undergoes partial glucuronidation and sulfation in healthy
761 humans. *Food Science and Nutrition*, *7*(9). https://doi.org/10.1002/fsn3.1168

762 Zhu, F. (2018). Interactions between cell wall polysaccharides and polyphenols. *Critical*
763 *Reviews in Food Science and Nutrition*, *58*(11), 1808–1831.
764 https://doi.org/10.1080/10408398.2017.1287659

765 Zhu, Y., Yang, S., Huang, Y., Huang, J., & Li, Y. (2021). Effect of in vitro gastrointestinal
766 digestion on phenolic compounds and antioxidant properties of soluble and insoluble dietary
767 fibers derived from hulless barley. *Journal of Food Science*, *86*(2), 628–634.
768 https://doi.org/10.1111/1750-3841.15592

769

Figure Captions

Figure 1. Effect of *in vitro* digestion of cocoa shell flour (CSF) and cocoa shell extract (CSE) on the total phenolic content (**A**), ABTS (**B**), and FRAP (**C**) antioxidant capacity from the digested (D) and non-digested (ND) cocoa shell, including free (FND), bound (BND), and total (TND) phenolic fractions. Behavior of total phenolic acids (PA) (hydroxybenzoic acids (HBA) + *N*-phenylpropenoyl-L-amino acids (NPAs)) (**D**), flavonoids (FVD) (flavan-3-ols (F3L) + flavonols (FVL)) (**E**), and total phenolics (TP) and methylxanthines (MTX) (**F**) throughout the simulated gastrointestinal digestion phases. The results are reported as mean ± SD (n = 3). Points with different letters significantly ($p < 0.05$) differ according to ANOVA and Tukey's multiple range test (between digestion phases, within the same sample).

Figure 2. Impact of *in vitro* digestion on the release of total phenolics (**A**), phenolic acids (**B**), flavonoids (**C**), and methylxanthines (**D**) from the cocoa shell flour (CSF) and the cocoa shell extract (CSE). The results are reported as mean ± SD (n = 3). Bars different letters significantly ($p < 0.05$) differ according to ANOVA and Tukey's multiple range test. *T*-test comparisons were performed between the retention indexes and the bioaccessibility of CSF and CSE. Differences were considered significant at $p < 0.05$.

Figure 3. Biplot (scores of samples and load factors of each variable) of the principal component analysis (PCA) (**A**) and agglomerative hierarchical cluster analysis coupled to heatmap (from the lowest (■) to the highest (■) value for each parameter) (**B**) illustrating the behavior of phenolic compounds and methylxanthines from the cocoa shell during simulated gastrointestinal digestion. CSE: Cocoa Shell Extract; CSF: Cocoa Shell Flour; ND: Non-

793 Digested; OP: Oral Phase; GP: Gastric Phase; IP: Intestinal Phase; CP: Colonic Phase; C2A:
794 Caco-2 Absorption; HIA: Human Intestinal Absorption.

795

796 **Figure 4**. Proposed pathways involved in the colonic metabolism of phenolic compounds from
797 the cocoa shell predicted *in silico*. 3,3-HPPA: 3-(3-Hydroxyphenyl)propionic acid; 3,4-HPPA:
798 3-(4-Hydroxyphenyl)propionic acid; 3-H-5(4-HP)VA: 3-Hydroxy-5-(4-hydroxyphenyl)valeric
799 acid; 4-H-5(3,4-HP)VA: 4-Hydroxy-5-(3,4-dihydroxyphenyl)valeric acid; 4-H-5(3-HP)VA: 4-
800 Hydroxy-5-(3-hydroxyphenyl)valeric acid; 4-H-5-PVA: 4-Hydroxy-5-phenyl-valeric acid; 5-
801 (3,4-HP)VA: 5-(3,4-Dihydroxyphenyl)valeric acid; 5(3,4-HP)γ-VL: 5-(3,4-Dihydroxyphenyl)-
802 γ-valerolactone; 5-(3-HP)VA: 5-(3-Hydroxyphenyl)valeric acid; 5(3-HP)γ-VL: 5-(3'-
803 Hydroxyphenyl)-γ-valerolactone; 5-(4-HP)VA: 5-(4-Hydroxyphenyl)valeric acid; 5(4-HP)γ-
804 VL: 5-(4'-Hydroxyphenyl)-γ-valerolactone; 5-PVA: 5-Phenylvaleric acid; 5-P-γ-VL: 5-Phenyl-
805 γ-valerolactone; API: Apigenin; API6G: Isovitexin; BA: Benzoic acid; CAT: Catechin; CHRY:
806 Chrysin; CiA: Cinnamic acid; dhCA: Dihydrocaffeic acid; dhPBCH: 2',4',6'-
807 Trihydroxydihydrochalcone; EPI: Epicatechin; GA: Gallic acid; hCiA: Hydrocinnamic acid;
808 m-SA: 3-Hydroxybenzoic acid; N-3,3-HPP-AA: N-[3-(3-Hydroxyphenyl)propionyl]-AA; N-
809 3,4-HPP-AA: N-[3-(4-Hydroxyphenyl)propionyl]-AA; NAR: Naringenin; NARCH:
810 Naringenin chalcone; N-C-AA: N-Caffeoyl-AA; N-Ci-AA: N-Cinnamoyl-L-AA; N-Cou-AA:
811 N-Coumaroyl-AA; N-dhC-AA: N-Dihydrocaffeic-acid-AA; N-hCi-AA: N-
812 (Hydroxycinnamoyl)-AA; PB: Pinocembrin; PBCH: Pinocembrin chalcone; PCA:
813 Protocatechuic acid; p-CouA: p-Coumaric acid; p-SA: 4-Hydroxybenzoic acid ; VIC2:
814 Vicenin-2; α-RA: α-Resorcylic acid. Complete information on the phenolic compounds can be
815 found in the **Supplementary Table 2**.

816

817 **Figure 5.** An illustrated diagram integrating the experimental design and the main results and
818 conclusions retrieved from this study. The cocoa shell (flour and aqueous extract) was digested
819 *in vitro,* and the phytochemicals released were evaluated using HPLC-DAD-MS/MS (phenolic
820 compounds, methylxanthines, and *N*-phenylpropenoyl-L-amino acids). Metabolites'
821 bioaccessibility and potential bioavailability were calculated, and the colonic biotransformation
822 of phenolic compounds and amino derivatives was studied *in silico*.

823

Table 1. Identification of phenolic compounds and methylxanthines in the cocoa shell flour and extract by HPLC-DAD-MS.

Compound	R_t (min)	λ_{max} (nm)	Molecular ion $[M-H]^-$ (*m/z*)	Fragments MS^2	Tentative identification	Common name	CID
1	4.70	270	169	–	3,4,5-Trihydroxybenzoic acid	Gallic acid	370
2	6.00	272	–	–	3,7-Dimethylxanthine	Theobromine	5429
3	6.10	290, 294	153	109	3,4-Dihydroxybenzoic acid	Protocatechuic acid	72
4	6.55	323	294	179	*N*-Caffeoyl-L-aspartate	–	23658567
5	6.63	270, 329	278	235, 163	*N*-Coumaroyl-L-aspartate, *cis*	–	165368074
6	8.50	279	289	245, 205, 151, 137	(+)-Catechin	Catechin	9064
7	9.50	310	278	235, 163	*N*-Coumaroyl-L-aspartate, *trans*	–	68537088
8	10.18	279	289	245, 205, 151, 137	(−)-Epicatechin	Epicatechin	72276
9	12.79	335	593	413, 383, 353, 297, 283	Apigenin 6,8-di-*C*-glucoside	Vicenin-2	3084407
10	12.98	321	358	179	*N*-Caffeoyl-L-DOPA, *cis*	Clovamide, *cis*	6443790
11	13.47	273	–	–	1,3,7-Trimethylxanthine	Caffeine	2519
12	14.51	323	358	179	*N*-Caffeoyl-L-DOPA, *trans*	Clovamide, *trans*	6506968
13	22.06	354	463	301	Quercetin-3-*O*-hexoside	–	5280804
14	23.08	303	326	282, 163, 146	*N*-Coumaroyl-tyrosine	Dideoxyclovamide	15825666
15	24.79	355	463	301	Quercetin-3-*O*-pentoside	–	12309865

Table 2. Concentration of individual phenolic compounds and methylxanthines (μg g^{-1}) in non-digested and digested cocoa shell flour and extract, and its potential Caco-2 absorption and human intestinal absorption throughout the different phases of the simulated gastrointestinal digestion.

Compounds	ND	OP	GP	IP	CP	C2A	HIA
Cocoa shell flour							
Hydroxybenzoic acids							
Gallic acid	160.1 ± 4.3^{a}	72.0 ± 5.8^{c}	70.6 ± 3.9^{c}	118.4 ± 7.6^{b}	165.6 ± 15.3^{a}	63.9 ± 4.1^{cd}	51.2 ± 3.3^{d}
Protocatechuic acid	65.5 ± 5.8^{b}	26.6 ± 0.7^{e}	54.1 ± 0.1^{c}	46.4 ± 2.7^{cd}	107.5 ± 9.8^{a}	40.5 ± 2.4^{d}	29.3 ± 1.7^{e}
N-Phenylpropenoyl-L-amino acids							
N-Coumaroyl-L-aspartate *cis*	8.8 ± 0.4^{b}	9.0 ± 0.4^{b}	8.3 ± 0.6^{b}	10.9 ± 0.8^{a}	8.5 ± 1.0^{b}	6.6 ± 0.5^{c}	5.4 ± 0.4^{d}
N-Coumaroyl-L-aspartate *trans*	4.7 ± 0.5^{d}	5.8 ± 0.5^{d}	5.2 ± 0.5^{d}	14.8 ± 1.0^{a}	–	8.9 ± 0.6^{b}	7.3 ± 0.5^{c}
N-Coumaroyl-L-tyrosine	–	0.7 ± 0.1^{c}	0.5 ± 0.0^{c}	3.8 ± 0.3^{a}	4.1 ± 0.4^{a}	2.1 ± 0.2^{b}	1.8 ± 0.1^{b}
N-Caffeoyl-L-aspartate	51.4 ± 2.9	–	–	–	–	–	–
N-Caffeoyl-L-DOPA *cis*	8.9 ± 0.4^{a}	5.6 ± 0.5^{c}	5.9 ± 0.2^{c}	4.2 ± 0.6^{d}	7.0 ± 0.8^{b}	0.6 ± 0.1^{e}	1.6 ± 0.2^{e}
Flavan-3-ols							
Catechin	115.2 ± 8.1^{b}	40.3 ± 4.2^{d}	100.0 ± 2.1^{b}	71.8 ± 5.5^{c}	164.6 ± 13.7^{a}	23.6 ± 1.8^{e}	39.1 ± 3.0^{de}
Epicatechin	13.3 ± 1.3^{bc}	15.1 ± 1.2^{b}	20.5 ± 0.6^{a}	22.5 ± 1.5^{a}	–	7.4 ± 0.5^{d}	12.3 ± 0.8^{c}
Flavonols							
Quercetin 3-*O*-hexoside	2.9 ± 0.1^{a}	0.9 ± 0.1^{b}	1.0 ± 0.1^{b}	–	–	–	–
Quercetin 3-*O*-pentoside	3.4 ± 0.0^{a}	0.8 ± 0.1^{b}	0.6 ± 0.0^{c}	–	–	–	–
Methylxanthines							
Theobromine	5258.0 ± 48.8^{a}	2360.0 ± 19.8^{e}	3822.1 ± 410.9^{d}	4701.2 ± 128.4^{b}	5105.9 ± 31.3^{a}	4248.4 ± 116.0^{c}	4181.2 ± 114.2^{cd}
Caffeine	1693.8 ± 3.7^{a}	558.2 ± 0.7^{f}	1009.9 ± 10.6^{e}	1303.5 ± 9.9^{bc}	1329.1 ± 30.8^{b}	1276.1 ± 9.7^{c}	1161.3 ± 8.9^{d}

Cocoa shell extract							
Hydroxybenzoic acids							
Gallic acid	739.3 ± 22.4^{a}	641.1 ± 16.6^{b}	311.4 ± 24.5^{d}	391.7 ± 26.1^{c}	666.6 ± 63.5^{a}	211.3 ± 14.1^{e}	169.3 ± 11.3^{e}
Protocatechuic acid	348.8 ± 24.2^{b}	401.9 ± 7.3^{a}	276.7 ± 7.2^{c}	334.0 ± 23.1^{b}	310.3 ± 11.2^{bc}	291.8 ± 20.2^{c}	210.8 ± 14.6^{d}
N-Phenylpropenoyl-L-amino acids							
N-Coumaroyl-L-aspartate *cis*	62.2 ± 6.2^{a}	59.3 ± 0.1^{a}	63.6 ± 6.0^{a}	–	–	–	–
N-Coumaroyl-L-aspartate *trans*	44.6 ± 3.1^{bc}	48.6 ± 4.0^{b}	40.9 ± 0.2^{c}	16.5 ± 0.2^{d}	55.7 ± 3.0^{a}	9.9 ± 0.1^{e}	8.1 ± 0.1^{e}
N-Coumaroyl-L-tyrosine	7.7 ± 1.5^{bc}	9.8 ± 0.2^{b}	12.6 ± 1.5^{a}	6.3 ± 0.7^{c}	–	3.5 ± 0.4^{d}	2.9 ± 0.3^{d}
N-Caffeoyl-L-aspartate	191.3 ± 18.0^{ab}	211.2 ± 7.0^{a}	173.0 ± 20.1^{b}	–	–	–	–
N-Caffeoyl-L-DOPA *cis*	37.9 ± 1.2^{b}	37.9 ± 3.9^{b}	46.7 ± 4.9^{a}	30.7 ± 3.1^{c}	24.3 ± 0.0^{d}	4.7 ± 0.5^{f}	11.3 ± 1.2^{e}
N-Caffeoyl-L-DOPA *trans*	10.5 ± 0.3^{a}	10.7 ± 0.7^{a}	10.1 ± 0.3^{a}	–	–	–	–
Flavan-3-ols							
Catechin	460.5 ± 14.2^{c}	416.0 ± 38.2^{c}	560.5 ± 56.4^{b}	403.5 ± 2.1^{c}	713.8 ± 83.3^{a}	132.6 ± 0.7^{d}	219.7 ± 1.2^{d}
Epicatechin	34.6 ± 0.6^{c}	30.4 ± 4.2^{cd}	60.1 ± 5.1^{b}	33.1 ± 3.4^{c}	137.7 ± 14.9^{a}	10.9 ± 1.1^{e}	18.0 ± 1.9^{de}
Flavonols							
Quercetin 3-*O*-hexoside	14.2 ± 0.1^{a}	12.6 ± 0.5^{b}	12.2 ± 0.4^{b}	–	–	–	–
Quercetin 3-*O*-pentoside	13.1 ± 0.6^{a}	12.6 ± 1.2^{a}	13.1 ± 0.4^{a}	–	–	–	–
Flavones							
Vicenin-2	28.7 ± 0.1^{b}	27.8 ± 1.7^{b}	33.6 ± 0.5^{a}	19.4 ± 2.2^{c}	–	2.3 ± 0.3^{d}	3.4 ± 0.4^{d}
Methylxanthines							
Theobromine	26053.1 ± 1255.0^{a}	22529.8 ± 205.7^{b}	17597.6 ± 1259.4^{c}	12498.3 ± 490.4^{d}	19193.5 ± 862.9^{c}	11294.4 ± 443.2^{d}	11115.7 ± 436.1^{d}
Caffeine	339.9 ± 19.5^{a}	281.1 ± 7.7^{b}	252.0 ± 15.9^{b}	199.6 ± 16.5^{c}	248.3 ± 10.0^{b}	195.4 ± 16.2^{c}	177.8 ± 14.7^{c}

Results are reported as mean ± SD (n = 3). Mean values within rows followed by different superscript letters (a, b, c, d, e, f) are significantly different when subjected to Tukey's test ($p < 0.05$). ND: Non-Digested; OP: Oral Phase; GP: Gastric Phase; IP: Intestinal Phase; CP: Colonic Phase; C2A: Caco-2 Absorption; HIA: Human Intestinal Absorption; nd: non-detected; t: traces.

Table 3. Retention index, bioaccessibility, and potential bioavailability (%) of individual phenolic compounds and methylxanthines from cocoa shell flour and extract after simulated *in vitro* gastrointestinal digestion.

Compounds	Retention Index		Bioaccessibility		Bioavailability	
	OP	GP	IP	CP	C2A	HIA
Cocoa shell flour						
Hydroxybenzoic acid derivatives						
Gallic acid	$45.0 \pm 4.8^{c***}$	44.1 ± 3.6^{c}	$73.9 \pm 6.8^{b**}$	103.4 ± 12.3^{a}	$39.9 \pm 3.6^{c**}$	$32.0 \pm 2.9^{c**}$
Protocatechuic acid	$40.6 \pm 4.7^{d***}$	82.6 ± 7.4^{b}	$70.7 \pm 10.4^{bc*}$	$164.1 \pm 29.4^{a**}$	$61.8 \pm 9.1^{bcd*}$	$44.7 \pm 6.6^{cd*}$
N-Phenylpropenoyl-L-amino acids						
N-Coumaroyl-L-aspartate *cis*	101.5 ± 9.7^{ab}	93.9 ± 11.1^{bc}	123.9 ± 15.7^{a}	96.1 ± 15.6^{bc}	74.3 ± 9.4^{cd}	60.7 ± 7.7^{d}
N-Coumaroyl-L-aspartate *trans*	123.4 ± 23.2^{bc}	109.6 ± 22.1^{c}	$314.4 \pm 54.8^{a***}$	–	$188.6 \pm 32.9^{b***}$	$154.0 \pm 26.8^{bc***}$
N-Coumaroyl-L-tyrosine	–	–	–	–	–	–
N-Caffeoyl-L-aspartate	–	–	–	–	–	–
N-Caffeoyl-L-DOPA *cis*	$62.4 \pm 8.3^{ab***}$	$65.9 \pm 5.8^{a***}$	$47.1 \pm 8.5^{b**}$	78.1 ± 12.6^{a}	$7.2 \pm 1.3^{c**}$	$17.4 \pm 3.1^{c**}$
Flavan-3-ols						
Catechin	$35.0 \pm 6.1^{d***}$	$86.8 \pm 7.9^{b**}$	$62.3 \pm 9.2^{c**}$	142.9 ± 21.9^{a}	$20.5 \pm 3.0^{d**}$	$33.9 \pm 5.0^{d**}$
Epicatechin	113.9 ± 19.9^{bc}	153.8 ± 19.5^{ab}	$169.3 \pm 27.1^{a**}$	–	$55.7 \pm 8.9^{d**}$	$92.2 \pm 14.8^{cd**}$
Flavonols						
Quercetin 3-*O*-glucoside	$32.1 \pm 3.2^{a***}$	$32.3 \pm 5.0^{a***}$	–	–	–	–
Quercetin 3-*O*-pentoside	$23.3 \pm 2.2^{a***}$	$18.7 \pm 0.9^{b***}$	–	–	–	–
Methylxanthines						
Theobromine	$44.9 \pm 0.8^{d***}$	72.7 ± 8.5^{c}	$89.4 \pm 3.3^{ab***}$	$97.1 \pm 1.5^{a***}$	$80.8 \pm 3.0^{bc***}$	$79.5 \pm 2.9^{c***}$
Caffeine	$33.0 \pm 0.1^{e***}$	$59.6 \pm 0.8^{d*}$	$77.0 \pm 0.8^{ab**}$	78.5 ± 2.0^{a}	$75.3 \pm 0.7^{b**}$	$68.6 \pm 0.7^{c**}$

Cocoa shell extract						
Hydroxybenzoic acid derivatives						
Gallic acid	86.7 ± 4.9^{a}	42.1 ± 4.6^{b}	53.0 ± 5.1^{b}	90.2 ± 11.3^{a}	28.6 ± 2.8^{c}	22.9 ± 2.2^{c}
Protocatechuic acid	115.2 ± 10.1^{a}	79.3 ± 7.5^{bc}	95.8 ± 13.3^{ab}	89.0 ± 9.4^{b}	83.7 ± 11.6^{b}	60.5 ± 8.4^{c}
N-Phenylpropenoyl-L-amino acids						
N-Coumaroyl-L-aspartate *cis*	95.3 ± 9.6^{a}	102.3 ± 19.8^{a}	–	–	–	–
N-Coumaroyl-L-aspartate *trans*	109.0 ± 16.5^{ab}	91.6 ± 6.8^{b}	37.0 ± 3.0^{c}	125.0 ± 15.2^{a}	22.2 ± 1.8^{c}	18.1 ± 1.5^{c}
N-Coumaroyl-L-tyrosine	127.7 ± 27.9^{ab}	163.4 ± 51.9^{a}	82.2 ± 25.9^{bc}	–	45.8 ± 14.4^{c}	37.4 ± 11.8^{c}
N-Caffeoyl-L-aspartate	110.4 ± 14.0^{a}	90.4 ± 19.0^{a}	–	–	–	–
N-Caffeoyl-L-DOPA *cis*	100.0 ± 13.6^{b}	123.0 ± 16.9^{a}	80.8 ± 10.9^{bc}	64.2 ± 2.2^{c}	12.3 ± 1.7^{d}	29.9 ± 4.0^{d}
N-Caffeoyl-L-DOPA *trans*	101.4 ± 9.6^{a}	96.2 ± 5.9^{a}	–	–	–	–
Flavan-3-ols						
Catechin	90.3 ± 11.1^{c}	121.7 ± 16.0^{b}	87.6 ± 3.2^{c}	155.0 ± 22.9^{a}	28.8 ± 1.0^{d}	47.7 ± 1.7^{d}
Epicatechin	87.8 ± 13.5^{c}	173.4 ± 17.8^{b}	95.5 ± 11.6^{c}	397.6 ± 50.0^{a}	31.4 ± 3.8^{d}	52.0 ± 6.3^{cd}
Flavonols						
Quercetin 3-*O*-glucoside	88.9 ± 4.2^{a}	85.7 ± 3.6^{a}	–	–	–	–
Quercetin 3-*O*-pentoside	96.6 ± 13.5^{a}	100.0 ± 7.5^{a}	–	–	–	–
Flavones						
Vicenin-2	96.9 ± 6.2^{b}	117.1 ± 2.3^{a}	67.6± 8.0^{c}	–	8.1 ± 1.0^{d}	11.8 ± 1.4^{d}
Methylxanthines						
Theobromine	86.5 ± 5.0^{a}	67.5 ± 8.1^{b}	48.0 ± 4.2^{c}	73.7 ± 6.9^{b}	43.4 ± 3.8^{c}	42.7 ± 3.7^{c}
Caffeine	82.7 ± 7.0^{a}	74.1 ± 8.9^{ab}	58.7 ± 8.2^{bc}	73.0 ± 7.1^{ab}	57.5 ± 8.1^{bc}	52.3 ± 7.3^{c}

Results are reported as mean ± SD (n = 3). Mean values within a line followed by different superscript letters (a, b, c, d, e) are significantly different when subjected to Tukey's test ($p < 0.05$). OP: Oral Phase; GP: Gastric Phase; IP: Intestinal Phase; CP: Colonic Phase; C2A: Caco-2 Absorption; HIA: Human Intestinal Absorption;

nd: non-detected; t: traces. Mean values followed by superscript asterisks significantly differ (CSF *vs.* CSE) when subjected to *T*-test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

Fig. 1.

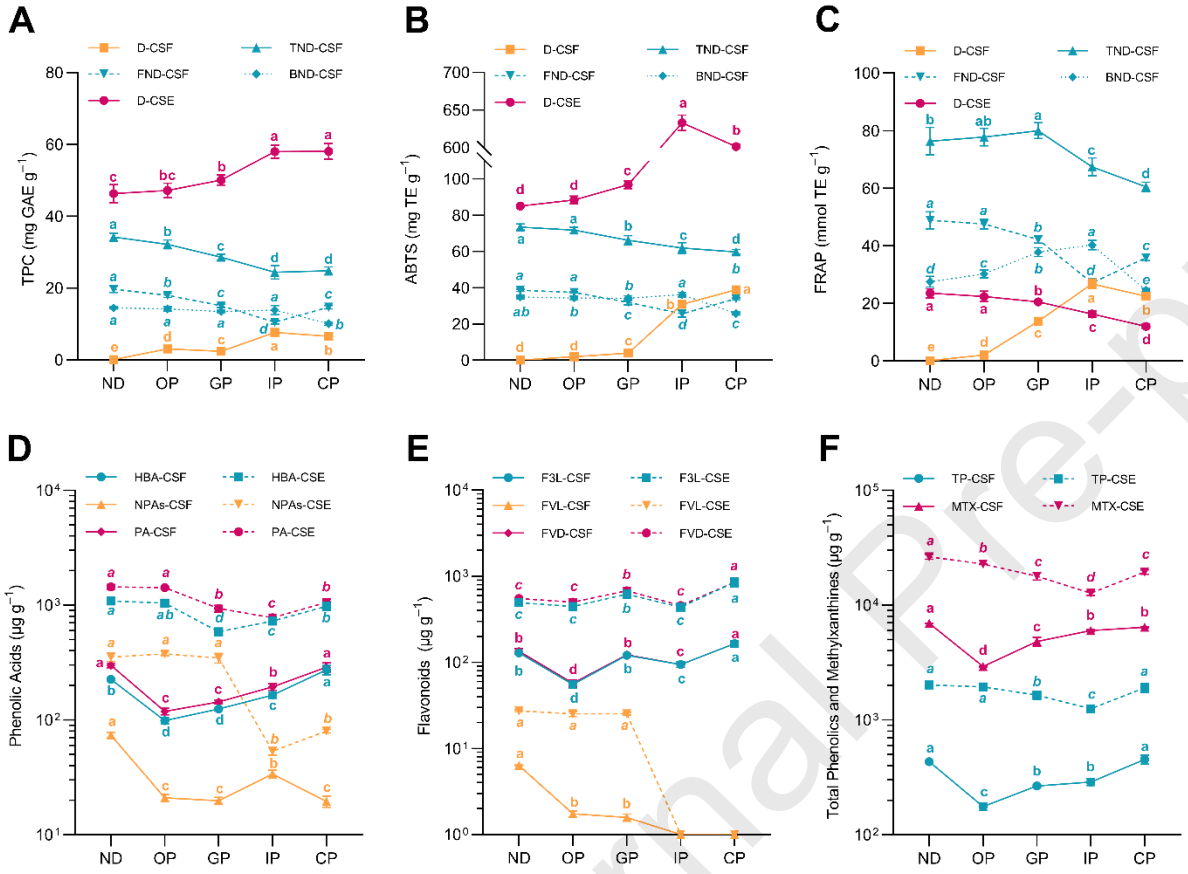


Fig. 2.

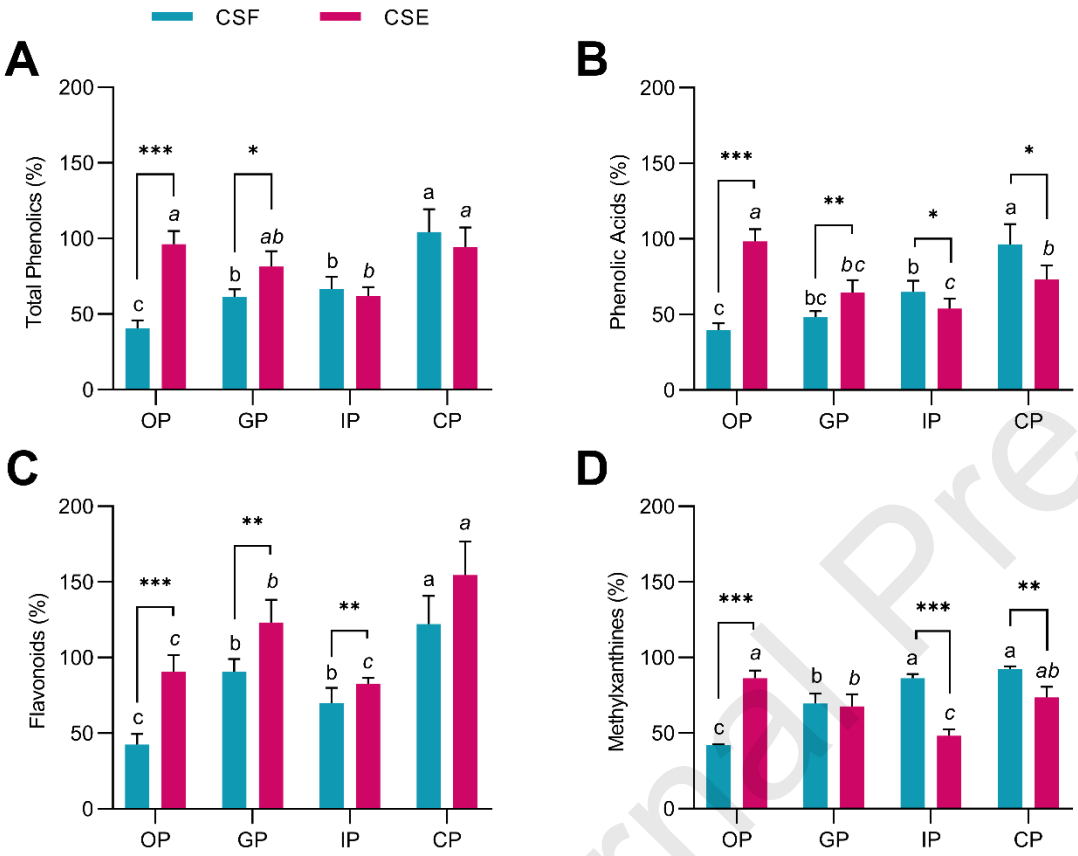


Fig. 3.

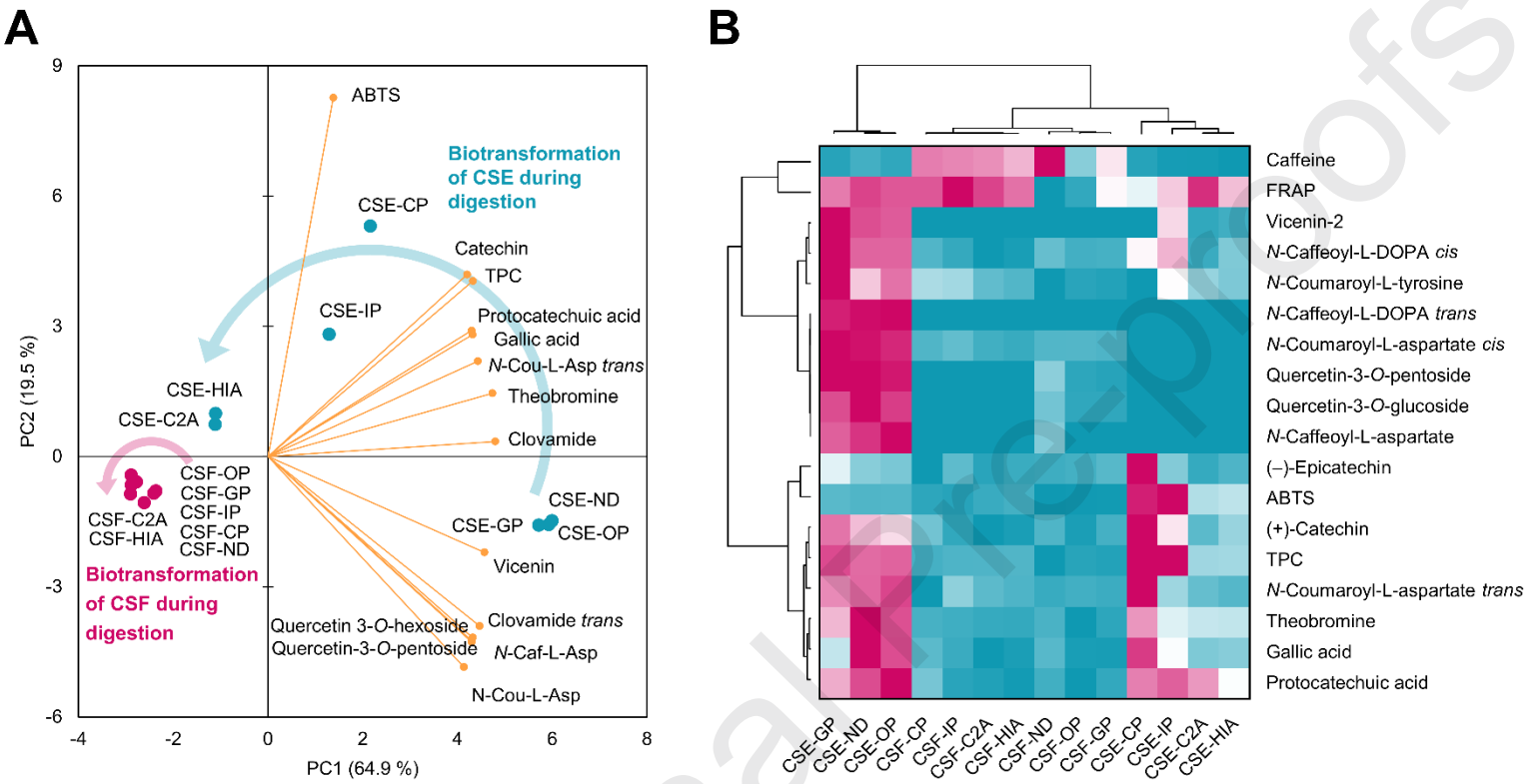


Fig. 4.

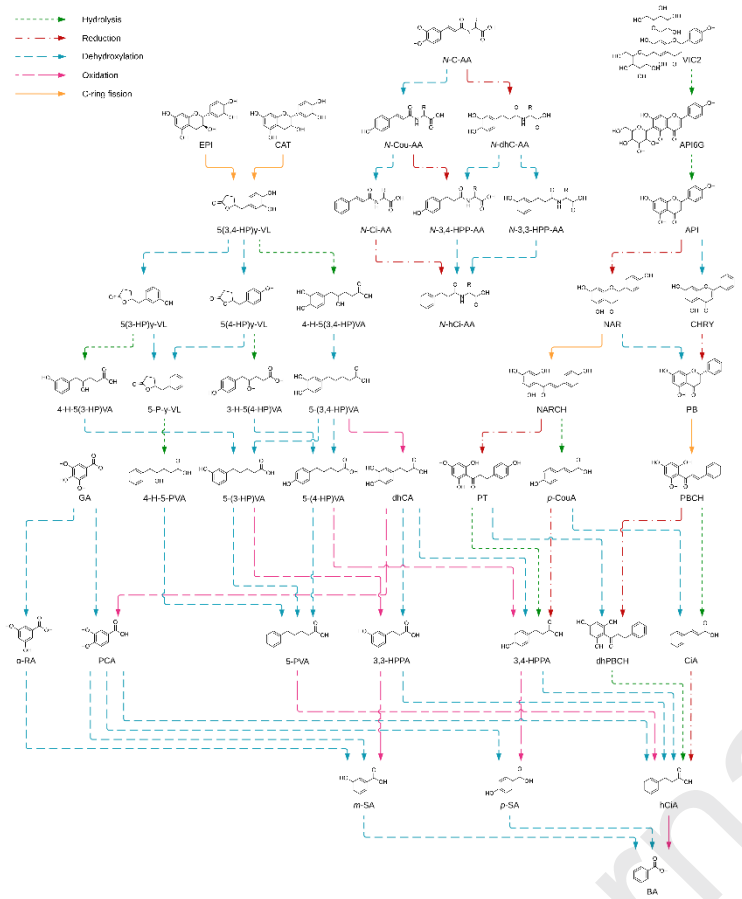
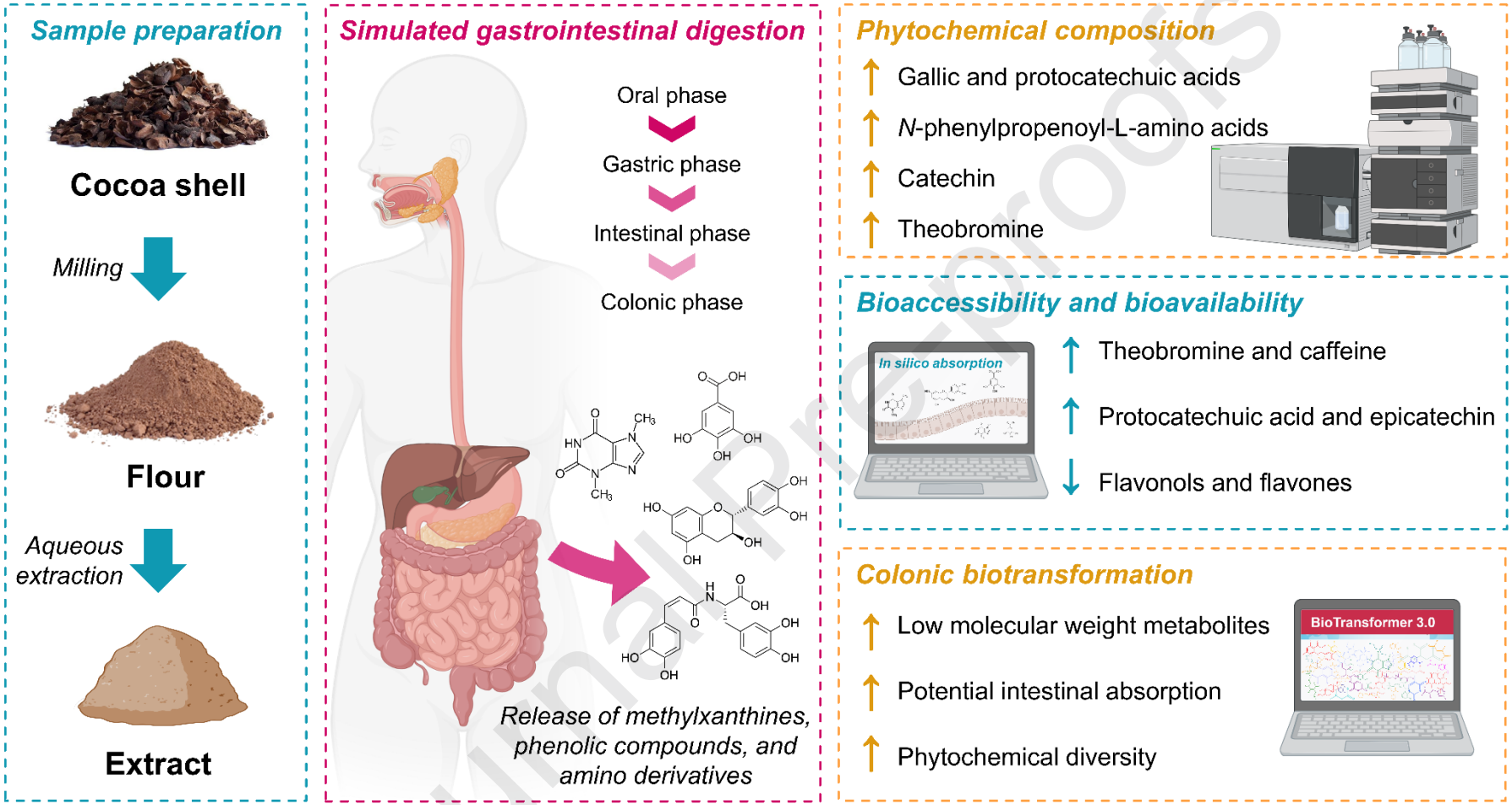
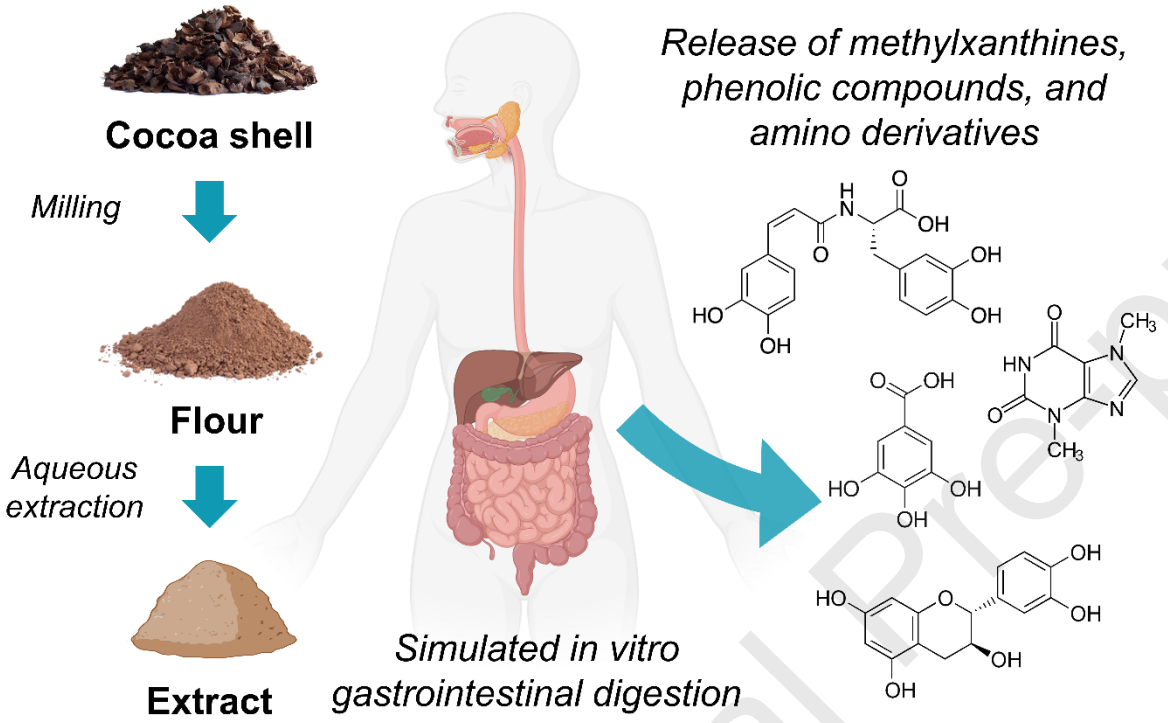


Fig. 5.



Graphical abstract



Highlights

- The CSF and the CSE cocoa shell phytochemical profiles were as evaluated during gastrointestinal digestion
- Phenolic compounds and methylxanthines were released throughout the digestion
- Phenolic compounds' bioaccessibility was higher in the CSE than in the CSF depended on the cocoa shell matrix
- Phenolic compounds were poorly absorbed in the gut in comparison to methylxanthines
- Colonic biotransformation could generate smaller and more adsorbable phenolic metabolites

Credit and authorship contribution statement

Silvia Cañas: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization; **Miguel Rebollo-Hernanz:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – review & editing, Visualization; **Cheyenne Braojos:** Investigation; **Vanesa Benítez:** Conceptualization; **Rebeca Ferreras-Charro:** Formal analysis, Investigation; **Montserrat Dueñas:** Formal analysis; **Yolanda Aguilera:** Writing – review & editing; **María A. Martín-Cabrejas:** Conceptualization, Writing – original draft, Writing – review & editing, Supervision, Project administration.

Conflict of Interest

The authors declare that there is no financial/personal interest or belief that could affect their objectivity.