JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY
ARTICLE

J. Agric. Food Chem. 2010, 58, 12105–12111 12105
DOI:10.1021/jf102985w

^{1}H NMR Study of Fermented Cocoa (*Theobroma cacao* L.) Beans

Augusta Caligiani,*,† Domenico Acquotti,‡ Martina Cirlini,† and Gerardo Palla†

†Dipartimento di Chimica Organica e Industriale, Università di Parma, Via Usberti 17A, 43124 Parma, Italy, and ‡Centro Interdipartimentale Misure, Università di Parma, Via Usberti 23A, 43124 Parma, Italy

This study reports for the first time the metabolic profile of cocoa (*Theobroma cacao* L.) beans using the ^{1}H NMR technique applied to polar extracts of fermented cocoa beans. The simultaneous detection and quantification of amino acids, polyalcohols, organic acids, sugars, methylxanthines, catechins, and phenols were obtained by assigning the major signals of the spectra for different varieties of cocoa beans (Forastero, Criollo, and Trinitario) from different countries (Ecuador, Ghana, Grenada, and Trinidad). The data set obtained, representative of all classes of soluble compounds of cocoa, was useful to characterize the fermented cocoa beans as a function of the variety and geographic origin.

KEYWORDS: ^{1}H NMR; cocoa beans; variety; geographic origin; metabolic analysis

INTRODUCTION

Cocoa trees (*Theobroma cacao* L.) grow in a limited geographical zone approximately 10° to the north and south of the equator (in particular, Central America, West Indian islands, South America, Africa). Three important varieties of cocoa exist: Forastero, Criollo, and Trinitario. Forastero comprises 95% of the world production of cocoa, and it is the most widely used. Overall, the highest quality of cocoa comes from the native Forastero variety of Ecuador (known as Arriba) and from the Criollo variety. Fresh cocoa seeds undergo fermentation and a drying process in the countries of origin to be ready for transport to the countries in which they will be processed to chocolate and related products. Local or regional variations in cocoa plant materials, fermentation procedures, and drying processes lead to a traded good typical of the country of origin. Therefore, the composition of the fermented cocoa beans, which is one of the most important factors influencing the taste and flavor of the cocoa products, depends both on the cocoa variety and on the geographical origin (*1*). Because commercial samples usually lack information on fermentation and drying practices as well as on planting material used, a rapid method for the quality control of traded cocoa beans would be desirable. Cocoa beans and derived products were extensively studied for their composition by several analytical methods, aimed at the detection of amino acids (*2*), sugars (*3*), polyphenols (*4*), methylxanthines (*5*), and aroma compounds (*6*), but to our knowledge, no complete assignment of the ^{1}H NMR spectra of the soluble components of fermented cocoa beans has been performed to date. The ^{1}H NMR technique has been successfully applied to determine the metabolite profile of other food matrices such as tomatoes (*7*), lettuce leaves (*8*), tea (*9*), and coffee (*10*).

^{1}H NMR was previously utilized for the quantification of organic acids and some other metabolites in cupuassu (*Theobroma grandiflorum* Spreng), a *Theobroma* species with composition and utilization similar to those of cocoa, but no full assignment of the NMR spectrum was performed (*11*). A NMR method aimed at the classification of cocoa beans according to the geographic origin was also developed, by applying a multivariate statistical analysis of heterocorrelated bidimensional NMR experiments (HMBC) of polyphenolic extracts, and patented, but this does not report the assignment of the signals (*12*).

In this context, the aim of this work has been the full assignment of the ^{1}H NMR spectra of hydroalcoholic extracts of a series of traded cocoa beans to provide enough qualitative and quantitative information on cocoa bean composition to discriminate between varieties (Forastero, Criollo, and Trinitario) and geographical origin (Ecuador, Ghana, Grenada, and Trinidad).

MATERIALS AND METHODS

Materials. Ten lots of fermented cocoa beans of different varieties and different geographical origins were considered, kindly provided by an Italian chocolate manufacturer (Streglio s.p.a., Turin, Italy). These were three lots of Forastero variety from Ecuador (Arriba), three lots of Forastero from Ghana, two lots of Trinitario variety from Trinidad, and two lots of Criollo variety from Grenada. All of the beans analyzed were classified as "well fermented cocoa beans" on the basis of the brown color.

Standards of organic acids, sugars, amino acids, alcohols, (−)-epicatechin, (+)-catechin, caffeine, theobromine, caffeic acid, D_2O, CD_3OD, and 3-(trimethylsilyl)propionate-d_4 (TSP, internal standard for NMR analysis) were purchased from Sigma-Aldrich (Milan, Italy). Standards of procyanidins B1 and B2 were from Extrasynthese (Lyon, France) and kindly provided by Dr. Daniele Del Rio (Dipartimento di Sanità Pubblica, Sezione di Igiene, Università di Parma, Italy).

Sample Preparation for NMR Analyses. Two hundred milligrams of fermented cocoa beans, finely ground, was extracted with 20 mL of a distilled water/methanol mixture (8:2 v/v), kept at the boiling point for 10 min under magnetic stirring. Extracts were cooled, filtered, taken to dryness, dissolved in 1 mL of D_2O/CD_3OD (8:2 v/v) containing 0.1% of TSP, filtered again, and transferred to a 5 mm NMR sample tube. TSP was used for chemical shift referencing (δ = 0 ppm) and as internal standard for the quantitative analysis. Extractions were performed in triplicate for each lot of fermented cocoa beans.

*Corresponding author (phone +39 0521 905407; fax +39 0521 905472; e-mail augusta.caligiani@unipr.it).

Published on Web 11/03/2010

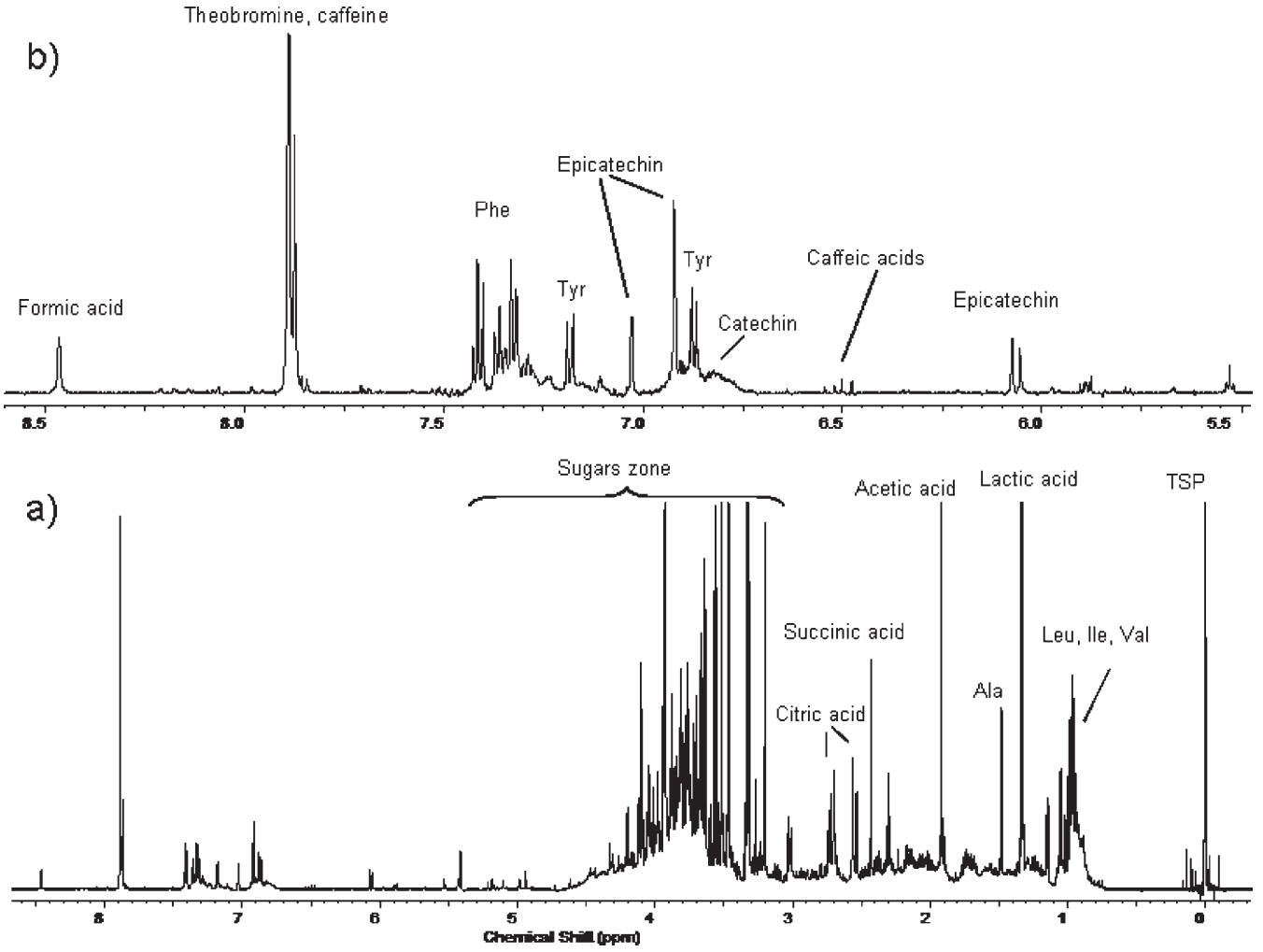

Figure 1. ^{1}H NMR spectrum (600 MHz, solvent D_2O/CD_3OD 8:2 v/v) of a Trinitario cocoa bean hydroalcoholic extract: (**a**) full spectrum; (**b**) expansion of the 5.5−8.5 ppm zone.

^{1}H NMR Acquisition. ^{1}H NMR spectra were acquired on a Varian-Inova 600 MHz spectrometer, equipped with a triple-resonance inverse probe (HCN), operating at 599.736 MHz for proton. The experiments were carried out with water suppression by low-power selective water signal presaturation of 1.5 s. Spectra were acquired at 308 K, with 32K complex points, using a 45° pulse length. One hundred and twenty-eight scans were acquired with a spectral width of 9611.9 Hz, an acquisition time of 1.3 s, and a relaxation delay (d1) of 3 s. The NMR spectra were processed by MestreC software. The spectra were Fourier transformed with FT size of 64K and a 0.2 Hz line-broadening factor, phased and baseline corrected, and referenced to the TSP peak (0 ppm).

Bidimensional NMR Experiments. Gradient-selected COSY spectra (*13*) (gCOSY from the Varian standard pulse sequence library) were acquired at 308 K, with 2048 data points covering a spectral width of 6596.9 Hz in both dimensions. Thirty-two scans were acquired for each of 256 increments, with an acquisition time of 0.155 s and a water presaturation during the relaxation delay of 1.5 s. Spectra were processed with a sinebell function in both dimensions.

TOCSY spectra (*14*) were acquired with water suppression with the same experimental parameters reported for gCOSY. Duration of the spinlock was 80 ms.

Gradient-selected HSQC spectra (*15*) (gHSQC from the Varian standard pulse sequence library) were obtained with water suppression during 1.5 s of relaxation delay, at 308 K, with spectral widths of 6596.9 Hz in F2 and 25632.8 Hz in F1. The other parameters were kept the same as those of the homocorrelated experiments. The gradient selected heteronuclear multiple bond correlation spectra (gHMBC from the Varian standard pulse sequence library) (*16*) were obtained in the same conditions, with spectral widths of 6596.9 Hz in F2 and 30911.9 Hz in F1.

RESULTS AND DISCUSSION

Signal Identification. Figure 1 shows the ^{1}H NMR spectrum of a hydroalcoholic extract of a fermented cocoa bean. The main signals of the spectra were assigned by comparison with spectra of pure compounds or with literature chemical shift data. Some minor signals or overlapped resonances were assigned by using connectivity information from bidimensional experiments (gCOSY, TOCSY, gHSQC, gHMBC). The resulting assignments are summarized in **Table 1**. The chemical structures of relevant aromatic metabolites in cocoa are reported in **Figure 2**.

Methylxanthines. The dominating signals of all the ^{1}H NMR spectra recorded are those of theobromine (3,7-dimethylxanthine), representing with caffeine (1,3,7-trimethylxanthine) the most abundant alkaloid of cocoa seeds. The ^{1}H NMR spectra of methylxanthines are very simple as all of the signals are singlets. The signals of caffeine and theobromine are very sharp and can be easily assigned because of their abundance and their chemical shifts falling in spectrum zones free of interferences. In particular, the hydrogen of the imidazole ring is very deshielded, giving signals at 7.862 ppm for caffeine and at 7.880 ppm for theobromine. The singlets at 3.462 and 3.508 ppm were used for the quantification of caffeine and theobromine, respectively (**Table 2**). The data obtained are in agreement with the contents of caffeine and theobromine reported in literature for cocoa beans (*17*, *18*); therefore, ^{1}H NMR can be considered a simple and rapid alternative method for the quantification of methylxanthines, which

Table 1. Summary of the Substances Identified in 600 MHz ^{1}H NMR Spectra of Fermented Cocoa Bean Hydroalcoholic Extracts[a]

δ	multiplicity	compound	group	J (Hz)	observable in 1D ^{1}H NMR[b]
0.945	t	isoleucine	CδH$_3$	7.42	y
0.970[c]	bt	leucine	CδH$_3$		y
0.999	d	valine	CγH$_3$	7.01	y
1.017[c]	d	isoleucine	CγH$_3$	7.05	y
1.050[c]	d	valine	Cγ'H$_3$	7.05	y
1.143[c]	d	2,3-butanediol	C1H$_3$ + C4H$_3$	5.16	y
1.332[c]	d	lactic acid	C3H$_3$	6.91	y
1.335		threonine	CγH$_3$		n
1.370	d	acetoin	C4H$_3$	7.30	y
1.484[c]	d	alanine	CβH$_3$	7.24	y
1.526	d	?		7.26	y
1.7	m	arginine + leucine			y/n
1.904	m	GABA	CβH2		y
1.927	s	acetic acid	C2H$_3$		y
2.05	m	glutamic acid			y/n
2.14	m	glutamic acid			y/n
2.303[c]	t	GABA	CαH$_2$	7.29	y
2.350		glutamic acid			y/n
2.450[c]	s	succinic acid	C2H$_2$ + C3H$_2$		y
2.552[c]	d	citric acid	C2Ha + C4Ha	16.74	y
2.632	dd	malic acid			n
2.670	dd	aspartic acid			y/n
2.714	d	citric acid	C2Hb + C4Hb	16.81	y
2.745	dd	(−)-epicatechin	C4Ha		y/n
2.748	dd	malic acid			n
2.814	dd	aspartic acid			n
2.850	dd	asparagine	CβH$_2$		y/n
2.900	dd	(−)-epicatechin	C4Hb		y/n
2.953	dd	asparagine			y/n
3.020	t	GABA	CγH$_2$	7.30	y
3.106	dd	tyrosine?		8.14; 14.59	y
3.199	s	choline ?	N(CH$_3$)$_3$		y
3.331	s	caffeine	N3CH$_3$		y
3.460		sucrose (glu)	C4H		y/n
3.462[c]	s	theobromine	N3CH$_3$		y
3.508[c]	s	caffeine	N1CH$_3$		y
3.545		sucrose (glu)	C2H		y/n
3.551	d	β-D-fructopyranose	C1Hb	11.81	y/n
3.559	dd	glycerol	C1Ha + C3Ha	6.53; 11.67	y/n
3.646	dd	glycerol	C1Hb + C3Hb	4.47; 11.69	y/n
3.670	dd	β-D-fructofuranose	C6H		y/n
3.671	s	sucrose	C1H$_2$ (fructose)		y/n
3.696	dd	β-D-fructopyranose	C6Hax	1.96; 12.70	y/n
3.704	d	β-D-fructopyranose	C1Ha	11.68	y/n
3.755		sucrose (glu)	C3H		y/n
3.770	m	glycerol + sucrose (fru)	C2H		y/n
3.805	m	β-D-fructofuranose			y/n
3.806	s	ferulic acid?	OCH3		y/n
3.817	m	sucrose	C5H (fructose)		y/n
3.841	m	sucrose (glu)	C5H + C6H2		y/n
3.886	dd	β-D-fructopyranose	C4H		y/n
3.919	s	theobromine	N7CH$_3$		y
3.930	s	caffeine	N7CH$_3$		y
4.026[c]	dd	β-D-fructopyranose	C6Heq		y
4.046		sucrose	C4H (fructose)		y
4.096[c]	m	β-D-fructofuranose	C3H, C4H		y
4.193	d	sucrose	C3H (fructose)	8.67	y
4.209	d	stachiose + raffinose		8.74	y/n
4.320	m	(−)-epicatechin	C3H		y/n
4.615[c]	d	β-D-glucopyranose	C1H	7.90	y
4.934	s	(−)-epicatechin	C2H		y
4.980	t	stachyose + raffinose		3.53	y
5.101	d	?		2.85	y
5.209[c]	d	α-D-glucopyranose	C1H	3.75	y
5.412[c]	d	sucrose	C1H (glucose)	3.84	y
5.432[c]	d	raffinose + stachyose	C1H (glucose)	3.57	y
5.524	t	?		4.85	y
5.887	dd	?		3.81; 11.74	y

Table 1. Continued

δ	multiplicity	compound	group	J (Hz)	observable in 1D ^{1}H NMR[b]
6.053	d	(−)-epicatechin	C8H	1.96	y
6.064	dd	?		1.49; 12.74	y
6.075	d	(−)-epicatechin	C6H	1.98	y
6.488[c]	d	caffeic acid derivative	C2H	15.75	y
6.518	s	fumaric acid?			y
6.535[c]	d	caffeic acid derivative	C2H	15.89	y
6.546[c]	d	caffeic acid derivative	C2H	15.88	y
6.869[c]	d	tyrosine	C3H, C5H	8.40	y
6.919	m	(+)-catechin			y
6.920	m	(−)-epicatechin	C5′H + C6′H		y
7.027[c]	s	(−)-epicatechin	C2′H		y
7.181	d	tyrosine	C2H, C6H	8.43	y
7.322	dd	phenylalanine	C2H, C6H	7.20; 1.46	y
7.356[c]	t	phenylalanine	C4H	7.53	y
7.405	td	phenylalanine	C3H, C5H	7.51; 1.74	y
7.862	s	caffeine	C6H		y
7.880	s	theobromine	C6H		y
8.460[c]	s	formic acid	C1H		y
8.830	dd	?		7.76; 19.3,7	y

[a] Solvent D_2O/CD_3OD 8:2 v/v. Chemical shifts are reported with respect to TSP (δ = 0.000 ppm). [b] y, observable; n, completely overlapped by stronger signals; y/n, partially overlapped. [c] Specific signals utilized for integration and quantitative analysis.

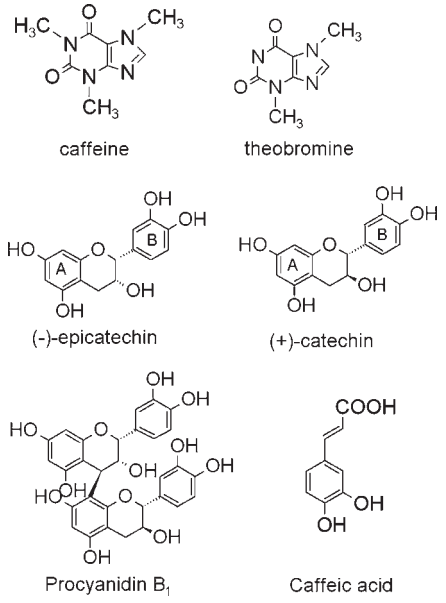

Figure 2. Structures of some aromatic compounds characteristic of cocoa bean.

are normally determined by various high-performance liquid chromatography protocols (*5*, *19*).

Organic Acids. The main organic acids identified by ^{1}H NMR in the cocoa bean extracts are lactic, acetic, succinic, citric, malic, and formic acids. Aromatic acids, such as caffeic acid, were also detected and are considered under Polyphenols. Acetic and citric acids are the most abundant in cocoa. The attribution of the doublet at 1.332 ppm to the methyl group of lactic acid was confirmed by the TOCSY correlation with the methyne group at 4.100 ppm. The signal of lactic acid partially overlaps with that of threonine at 1.335 ppm, evidenced by the TOCSY signals at 3.568 and 4.248 ppm. Acetic, succinic, and formic acids are characterized by unique singlets at 1.927, 2.450, and 8.460 ppm, respectively. Citric acid has two well recognizable signals at 2.552 and 2.714 ppm; malic acid, instead, was very difficult to detect because it is less abundant in cocoa and its signals in the high-field region overlap with those of aspartate and catechins. Spectra interpretation in this region is further complicated by the fact that several methylene groups are prochiral, with the two nonequivalent hydrogens increasing the multiplicity of the signals (generally double doublets). The monodimensional protonic spectra do not permit a certain assignment of the crowded resonance in the 2.6−3 ppm zone, whereas the spread in the second dimension of the bidimensional TOCSY experiment permitted the full assignment of aspartate, asparagine, malic acid, and epicatechin to be obtained (**Figure 3**). Malic acid was detected by the TOCSY correlation of the signals centered at 2.632 and 2.748 ppm with the signal at 4.560 ppm. All of the malic acid signals are partially overlapped and hinder the quantification, whereas the other organic acids considered show at least one well-resolved signal that allows quantification (**Table 2**).

Amino Acids. Aliphatic amino acids were detected in the high-field region of the spectra (0.5−2.5 ppm), where the signals of the lateral chains are present. The signals of the hydrogen bound to the α carbon are more difficult to assign, because they fall in a spectrum zone crowded with the abundant sugar signals. Non-overlapped signals useful for quantification were detected for isoleucine, leucine, valine, alanine, γ-aminobutyric acid (GABA) and, in the aromatic zone, for tyrosine and phenylalanine. The arginine multiplet is centered at 1.7 ppm and overlaps with the methylene of leucine. Partially overlapped signals of aspartic acid (2.670 and 2.814 ppm) and asparagine (2.850 and 2.953 ppm) were assigned on the basis of the TOCSY correlations with the α hydrogens centered at 3.877 and 3.976 ppm, respectively (**Figure 3**). Glutamic acid and glutamine, as aspartic acid and asparagine, give weak overlapped signals in the 2.05−2.15 ppm zone, which can be separated in the second dimension of the TOCSY experiment. The threonine methyl group partially overlaps the strong lactic acid signal.

Sugars and Alcohols. Among carbohydrates, the predominant sugars are fructose and sucrose. The sugar zone (3.0−4.5 ppm) is very crowded; their resonances are often overlapped, especially those of fructose with their corresponding di- and oligosaccharides (e.g., sucrose, raffinose, stachiose), so the assignments were performed with the aid of bidimensional experiments (TOCSY and HSQC). Some signals of the two main forms of fructose (β-D-fructopyranose, β-D-fructofuranose) and of sucrose were assigned by their TOCSY correlations. Sucrose is well recognizable in the monodimensional ^{1}H NMR experiments for

Table 2. Quantitative Results (Grams per Kilogram) Obtained by ^{1}H NMR for Some Selected Substances Characteristic of Cocoa Beans and Observable in ^{1}H NMR Spectra[a]

	Forastero 'Arriba' from Ecuador (3 lots)	Forastero from Ghana (3 lots)	Trinitario from Trinidad (2 lots)	Criollo from Grenada (2 lots)
leucine	0.56 ± 0.15	1.05 ± 0.44	1.03 ± 0.10	1.15 ± 0.06
valine	0.29 ± 0.05	0.45 ± 0.15	0.47 ± 0.09	0.48 ± 0.05
isoleucine	0.22 ± 0.04	0.33 ± 0.14	0.35 ± 0.10	0.36 ± 0.03
2,3-butanediol	0.23 ± 0.11	0.16 ± 0.06	0.19 ± 0.05	0.36 ± 0.01
lactic acid	1.28 ± 0.61	1.20 ± 0.31	1.78 ± 0.20	2.28 ± 0.15
alanine	0.29 ± 0.06	0.49 ± 0.18	0.49 ± 0.05	0.58 ± 0.05
acetic acid	1.47 ± 0.58	0.80 ± 0.63	1.02 ± 0.17	1.33 ± 0.18
γ-aminobutyric acid	0.71 ± 0.17	0.83 ± 0.33	0.92 ± 0.08	0.77 ± 0.01
succinic acid	0.32 ± 0.14	0.24 ± 0.09	0.28 ± 0.10	0.29 ± 0.04
citric acid	3.39 ± 0.95	3.43 ± 1.01	2.92 ± 0.52	2.66 ± 0.35
theobromine	6.95 ± 0.29	6.06 ± 1.46	6.58 ± 0.93	4.73 ± 0.60
caffeine	1.64 ± 0.21	0.81 ± 0.26	1.87 ± 0.44	1.68 ± 0.31
fructose	1.37 ± 0.05	1.85 ± 0.61	1.60 ± 0.12	1.63 ± 0.02
glucose	0.58 ± 0.01	0.40 ± 0.03	0.35 ± 0.07	0.30 ± 0.03
sucrose	4.16 ± 0.09	0.98 ± 0.30	1.24 ± 0.21	0.14 ± 0.08
oligosaccharides[b]	1.54 ± 0.13	0.84 ± 0.15	0.88 ± 0.08	0.40 ± 0.08
caffeic acid derivatives[c]	0.31 ± 0.17	0.16 ± 0.13	0.18 ± 0.02	0.20 ± 0.01
tyrosine	0.64 ± 0.32	0.54 ± 0.30	0.62 ± 0.11	0.66 ± 0.03
epicatechin	3.24 ± 1.18	1.77 ± 1.29	1.94 ± 1.01	1.81 ± 0.41
phenylalanine	0.48 ± 0.17	0.89 ± 0.38	0.83 ± 0.06	0.89 ± 0.02
formic acid	0.08 ± 0.03	0.08 ± 0.07	0.12 ± 0.01	0.12 ± 0.02

[a] Results are the means of all lots with similar characteristics. [b] Expressed as raffinose. [c] Expressed as caffeic acid.

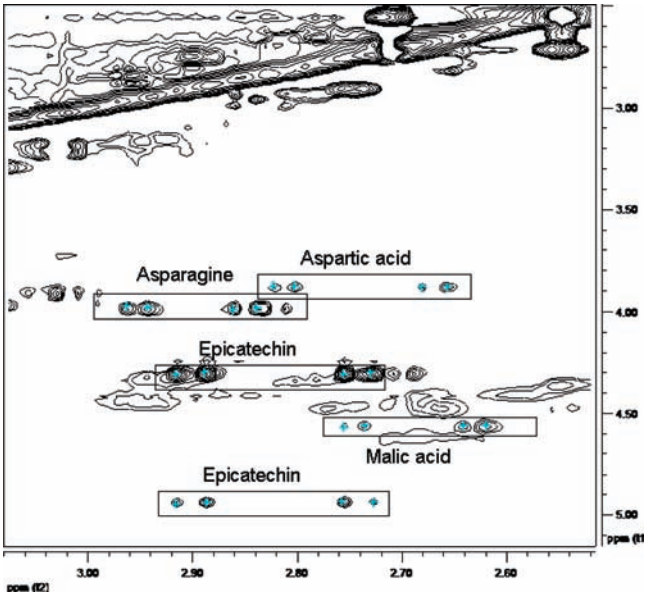

Figure 3. Bidimensional TOCSY of cocoa bean hydroalcoholic extract: expansion of the 2.5−3.5 ppm spectral zone. Solvent was D_2O/CD_3OD 8:2 v/v.

the anomeric glucose signal at 5.412 ppm. The anomeric signals permitted also the identification of the α and β forms of glucose, whereas the other glucose signals overlap with those of other abundant sugars. Raffinose and stachyose can be recognized by their glucose anomeric signal at 5.432 ppm, but they are not separately determined as their resonances overlap. In some samples it was also possible to detect the glycerol resonances and those of two alcohols characteristic of the fermentation process, 2,3-butanediol (1.143 ppm) and acetoin (1.370 ppm). 2,3-Butanediol, in particular, is quite abundant and easily quantifiable.

Polyphenols. One of the most studied class of substances in cocoa is the polyphenols, which give ^{1}H NMR signals in the aromatic zone, centered at 7 ppm. It is well-known that cocoa is a rich source of polyphenols, in particular, flavan-3-ols such as epicatechin and catechin (**Figure 2**), representing > 30% of total polyphenols, and procyanidins (about 60%), which are oligomeric compounds composed by catechin and epicatechin, primarily joined via 4−8 linkage (*20*, *21*). The structure of procyanidin B1 is reported in **Figure 2**.

The identification of polyphenols in cocoa bean ^{1}H NMR spectra has required the registration of the ^{1}H NMR spectra of catechin, epicatechin, and procyanidin B1 and B2 standards to facilitate the assignments of the corresponding signals in the spectra of cocoa bean extracts. Registering under the same condition the ^{1}H NMR spectra of the polyphenol standards, only very small differences in chemical shifts were observed between monomers and dimers; therefore, it is possible to hypothesize that the signals of the different dimeric procyanidins are superimposed with those of the monomeric epicatechin and catechin. In particular, in the ^{1}H NMR spectra of cocoa bean extract (**Figure 1b**) two important signals at 6.920 and 7.027 ppm, corresponding to hydrogens of the B ring of epicatechin, are well detectable; the homonuclear bidimensional experiments confirmed that these signals are correlated each other and can be attributed also to some procyanidins containing the epicatechin moiety. The signals of the A ring fall near 6 ppm, but their intensity in the spectra of the standard epicatechin and catechin are lower than expected. This was explained considering a proton exchange with the deuterium oxide for a tautomeric equilibrium between phenolic and chinonic forms. This was confirmed by the correct intensities obtained in aprotic deuterated solvent as DMSO. Therefore, the signals of ring A, even if well detectable, cannot be utilized for quantitative analysis. The signals of the heterocyclic ring are not easily detected because they fall in the high- and medium-field regions of the spectra, crowded with other resonances, as previously discussed. The assignment of the signals centered at 2.745 and 2.900 ppm to epicatechin was done on the basis of their TOCSY correlation with the signals centered at 4.320 and 4.934 ppm. The epicatechin content in fermented cocoa beans is largely higher with respect to catechin (about 20−25 times), and for this reason the signals of catechin monomer in the ^{1}H NMR spectra of cocoa beans were not observable. However, (+)-catechin can be recognized in cocoa as a component of procyanidin, for example, in procyanidins B1 and B3, and the amount of these substances is relevant in cocoa (*22*). The signals of procyanidins

containing catechin monomer (for example, procyanidin B1) appear as a broad multiplet centered at 6.919 ppm.

By integrating the signals of epicatechin at 6.920 and 7.027 ppm, an average content of 2.5 g/kg was obtained, which is in agreement with literature data. However, higher values were expected, as these signals should contain also procyanidins; hence, we realized that the extraction conditions utilized (water/methanol 8:2) were not effective for the exhaustive extraction of procyanidins, which are normally recovered under acidic media. Therefore, further extraction conditions were applied: 100% methanol, 100% water, and a mixture of acetone/water/acetic acid, as reported in the literature (*4*). All of the extractions were carried out at the boiling point, for 10 min. The most effective mixtures for the extraction of catechins were those of water/methanol (8:2) and acetone/water/acetic acid (70:28:2), but no significant differences in catechins were observed between the two different extractions (measured as peak areas at 7.027 ppm). The signals' shape in the zone 6.5−7.5 ppm was considered. Excluding the extraction effect (**Figure 1b**), the baseline is rising in this zone, possibly for the presence of several minor phenol-type substances that contribute, together with catechins, to the total polyphenol content reported in the cocoa literature. The integration of the entire area and the expression as gallic acid equivalents give results comparable with those obtained by measuring the total phenol content with the traditional Folin method. Moreover, procyanidins previously studied by ^{1}H NMR (*23*), at room temperature, show a broadening of the ^{1}H NMR signals by atropisomerism, caused by steric interactions among free rotating flavonoids in the vicinity of the interflavonoid bonds. For this reason, the shifts of catechin or epicatechin due to the polymerization could be masked in our experiments by the broadening of the signals. To obtain a better resolution of the procyanidin signals, a series of experiments was performed at lower temperatures (from 308 to 253 K, physical lower limit of the instrument probe), but without significant improvement of the spectra, which suggested that even lower temperatures were probably required.

Another class of phenolic compounds present in cocoa is that of gallic, caffeic, and ferulic acids (*24*). Gallic acid gives a ^{1}H NMR spectrum with a broad singlet at about 7 ppm, which probably contributes to the high baseline recorded in this zone, as previously reported. The methoxy group of ferulic acid can be tentatively assigned to a singlet at 3.806 ppm. The group of signals centered at 6.5 ppm can be attributed to derivatives of caffeic acid. The ^{1}H NMR spectrum of pure caffeic acid shows the doublet of the olefinic hydrogen near the carboxylic group at 6.197 ppm, with a coupling constant of 15.86 Hz, characteristic of the hydrogen in a double bond. The spectra of cocoa extracts do not show this specific signal, but three doublets around 6.5 ppm (6.488, 6.535, and 6.546 ppm) with signal shapes and coupling constants (15.75, 15.88, and 15.89 Hz, respectively) very similar to those of caffeic acid. The significant shift to lower field with respect to caffeic acid can be explained by the presence of an electron attractor group, probably an ester or an amide. Some derivatives of caffeic acids are reported in cocoa, for example, clovamide (*N*-caffeoyl-L-3,4-dihydroxyphenylalanine) and a deoxy analogue of clovamide (*25*, *26*). The ^{1}H NMR assignments reported for olefinic hydrogens in deoxyclovamide (7.37 and 6.48 ppm) are compatible with the signal at 6.488 ppm, which shows a TOCSY correlation with a signal at 7.37 ppm, overlapped in the monodimensional spectra by the more intense phenylalanine signals.

Recently, other derivatives of caffeic acid and amino acids (*N*-phenylpropenoyl amino acids) were determined in cocoa, for example, *N*-(*E*)-caffeoyl-L-aspartate (caffeic acid aspartate) (*27*), and the two signals at 6.535 and 6.546 ppm can be tentatively attributed to these derivatives, in agreement with the ^{1}H NMR assignments previously obtained (*28*). However, due to the uncertainty of the assignments, these three substances in **Table 1** are reported as generic caffeic acid derivatives and integrated together for the quantitative analysis.

Unidentified Signals. Some signals between 5.5 and 7.7 ppm remain unassigned. For example, the complex signal at 5.887 ppm correlates in the TOCSY experiment with a signal falling at 7.864 ppm, assignable to a purine alkaloid, covered by the strong signal of caffeine. Another unresolved signal occurring at 7.090 ppm shows a TOCSY correlation with a 6.801 ppm resonance and could be attributed to another procyanidin. A further signal at 7.445 ppm, partially overlapped with phenylalanine, correlates with a 7.66 ppm signal. Other minor unidentified signals are present in the lower field zone (8−9.5 ppm).

Quantification of Selected Substances. According to the assignments obtained, the integration of the nonoverlapped signals was performed as follows. Integrals were normalized to the area of TSP, added to each sample in an exactly known amount, and the values were converted in mass/mass value (mg/g) as previously reported (*29*). The accuracy of the quantitative data was assured by the relaxation delay, determined by T1 measurements, which was set to allow the complete relaxation of the nuclei. T1 varied from 0.7 to 1.5 s for high-field protons and low-field protons, respectively, and theobromine shows the longest relaxation time. The data that were possible to obtain by a single ^{1}H NMR experiments are reported in **Table 2** as a mean of all the values obtained for lots of fermented cocoa beans with analogous characteristics in terms of variety and geographic origin. The data are representative of all the most important classes of cocoa bean compounds and can be considered a useful data set for the quality evaluation of traded cocoa bean lots. A more detailed evaluation of the preliminary data obtained revealed significant differences arising for the different cocoa samples. In particular, Arriba subspecies of Forastero from Ecuador are the samples with the most peculiar characteristics, and they were very different from the other Forastero variety (Ghana), confirming the importance of the fermentation and drying processes for the composition of traded beans. Arriba beans are characterized by a higher content of phenolic substances (epicatechin and caffeic acid) and sugars, sucrose and oligosaccharides in particular, and a lower content of amino acids. On the contrary, the other "fine" cocoa, Criollo, showed a very low amount of carbohydrates and a higher content of amino acids. Theobromine and caffeine showed high variability also among samples of the same variety and geographic origin. The same variability was observed for phenolic compounds.

In conclusion, analysis by ^{1}H NMR appears promising for the simultaneous detection of the main classes of cocoa bean components and can be proposed as a rapid method for cocoa bean quality control and for the quantification of some important substances such as methylxanthines and catechins. Further work is in progress on a more significant number of cocoa bean samples and also on roasted derived products to develop a NMR-based method for the classification of cocoa beans and chocolate. Examination of a significant number of samples for each variety of cocoa beans would permit also more considerations of the biochemical reactions involved in the metabolism of each specific variety of cocoa plant.

LITERATURE CITED

(1) Roelofsen, P. A. Fermentation, drying, and storage of cacao beans. In *Advances in Food Research*; Mrak, E. M., Stewart, G. F., Eds.; Academic Press: New York, 1958; Vol. 8, pp 225−296.

(2) Adeyeye, E. I.; Akinyeye, R. O.; Ogunlade, I.; Olaofe, O.; Boluwade, J. O. Effect of farm and industrial processing on the amino acid profile of cocoa beans. *Food Chem.* **2010**, *118*, 357−363.

(3) Redgwell, R. J.; Trovato, V.; Curti, D. Cocoa bean carbohydrates: roasting-induced changes and polymer interactions. *Food Chem.* **2003**, *80*, 511−516.
(4) Cooper, K. A.; Campos-Gimenez, E.; Jimenez Alvarez, D.; Nagy, K.; Donovan, J. L.; Wlliamson, G. Rapid reversed phase ultra-performance liquid chromatography analysis of the major cocoa polyphenols and inter-relationships of their concentrations in chocolate. *J. Agric. Food Chem.* **2007**, *55*, 2841−2847.
(5) Pura Naik, J. Improved high-performance liquid chromatography method to determine theobromine and caffeine in cocoa and cocoa products. *J. Agric. Food Chem.* **2001**, *49*, 3579−3583.
(6) Frauendorfer, F.; Schieberle, P. Changes in key aroma compounds of criollo cocoa beans during roasting. *J. Agric. Food Chem.* **2008**, *56*, 10244−10251.
(7) Sobolev, A. P.; Segre, A. L.; Lamanna, R. Proton high-field NMR study of tomato juice. *Magn. Reson. Chem.* **2003**, *41*, 237−245.
(8) Sobolev, A. P.; Brosio, E.; Gianferri, R.; Segre, A. L. Metabolic profile of lettuce leaves by high-field NMR spectra. *Magn. Reson. Chem.* **2005**, *43*, 625−638.
(9) Tarachiwin, L.; Ute, K.; Kobayashi, A.; Fukusaki, A. ^{1}H NMR based metabolic profiling in the evaluation of Japanese green tea quality. *J. Agric. Food Chem.* **2007**, *55*, 9330−9336.
(10) Del Campo, G.; Berregi, I.; Caracena, R.; Zuriarrain, J. Quantitative determination of caffeine, formic acid, trigonelline and 5-(hydroxymethyl)furfural in soluble coffees by ^{1}H NMR spectrometry. *Talanta* **2010**, *81*, 367−371.
(11) Figueiredo, I. M.; Pereira, N. R.; Efraim, P.; Garcia, N. H. P.; Rodrigues, N. R.; Marsaioli, A., Jr.; Marsaioli, A. J. ^{1}H NMR, a rapid method to monitor organic acids during cupuassu (*Theobroma grandiflorum* Spreng) processing. *J. Agric. Food Chem.* **2006**, *54*, 4102−4106.
(12) Soulet, S.; Vercauteren, J.; Troplin, P.; Raybaud, G. NMR determination of the geographic origin of cocoa beans and their products. Fr. Demande FR 2857450 A1 20050114, 2005.
(13) Hurd, R. E. Gradient-enhanced spectroscopy. *J. Magn. Reson.* **1990**, *87*, 422−428.
(14) Levitt, M. H.; Freeman, R.; Frenkiel, T. Broadband heteronuclear decoupling. *J. Magn. Reson.* **1982**, *47*, 328−330.
(15) Wilker, W.; Leibfritz, D.; Kerssebaum, R.; Bermel, W. Gradient selection in inverse heteronuclear correlation spectroscopy. *Magn. Reson. Chem.* **1993**, *31*, 287−292.
(16) Rinaldi, P. L.; Keifer, P. A. The utility of pulsed-field gradient HMBC for organic structure determination. *J. Magn. Reson. Ser. A* **1994**, *108*, 259−262.
(17) Risner, C. H. Simultaneous determination of theobromine, (+)-catechin, caffeine, and (−)-epicatechin in Standard Reference Material Baking Chocolate 2384, cocoa, cocoa beans, and cocoa butter. *J. Chromatogr. Sci.* **2008**, *46*, 892−899.
(18) Del Rosario Brunetto, M.; Gutierrez, L.; Delgado, Y.; Gallignani, M.; Zambrano, A.; Gomez, A.; Ramos, G.; Romero, C. Determination of theobromine, theophylline and caffeine in cocoa samples by a high-performance liquid chromatographic method with on-line sample cleanup in a switching-column system. *Food Chem.* **2007**, *100*, 459−467.
(19) Bispo, M. S.; Veloso, M. C. C.; Pinheiro, H. L. C.; DE Oliveira, R. F. S.; Reis, J. O. N.; De Andrade, J. B. Simultaneous determination of caffeine, theobromine, and theophylline by high-performance liquid chromatography. *J. Chromatogr. Sci.* **2002**, *40*, 45−48.
(20) Sanchez-Rabaneda, F.; Jauregui, O.; Casals, I.; Andres-Lacueva, C.; Izquierdo-Pulido, M.; Lamuela-Raventos, R. M. Liquid chromatographic/electrospray ionization tandem mass spectrometric study of the phenolic composition of cocoa *Theobroma cacao*. *J. Mass Spectrom.* **2003**, *38*, 35−42.
(21) Wollgast, J.; Anklam, E. Review on polyphenols in *Theobroma cacao*: changes in composition during the manufacture of chocolate and methodology for identification and quantification. *Food Res. Int.* **2000**, *33*, 423−447.
(22) Gu, L.; House, S. E.; Wu, X.; Ou, B.; Prior, R. L. Procyanidin and catechin contents and antioxidant capacity of cocoa and chocolate products. *J. Agric. Food Chem.* **2006**, *54* (11), 4057−4061.
(23) Abe, Y.; Shoji, T.; Kawahara, N.; Kamakura, H.; Kanda, T.; Goda, Y.; Ozeki, Y. Structural characterization of a procyanidin tetramer and pentamer from the apple by low-temperature NMR analysis. *Tetrahedron Lett.* **2008**, *49*, 6413−6418.
(24) Belščak, A.; Komes, D.; Horzic, D.; Kovacevic, K.; Karlovic, D. Comparative study of commercially available cocoa products in terms of their bioactive composition. *Food Res. Int.* **2009**, *42*, 707−716.
(25) Arlorio, M.; Locatelli, M.; Travaglia, F.; Coisson, J. D.; Del Grosso, E.; Minassi, A.; Appendino, G.; Martelli, A. Roasting impact on the contents of clovamide (*N*-caffeoyl-L-DOPA) and the antioxidant activity of cocoa beans (*Theobroma cacao* L.). *Food Chem.* **2008**, *106*, 967−975.
(26) Sanbongi, C.; Osakabe, N.; Natsume, M.; Takizawa, T.; Gomi, S.; Osawa, T. Antioxidative polyphenols isolated from *Theobroma cacao*. *J. Agric. Food Chem.* **1998**, *46* (2), 454−457.
(27) Elwers, S.; Zambrano, A.; Rohsius, C.; Lieberei, R. Differences between the content of phenolic compounds in Criollo, Forastero and Trinitario cocoa seed (*Theobroma cacao* L.). *Eur. Food Res. Technol.* **2009**, *229*, 937−948.
(28) Stark, T.; Hofmann, T. Isolation, structure determination, synthesis, and sensory activity of *N*-phenylpropenoyl-L-amino acids from cocoa (*Theobroma cacao*). *J. Agric. Food Chem.* **2005**, *53*, 5419−5428.
(29) Caligiani, A.; Acquotti, D.; Palla, G.; Bocchi, V. Identification and quantification of the main organic components of vinegars by high resolution ^{1}H NMR spectroscopy. *Anal. Chim. Acta* **2007**, *585*, 110−119.

Received for review July 30, 2010. Revised manuscript received October 14, 2010. Accepted October 16, 2010.