J Sci Food Agric 84:693–700 (online: 2004)
DOI: 10.1002/jsfa.1717

Comparative study on the proteolytic activities and storage globulins in seeds of *Theobroma grandiflorum* (Willd ex Spreng) Schum and *Theobroma bicolor* Humb Bonpl, in relation to their potential to generate chocolate-like aroma

Christoph Reisdorff,[1]* Christina Rohsius,[1] Aparecida das Graças Claret de Souza,[2] Luadir Gasparotto[2] and Reinhard Lieberei[1]

[1]Institute of Applied Botany, University of Hamburg, Ohnhorststraße 18, 22609 Hamburg, Germany
[2]Empresa Brasileira de Pesquisa Agropecuária — Embrapa Amazônia Ocidental, Caixa postal 319, 69.011–970 — Manaus, AM, Brazil

Abstract: The cocoa relatives *T grandiflorum* (cupuaçu) and *T bicolor* (macambo) are promising crop plants for sustainable agroforestry in the Amazon region of South America. The market for cupuaçu is expanding since the fruit flesh is utilised by the foodstuffs industry. Attempts to commercialise chocolate-like wares from the seeds have failed so far because of unreliable product quality. It is not known whether this is due to an insufficient aroma potential of cupuaçu seeds. We therefore investigated the proteolytic enzymes and the seed storage globulins which are both decisive for the formation of aroma precursors in cocoa. We found that the activities of the aspartic endopeptidase and the carboxypeptidase in *T bicolor* and *T grandiflorum* differed slightly from those in cocoa. The specificity of the carboxypeptidase for hydrophobic amino acids was quite similar across the three species, while the optimal pH of the *T grandiflorum* enzyme was lower than that of the other species. The qualitative and quantitative differences between the globulins indicate a lower maximum yield of aroma precursors in *T grandiflorum* and a higher maximum yield of aroma precursors in *T bicolor*, compared to cocoa. We conclude that the quality of chocolate-like products made from the studied cocoa relatives can be improved by adapting fermentation procedures to particular biochemical features of these seeds.

Keywords: aspartic endopeptidase; chocolate aroma; cocoa; cupuaçu; fermentation; globulins; macambo; seed storage proteins; serine-type carboxypeptidase; *Theobroma bicolor*; *Theobroma cacao*; *Theobroma grandiflorum*

INTRODUCTION

The close phylogenetic relationship between the cupuaçu tree (*Theobroma grandiflorum* Willd ex Spreng, Schum), the macambo-tree (*Theobroma bicolor* Humb Bonpl) and the cocoa tree (*Theobroma cacao* L) is reflected in their growth habit, in their floral characteristics and in the morphology of their fruits and seeds.[1,2] The close relationship of these species has been confirmed in terms of the biochemical properties of their seeds, eg the purine alkaloids, the composition of lipids and the vicilin class storage proteins.[3–7]

In tropical South America the cultivation of *T grandiflorum* trees is currently expanding for two main reasons: the plant's high suitability for sustainable agroforestry systems and the high market potential of the aromatic fruit flesh.[8] The seeds of *T grandiflorum* are usually discarded, despite their high fat and protein content which is comparable to that of cocoa seeds.[6] This results in a loss of valuable dietary resources accounting for 350–450 g kg^{-1} of the dry weight of fruit. Some efforts have been made to produce types of chocolate wares from *T grandiflorum* seeds such as the so-called cupulate.[9,10] However, attempts to commercialise those products have not yet been successful, mostly owing to problems in obtaining a satisfactory and uniform aroma quality

* Correspondence to: Christoph Reisdorff, Institute of Applied Botany, University of Hamburg, Ohnhorststraße 18, 22609 Hamburg, Germany
E-mail: christoph.reisdorff@iangbot.uni-hamburg.de
Contract/grant sponsor: Federal Ministry of Education and Research; contract/grant number: Förderkennzahl 0339457A, project ENV23/2
Contract/grant sponsor: Council for Scientific and Technological Development
(*Received 29 November 2003; accepted 22 January 2004*)
Published online 5 April 2004

in the processed seeds. It is not known whether these problems are due to a lack of aroma potential of the primary material or whether the processing procedure used for cocoa seeds is not suitable for seeds of *T grandiflorum* or *T bicolor*, thus requiring the processing procedure to be adapted to particular biochemical and morphological features of the different seeds. In our studies we addressed these questions by comparative investigations on those inherent seed factors which are known to be decisive for the generation of aroma precursors in the case of cocoa seeds.

It has been shown in recent years that the typical aroma of cocoa is formed in stages in the course of fermentation, drying and roasting of the seeds. During the anaerobic phase of fermentation, yeasts degrade the sugars of the adhering pulp to ethanol which is oxidised by bacteria to acetic acid in the subsequent aerobic phase. The acetic acid passes through the shell of the cocoa seeds and penetrates the cotyledon tissue, leading to cell death, disintegration of cell compartments and formation of cellular aqueous reaction phases with lowered pH values.[11–14] These alterations trigger a cascade of proteolytic reactions which result in the formation of aroma precursors from the major seed storage proteins which belong to the vicilin-class globulins.[15,16] The reliable formation of aroma precursors during fermentation depends on the simultaneous activity of two proteolytic enzymes, an aspartic endopeptidase (EC 3.4.23) and a serine-type carboxypeptidase (EC 3.4.16).[17] After fermentation—during drying, roasting and conching—the mixture of hydrolysates undergoes Maillard reactions in the presence of reducing sugars. These Maillard products contribute essentially to the characteristic aroma of cocoa products.[18–21]

It has not yet been established to what extent the seeds of *T grandiflorum* (and *T bicolor*) possess the fundamental biochemical prerequisites for the formation of aroma precursors and we therefore studied the proteolytic activities and the protein patterns of unfermented ripe seeds of *T grandiflorum* in comparison with those of *T cacao*. Most of the experiments were carried out in parallel with seeds of *T bicolor*, which is considered to be of less economic relevance but nevertheless is a promising species for agroforestry systems in the humid tropics of south America.[8]

EXPERIMENTAL

Ripe fruits of *T grandiflorum* and *T bicolor* were harvested from primary selected plants grown on an experimental area of the West Amazonian Agroforestry Research Centre EMBRAPA near the city Manaus, Brazil. Cocoa fruits of the forastero type were obtained from local producers on the outskirts of Manaus. The fruits were opened and the seeds were separated from the adhering pulp. After careful removal of the seed shells the cotyledons were shock frozen in liquid nitrogen and subsequently freeze-dried. After removal of phenolic compounds and fat, the activities of the respective seed proteases were tested in crude extracts. The protein patterns of the seeds were analysed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and paying special attention to the seed globulin fraction.

Chemicals

If not otherwise stated, all chemicals used were pa (pro analysis) grade. Acetic acid, acetone, ascorbic acid, bovine serum albumin (fraction V, for biochemistry), citric acid, Coomassie Blue R-250 (for electrophoresis), $CuSO_4 \cdot 2H_2O$, glycerol, glycine, iodoacetic acid, leucine (for biochemistry), mercaptoethanol, methanol, $Na_2B_4O_7 \cdot 10H_2O$, $Na_2HPO_4 \cdot 2H_2O$, light petroleum (bpt 40–60 °C) (purest grade), trichloroacetic acid and Triton X-100 were obtained from Merck, Darmstadt, Germany. From Fluka, Steinheim, Germany we obtained 2,4,6-trinitro benzene sulfonic acid (1 M solution in water), bromophenol blue (Reag Ph Eur), Na_2-EDTA$\cdot 2H_2O$ (BioChemika Ultra), pepstatin A (BioChemika), phenylmethylsulfonyl fluoride (BioChemika) and sodium ascorbate (Biochemika). Sodium dodecyl sulfate and Tris-(hydroxymethyl)aminomethane were obtained from Serva, Heidelberg, Germany, urea (ultrapure) from Biomol, Hamburg, Germany, and *N*-benzyloxycarbonyl-dipeptides from Bachem Biochemica, Heidelberg, Germany.

Removal of phenolic compounds and fat

The fat-free acetone dry powder (acdp) was prepared according to the procedure described by Kirchhoff *et al*.[22] In order to prevent interactions between proteins and phenolic compounds the dried material was homogenised, using a stainless steel ball homogeniser (Retsch, Haan, Germany) at −5 °C, in 700 ml l^{-1} aqueous acetone containing 5 mmol l^{-1} ascorbic acid. The homogenates were cooled to −20 °C in order to ensure the precipitation of proteins. After centrifuging at 15 000 × *g*, the removal of phenolic compounds was continued by extracting the precipitate once with 800 ml l^{-1} and twice with 700 ml l^{-1} ice-cold aqueous acetone in the presence of 5 mmol l^{-1} ascorbic acid. Afterwards, the samples were washed twice with water free, ice-cold acetone and the acetone remaining in the water free precipitates was evaporated under vacuum at 20 °C. The fat of the resultant dry powders was extracted twice by stirring in light petroleum (40–60 °C) at 20 °C. After filtering, the residual solvent was removed under vacuum at 20 °C. The fat-free acdps were stored at −20 °C.

Extraction of proteases

All extraction steps were undertaken at 2–4 °C. The aspartic endopeptidase was extracted according to Hansen *et al*[23] by adding 30 ml of 500 mmol l^{-1} sodium borate buffer (pH 9) to 1 g of acdp. After suspending in a glass tissue grinder (Kimbel/Kontes, Vineland, NJ), the homogenate was centrifuged for

20 min at 14 000 × *g*. Aliquots of the supernatant were utilised for activity tests (see below).

Serine-type carboxypeptidase was extracted according to Bytof *et al*[24] by a phosphate–citrate buffer of pH 6.8 (20 mmol l^{-1} disodium hydrogen phosphate and 10 mmol l^{-1} citric acid) containing 1 g l^{-1} Triton-X-100, 20 mg l^{-1} pepstatin-A and 2.5 mmol l^{-1} iodoacetic acid. Pepstatin-A and iodoacetic acid were added in order to inhibit protein degradation through aspartic endopeptidase and peptidases of the cysteine type. One gram of the acdp was suspended in 50 ml buffer using a tissue grinder (Kimbel/Kontes, Vineland, NJ, USA). The suspension was pre-incubated for 10 min in an ice bath in order to allow the inhibitors to interact with the respective peptidases. The homogenate was centrifuged for 40 min at 12 000 × *g*. The serine-type carboxypeptidase activity was determined in the supernatant.

Protease activity assays

The aspartic endopeptidase was determined as pepstatin A-sensitive activity. 100 µl of the enzyme extract was pre-incubated with 100 µl of 200 mg l^{-1} pepstatin A or water at 35 °C for 10 min. The proteolysis was started by adding 800 µl of substrate solution consisting of 6.25 mg ml^{-1} bovine serum albumin in McIlvaine buffer (phosphate citrate buffer) adjusted to specific pH values between pH 2.8 and pH 8.2. The pH adjustments of the buffer were made by mixing 200 mM disodium hydrogen phosphate solution with 100 mM citric acid solution in varying volumetric proportions. The effective initial pH values of the assays were measured at the start of the proteolysis with a glass electrode (Schott, Hofheim, Germany, pH Meter CG 825). After incubation at 42 °C for 2 h, the reaction was stopped by adding 0.2 ml of ice-cold 250 g l^{-1} trichloroacetic acid (TCA). After centrifuging at 10 000 × *g*, the TCA-soluble products of the proteolysis were quantified (see below).

The seryl-type carboxypeptidase activity was determined using different N-masked dipeptides (*N*-benzyloxycarbonyl dipeptides, referred to as Z-dipeptides below) at a concentration of 2 or 5 mmol l^{-1} as substrates in the presence (1 mmol l^{-1}) or absence of phenylmethylsulfonyl fluoride (PMSF), which is a potent inhibitor of seryl-type peptidases. One hundred microlitre of the enzyme extract was transferred to 860 µl McIlvaine buffer of a chosen pH value with or without PMSF and incubation at 42 °C was started after the addition of substrate. The initial pH value of the assay was recorded. Proteolysis was stopped by adding 0.2 ml ice-cold 250 g l^{-1} TCA.

The proteolytic activities of both the aspartic endopeptidase and the seryl-type carboxypeptidase were determined by quantifying the liberated α-amino groups in the supernatant of the TCA mixture by the method of Shutov *et al*[25] using tri-nitrobenzene sulfonic acid (TNBS). The reaction of TNBS with α-amino groups leading to coloured sulfonamides is reliable only in the pH range from 8.8 to 9.3. Thus, the pH of the reaction mixture was adjusted to pH 9.0 ± 0.5 by adding 600 µl of a 40 g l^{-1} sodium borate buffer (pH between 9.5 and 9.9) to 150 or 200 µl of the TCA supernatant. The tuning of both the pH of the borate buffer and the volume of the TCA supernatant was necessary to compensate for the different proton concentrations in the proteolysis assays. The colour reaction was started by adding 100 µl of 3 g l^{-1} TNBS solution. The pH of the reaction mixture was monitored in a parallel assay with a glass electrode (Schott pH-Meter CG 825). After shaking for 1 h at 30 °C the reaction was stopped by adding 1.6 ml of 1 mol l^{-1} HCl and the absorption of the solutions was measured photometrically at $\lambda = 340$ nm. A calibration curve was constructed using leucine solutions at concentrations from 0.1 to 5.0 mmol l^{-1} treated in parallel with the samples. The protein contents of the acdp and of the respective extracts were measured photometrically at $\lambda = 546$ nm after biuret reaction in 750 mmol l^{-1} NaOH with 20 mmol l^{-1} $CuSO_4$ using bovine serum albumin as standard.

Seed protein pattern

The patterns of the entire seed proteins and the globulin fractions were studied by SDS-PAGE using BIO-RAD (Munich, Germany) Tris-HCl gradient ready gels with a separation range from 10 to 100 kDa (100–200 g l^{-1} acrylamide concentration). The sample buffer contained 8 mol l^{-1} urea, 20 g l^{-1} SDS, 10 mmol l^{-1} Na_2-EDTA, 20 mmol l^{-1} Tris-HCl, 200 mmol l^{-1} mercaptoethanol, 100 ml l^{-1} glycerol and 0.5 g l^{-1} bromophenol blue. A 25 mmol l^{-1} Tris-HCl solution of pH 6.8 containing 1 g l^{-1} SDS and 0.19 mol l^{-1} glycine was used as the electrode buffer. Staining was done using 1 g l^{-1} Coomassie Blue R-250 dissolved in a mixture of water, methanol and acetic acid in the ratio 5/4/1 (v/v/v). The gels were analysed densitometrically using a Pharmacia (Amersham, Bucks, UK) Image Master device.

A solubility fractionation procedure was applied to purify the globulins. The acdp was extracted three times with low-salt buffer of pH 7.5 containing 10 mmol l^{-1} NaCl, 5 mmol l^{-1} sodium ascorbate, 2 mmol l^{-1} Na_2-EDTA and 50 mmol l^{-1} Tris-HCl in order to remove the bulk albumins. Subsequently the precipitate was extracted twice with high-salt buffer of pH 8.0 containing 500 mmol l^{-1} NaCl and 200 mmol l^{-1} Tris-HCl in an ultrasonic bath (Bandelin Electronic, Berlin, Germany). The combined high-salt extracts were dialysed against distilled water in order to precipitate the globulins. After centrifuging at 47 000 × *g*, the globulins were suspended in small volumes of high-salt buffer. Aliquots of the globulin suspensions containing 2.5 µg of protein were subjected to SDS-PAGE together with 5 µg per track of entire seed proteins which had been extracted by dispersing the acdp protein in SDS-PAGE sample buffer and heating to 95 °C.

For the interspecific comparison of the enzyme activities, six replicates of different acdp were

prepared. After analysis of variance the mean values were compared by least significant difference (LSD) test at $p < 0.01$. The pH dependence of the enzymes was measured in four replicates at each pH step deriving from two enzyme extracts from two acdp preparations.

RESULTS

In addition to a high fat content between 520 and 560 mg g^{-1} DW (dry weight), the cotyledons of *T bicolor* and *T grandiflorum* contained considerable amounts of protein (149 ± 25 mg g^{-1} and 101 ± 19 mg g^{-1} DW respectively) which did not differ significantly from the fat and protein contents of *T cacao* seeds (fat 510 mg g^{-1}, protein 121 ± 18 mg g^{-1} DW).

Proteolytic activities

The endoproteolytic activities at pH 3.6 of *T bicolor* and *T grandiflorum* were 177 and 207 nkatal g^{-1} seed protein respectively, which is about half of the activity observed in *T cacao* seeds (Table 1). In all three cases the endoproteolytic activities at pH 3.6 were completely inhibited by pepstatin A. Thus, the respective activities can be attributed to endopeptidases of the aspartic type (EC 3.4.23). When activity was calculated on the basis of fresh weight the difference between *T grandiflorum* and the other two species was increased owing to the higher water content (0.39 g g^{-1}) in this sample (results not shown). This is an exceptionally high value in comparison to the mean of all examined *T grandiflorum* samples (0.33 ± 0.03 g g^{-1}). The water content of the cotyledons of the other two species showed little variation from 0.30 g g^{-1}.

Proteolysis conducted at different pH levels revealed that the aspartic endopeptidase of both *T cacao* and *T grandiflorum* showed very similar patterns of relative activity (Fig 1) except for the most acidic measuring point, where the enzyme of *T grandiflorum* showed 10% higher activity. A more precise pH curve for the aspartic endopeptidase of *T grandiflorum* was determined by measuring the pepstatin-A sensitive proteolysis between pH 3.0 and pH 8.7 at closer pH intervals (Fig 2). To determine the exact position of optimal activity, a distance weighted least-squares regression analysis was performed on 56 measuring points between pH 3.0 and pH 4.5. The resulting curve explained 96.5% of the observed variation and indicated a most probable maximum activity of the aspartic endopeptidase of *T grandiflorum* as lying between pH 3.45 and pH 3.55. The lowest activity was observed at pH 5.7, whilst at higher pH values there was a slight increase in activity which reached 10% of the measured maximum activity (Fig 2).

Table 1. Activities of peptidases of seeds of three *Theobroma* species at optimum pH (means and SD)

Enzyme activity (nkatal)	Species		
	T bicolor	*T cacao*	*T grandiflorum*
Aspartic endopeptidase			
g^{-1} seed protein	177 (29)^{a}	388 (29)^{b}	207 (19)^{a}
g^{-1} dry weight	24.3 (6.4)^{a}	46.7 (7.8)^{b}	20.2 (4.0)^{a}
g^{-1} fresh weight	17.2 (4.5)^{a}	33.4 (5.6)^{b}	12.4 (2.5)^{a}
Carboxypeptidase			
g^{-1} seed protein	37.8 (2.6)^{b}	25.3 (6.0)^{a}	22.5 (1.9)^{a}
g^{-1} dry weight	5.7 (0.4)^{b}	3.3 (1.0)^{a}	2.1 (0.2)^{a}
g^{-1} fresh weight	4.0 (0.3)^{b}	2.4 (0.7)^{a}	1.3 (0.1)^{a}

Aspartic endopeptidase was determined as pepstatin-A sensitive activity on bovine serum albumin at pH 3.6.
Carboxypeptidase was measured with Z-phenylalanine-alanine as substrate at pH 5.6.
Different letters within a row indicate significant differences at $p < 0.01$ (LSD test, $n = 6$).

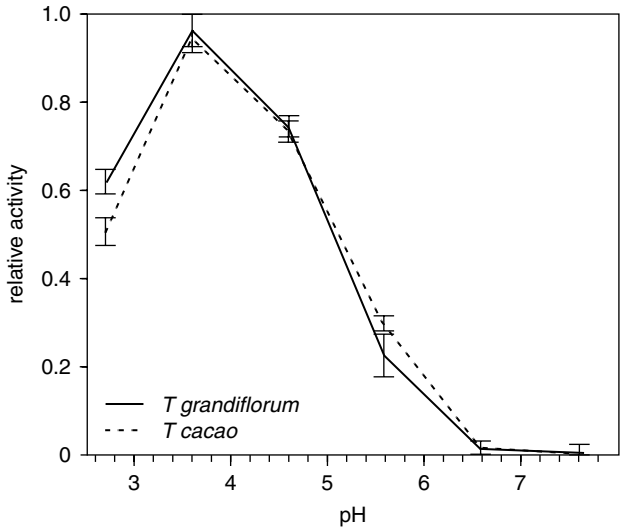

Figure 1. Comparison of the proteolytic activities of aspartic endopeptidase in seeds of *T cacao* and *T grandiflorum* at various pH values. Mean, minimum and maximum activities ($n = 4$) are plotted as proportions of the highest value of the respective species.

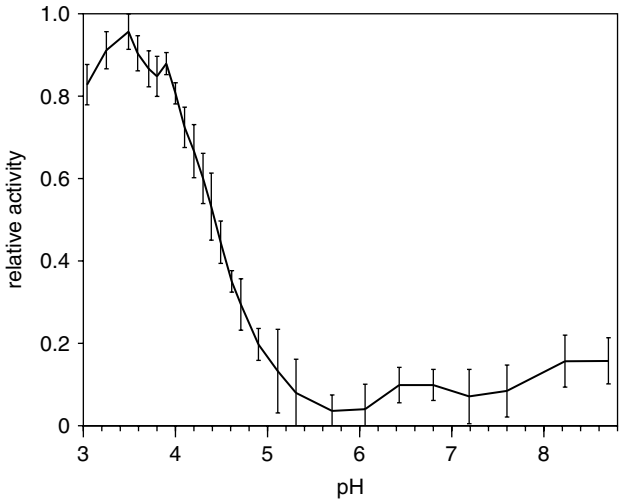

Figure 2. pH dependence of the aspartic endopeptidase activity of seeds of *T grandiflorum* measured as pepstatin-A sensitive proteolysis of bovine serum albumin. Mean values are connected by a line, bars represent SD ($n = 4$).

The extraction of carboxypeptidase from the acdp was most effective when 1 g l^{-1} of Triton-X-100 was added to the extraction buffer, leading to a doubling of activity yield. At higher concentrations of Triton-X-100 (at 5, 10 and 20 g l^{-1}) this positive effect progressively faded away.

The activity of the carboxypeptidase at pH 5.6 with Z-phenylalanine-alanine as substrate was around a tenth of the respective aspartic endopeptidase activity. No significant differences could be determined between *T cacao* and *T grandiflorum*, whereas *T bicolor* showed significantly higher activity (Table 1). The effect of 1 mmol l^{-1} PMSF was the same in all three species producing an inhibition of 70.3 ± 7.35% (results not shown) of the carboxypeptidase activity, indicating that the activity was due to a serine type enzyme (EC 3.4.16) in all three species.

The pH dependence of the carboxypeptidase with Z-phenylalanine-alanine as substrate was quite similar in *T cacao* and *T bicolor* with an optimum between pH 5.5 and 6.0, whilst the enzyme of *T grandiflorum* reached its maximum activity at a value of around pH 5.4 (Fig 3). Further investigations on the *T grandiflorum* carboxypeptidase revealed that the shape and the maximum value of the pH curve depended on the carboxyterminal amino acid residue of the Z-dipeptide substrate: With bulky hydrophobic side chains of the carboxyterminal amino acid residues (Z-phenylalanine-leucine and Z-phenylalanine-phenylalanine), the maximum activity was found at pH 5.7. Small nonpolar amino acid residues at the carboxyterminal site of the Z-dipeptide (Z-phenylalanine-glycine and Z-phenylalanine-alanine) led to a shift of the optimal activity to pH 5.4. In the case of the polar carboxyterminal amino acid residue in Z-phenylalanine-glutamate a flat maximum was observed at pH 4.8 but overall the activities were rather low (Fig 4).

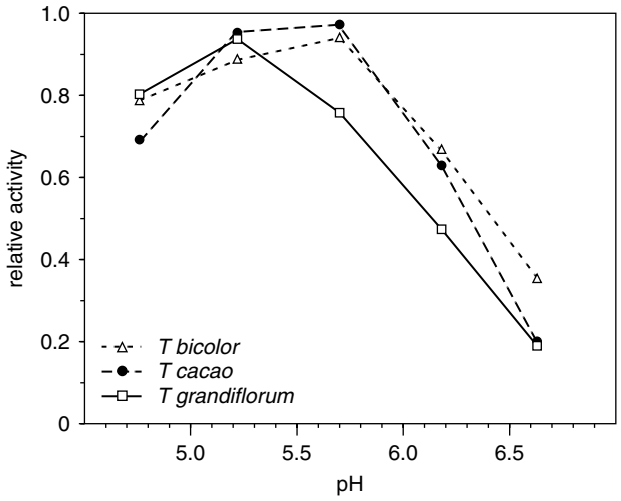

Figure 3. Comparison of the proteolytic activities of the carboxypeptidase in seeds of *T bicolor, T cacao* and *T grandiflorum* at various pHs of the incubation buffer using Z-phenylalanine-alanine as substrate. Mean values are shown as proportions of the maximum value of the respective species.

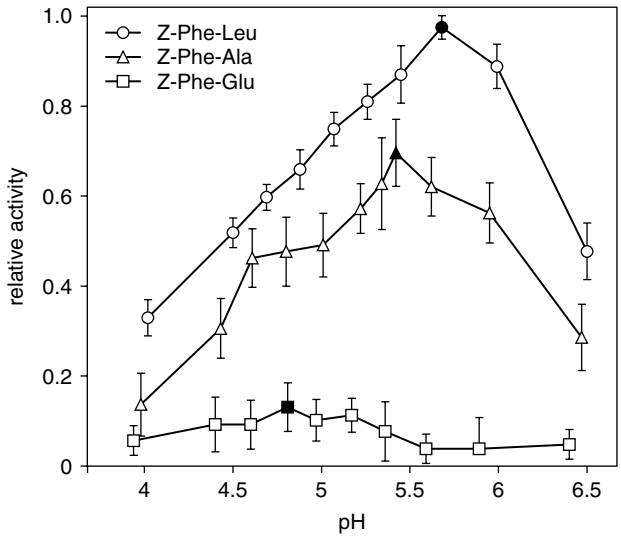

Figure 4. pH dependence of the carboxypeptidase in seeds of *T grandiflorum* as influenced by the carboxyterminal amino acid of the Z-dipeptide substrates (Z-phenylalanine-leucine, Z-phenylalanine-alanine, Z-phenylalanine-glutamate). Mean values and standard deviations ($n = 4$) are shown as proportions of the maximum value measured with Z-phe-leu. The pH maxima of the particular substrates are highlighted by filled symbols.

In all three species the turnover number of the carboxypeptidase at saturated substrate concentrations depended strongly on both amino acid residues of the Z-dipeptides. With reference to the terminal amino acid, hydrophobic residues (phenylalanine, leucine and alanine) were preferentially cleaved by the enzymes of all three species. The carboxypeptidase of *T grandiflorum* enzyme showed a clear preference for phenylalanine among these hydrophobic amino acids, whereas the enzyme of *T bicolor* was the least specific. The respective specificity of the *T cacao* enzyme was intermediate, showing a tendency to cleave bulky hydrophobic amino acids at higher rates. The activities were very low in all three species if the terminal amino acid residue was polar as in Z-phenylalanine-glutamate (Fig 5A).

The activity of the carboxypeptidase depended strongly on the presence of a hydrophobic amino acid residue adjacent to the cleaved one. Alanine adjacent to a terminal phenylalanine led to the highest activity observed. Bulky amino acid residues were less effective in that position. With glycine or glutamate adjacent to the terminal phenylalanine, the activity of the carboxypeptidase dropped towards the limit of detection. The influence of the amino acid residue adjacent to the carboxyterminal was the same in all three species (Fig 5B).

Protein pattern

The SDS-PAGE patterns of seed proteins of all three *Theobroma* species are marked by two predominant polypeptides occurring in the globulin fraction, and at least one predominant albumin at around 20 kDa. Regarding the globulin fraction, a protein in the

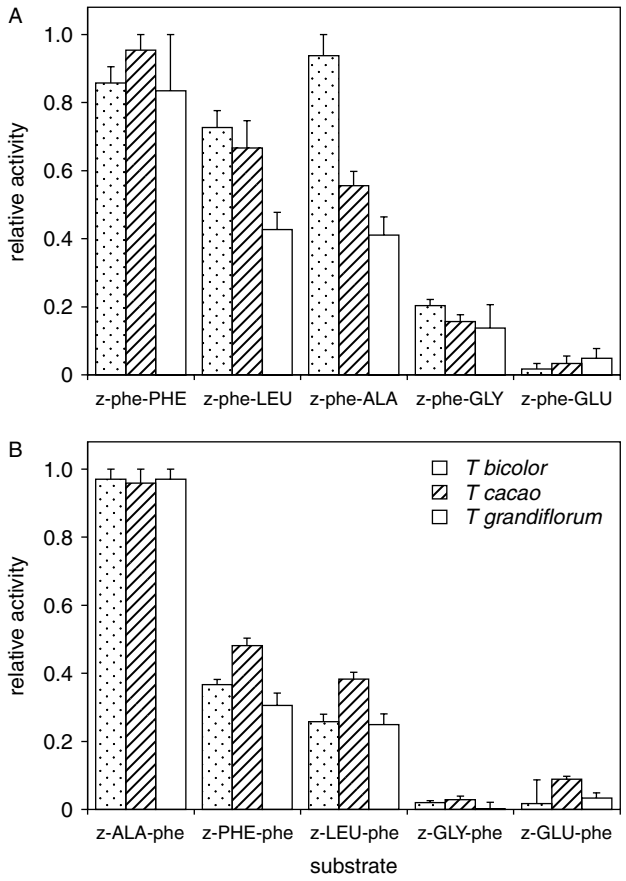

Figure 5. Turnover rates at pH 5.6 of carboxypeptidase of seeds of *T bicolor, T cacao* and *T grandiflorum* using Z-dipeptides as substrates with varying carboxyterminal amino acids attached to a phenylalanine (A) and with varying amino acids between a carboxyterminal phenylalanine and the *N*-benzyloxycarbonyl residue (B). Mean values and standard deviations ($n = 4$) are plotted as proportions of the maximum value measured with Z-phe-phe (A) and Z-phe-leu (B).

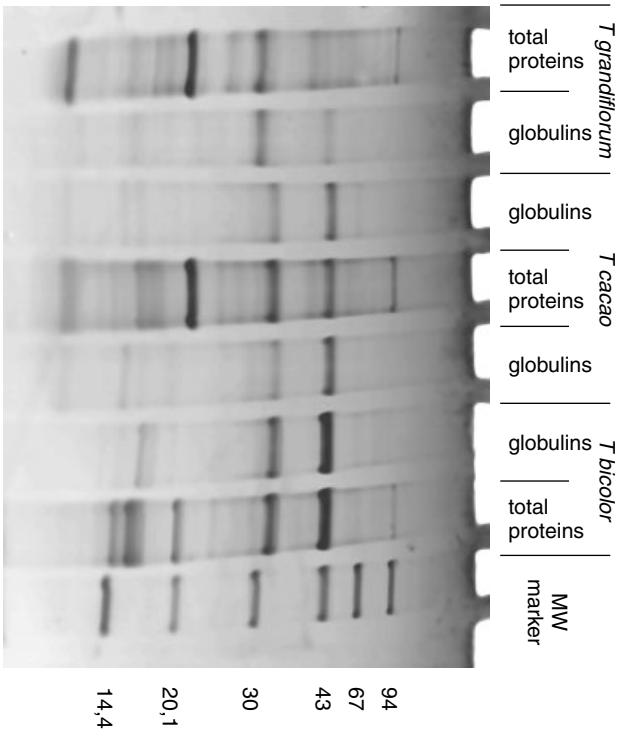

Figure 6. SDS gradient (100–200 g l^{-1}) polyacrylamide gel electrophoresis of seed proteins of *T bicolor, T cacao* and *T grandiflorum*. Globulins were obtained by high-salt (0.5 mol l^{-1} NaCl) and low-salt (0.01 mol l^{-1} NaCl) solubility fractionation procedures (for details refer to 'experimental'); total proteins were obtained by dispersing acdp in a sample buffer containing 70 mmol l^{-1} SDS. The amount of total protein and globulin per track was 5 µg and 2.5 µg respectively. Molecular weights are given in kDa.

15 kDa range was detected but this was present only in very small amounts in all three species (Fig 6).

The separation patterns of the globulins of *T bicolor* were very similar to those of cocoa. The globulin at 31 kDa shows no difference from the corresponding globulin of cocoa. At about 46 kDa, a very evident band appears to correspond to the 47 kDa globulin of cocoa. *T bicolor* was distinguished by its very high content of these two globulins, with each exceeding the quantity of the corresponding cocoa globulin. The albumin corresponding to the predominant albumin of *T cacao* and *T grandiflorum* at 20 kDa was less abundant in *T bicolor* and had a slightly lower molecular weight.

The protein patterns of *T cacao* and *T grandiflorum* were less similar with respect to both the separation properties and the relative quantities of the globulins. In *T grandiflorum* only a slight band could be detected in the range of the 47 kDa globulin, whereas this band was predominant in the globulin fraction of the other two species. The light globulin chain, which differed from the corresponding polypeptide in *T cacao* by 2 kDa, constituted most of the globulin fraction of *T grandiflorum*.

The seeds of *T bicolor* were characterised by their very high relative content of storage globulins (nearly 500 g kg^{-1} of total protein), while it was evident that the seeds of *T grandiflorum* contained more albumins than globulins. However, densitometric determination of the protein bands revealed that in all three species, the quantity of the light globulin chain was around 150 g kg^{-1} of the total seed protein.

DISCUSSION AND CONCLUSIONS

It is known that the reliable generation of the precursors of chocolate-like aroma depends on the co-operative activities of the aspartic endopeptidase and the seryl-type carboxypeptidase, which cleave the vicilin-class globulins in a specific manner.[26] We were able to show that these basic components are also present in the seeds of *T grandiflorum* and *T bicolor*. However, there are some distinct features of these two cocoa relatives which need to be considered in order to arrive at an adequate fermentation procedure.

Due to the exceptionally high amount of aspartic endopeptidase in cocoa seeds[27] the vicilin class globulins disappear within the first two or three days of fermentation.[28] Since the activities of the two relatives were half of that observed in *T cacao*, the maximum yield of aroma precursors in fermenting seeds of *T bicolor* and *T grandiflorum* might occur one or two days later.

Voigt and Biehl[29] found that the cotyledon tissue must be acidified to values between pH 5.0 and pH 5.5 during fermentation in order to guarantee satisfactory formation of aroma precursors. These pH values indicate the range where the pH dependencies of the aspartic endopeptidase and the seryl-type carboxypeptidase activities of cocoa overlap.[26] In comparison to cocoa, the pH range for simultaneous activities of the two enzymes of *T grandiflorum* is shifted to slightly lower pH values, due to the more acidic pH optimum of the carboxypeptidase. In addition, the aspartic endopeptidase activity of *T grandiflorum* was close to its minimum at pH 5.0 whereas, towards a slightly more acidic value of pH 4.5, the activity approached 50% of its maximum. From these *in vitro* observations we conclude that, during the fermentation of seeds of *T grandiflorum,* the pH in the cotyledons needs to be lowered from initial values of about pH 6.0 (data not shown) to a range between pH 5.0 and 4.5 in order to trigger sufficient activity of both enzymes.

However, when Bytof *et al*[24] used globulin derived peptides instead of Z-dipeptides as substrates for cocoa carboxypeptidase, the pH optimum of this enzyme was found at slightly higher pH values. We also observed an evident influence of the substrate on the pH dependence of the *T grandiflorum* carboxypeptidase. Thus, the effective optimum pH for the generation of aroma precursors should be the subject of further studies using globulin derived peptides as substrate.

The proteolysis of cocoa globulins during fermentation produces a mixture of hydrophilic oligopeptides and hydrophobic amino-acids,[26,29,30–32] which undergo Maillard reactions during roasting, thus providing the characteristic aroma constituents of cocoa products.[20,21] Voigt *et al*[33] showed that the yield of aroma precursors depends strongly on the specificity of the endopeptidase to hydrolyse the cocoa globulins at hydrophobic amino acid residues. But if the globulins are cleaved under conditions where only aspartic endopeptidase is active, the resulting hydrophobic peptides are not capable of developing a chocolate-like aroma.[29] The aroma emerges only if the carboxy-terminal hydrophobic amino acid residues of these peptides are hydrolysed.[26] Thus, the carboxypeptidase's specificity to cleave hydrophobic amino acids is crucial for the generation of aroma precursors.[24] This enzyme specificity, being a premise for the potential to form a chocolate-like aroma, was detected in both *T bicolor* and *T grandiflorum* seeds.

The pattern of proteins obtained from cocoa cotyledons was similar to that described in the literature[15,34] except for the 14.5-kDa globulin, which we found only in very low relative concentrations. The observed similarity of the globulin pattern between the studied relatives can be explained by the sequence homology of the vicilin genes within the genus *Theobroma* which amounts to 99%.[7] However, the quantity of the respective peptides varied markedly among the studied species. Thus, it is expected that under optimised fermentation conditions, *T grandiflorum* possesses the potential to generate about half the amount of aroma precursors that are present in properly fermented seeds of *T cacao*. From this point of view *T bicolor* can be considered to have a higher potential to generate a chocolate-like aroma.

In general, the results demonstrate that *T grandiflorum* and *T bicolor* meet the biochemical requirements to generate a chocolate-like aroma. Fermentation technologies could be adapted to the proteolytic characteristics of *T grandiflorum* by setting a slightly lower pH value than that in cocoa fermentation. In a first trial we put this insight into practice on a small scale (batches of 20 kg of fresh seeds) by starting the fermentation in polypropylene woven bags (70cm × 100mm) which are commonly used by farmers for the transport of bulk goods. This treatment caused an exhaustive anaerobic initial phase of fermentation which was maintained for 3 days at an ambient temperature of about 25 °C. During this time the mass reached a maximum temperature of 39 °C. Afterwards the aerobic fermentation phase was initiated by transferring the seeds to a wooden box where they were mixed once a day. During the 3 days of aerobic fermentation the temperature of the mass rose to 42 °C.

This procedure of the 'sack fermentation' appears to be promising as it allows for two mechanisms leading to lowered pH values: higher yields of ethanol as a basis for the formation of sufficient acetic acid and a prolonged time span for the diffusion of acetic acid into the cotyledons. Prolongation of the fermentation could additionally compensate for the lower activity of the endopeptidase of *T grandiflorum* in comparison to that of cocoa seeds. In addition, the use of polypropylene woven bags prevented humidity loss from the residual pulp adhering to the seeds after the fruit flesh had been mechanically removed for its commercialisation. The exhaustive and prolonged anaerobic initial phase during fermentation of *T grandiflorum* seeds led to palatable chocolate aroma after roasting.

ACKNOWLEDGEMENTS

The studies have been done as part of the German Brazilian SHIFT project—Studies on Human Impacts on Forests and Floodplains in the Tropics. The German partner was sponsored by the Federal Ministry of Education and Research (BMBF—Förderkennzahl 0339457A, project ENV23/2). The Brazilian partner was sponsored by the Council for Scientific and Technological Development (CNPq).

We thank Professor Dr Böhle Biehl of the Technical University of Braunschweig for many fruitful discussions.

REFERENCES

1 Ducke A, As espécies brasileiras do gênero *Theobroma* L. *Boletim do Inst Agrón Norte* **28**:3–20 (1953).

2 Cuatrecasas J, Cacao and its allies: A taxonomic revision of the genus *Theobroma*. *Bulletin of the United States National Museum* **35**:377–614 (1964).
3 Hammerstone JF Jr, Romanczyk LJ Jr and Aitken M, Purine alkaloid distribution within *Herrania* and *Theobroma*. *Phytochemistry* **35**:1237–1240 (1994).
4 Carpenter DR, Hammerstone JF Jr, Romanczyl LJ Jr and Martin Aitken W, Lipid composition of *Herrania* and *Theobroma* seeds. *J Am Oil Chem Soc* **71**:845–851 (1994).
5 Gilabert-Escrivá MV, Gonçalves LAG, Silva CRS and Figueira A, Fatty acid and triacylglycerol composition and thermal behaviour of fats from seeds of Brazilian Amazonian *Theobroma* species. *J Sci Food Agric* **82**:1425–1431 (2002).
6 Venturieri GA, *Cupuaçu: a Espécie, sua Cultura, Usos, e Processamento*. Clube do cupu, Belém, 108 pp (1993).
7 Whitlock BA and Baum DA, Phylogenetic relationships of *Theobroma* and *Herrania* (Sterculiaceae) based on sequences of the nuclear gene vicilin. *Syst Bot* **24**:128–138 (1999).
8 FAO, Food and fruit-bearing forest Species—3: Examples from Latin America. *FAO-Forestry Paper* **44**:289–308 (1986).
9 Venturieri GA and Aguiar JPL, Composiçao do chocolate caseiro de amêndoas de cupuaçu (*Theobroma grandiflorum* (Willd ex Spreng) Schum). *Acta Amazonica* **18**:3–8 (1988).
10 Ribeiro De Nazaré RF, Barbosa WC and Viegas RMF, Processamento das sementes de Cupuaçu para a obtençao de cupulate, *Boletim de Pesquisa*, No **108**; *EMBRAPA-CPATU, Belém, Brazil*: 1–38 (1990).
11 Biehl B, Cocoa fermentation and problem of acidity, overfermentation and low cocoa flavour, in *1984 International conference on cocoa and coconuts: progress and outlook,* ed by Pushparajah E and Chew Poh Soon, Inc Soc Planters, Kuala Lumpur, pp 561–568 (1986).
12 Biehl B and Adomako D, Die Kakaofermentation (Steuerung, Acidation, Proteolyse). *Lebensmittelchem Gerichtl Chem* **37**:57–63 (1983).
13 Biehl B, Passern U and Passern D, Subcellular structures in fermenting cocoa beans. Affect of aeration and temperature during seed and fragment incubation. *J Sci Food Agric* **28**:41–52 (1977).
14 Biehl B, Passern D and Sagemann W, Effect of acetic acid on subcellular structures of cocoa bean cotyledons. *J Sci Food Agric* **33**:1101–1109 (1982).
15 Voigt J, Biehl B and Wazir SKS, The major seed proteins of *Theobroma cacao* L. *Food Chem* **47**:145–151 (1993).
16 Voigt J, Wrann D, Heinrichs H and Biehl B, The proteolytic formation of essential cocoa-specific aroma precursors depends on particular chemical structures of the vicilin-class globulin of the cocoa seeds lacking in the globular storage proteins of coconuts, hazelnuts and sunflower seeds. *Food Chem* **51**:197–205 (1994).
17 Biehl B, Heinrichs H, Ziegler-Berghausen H, Srivastava S, Xiong Q, Passern D, Senyuk VI and Hammoor M, The proteases of ungerminated cocoa seeds and their role in the fermentation process. *J Appl Bot-Angew Bot* **67**:59–65 (1993).
18 Mohr W, Röhrle M and Severin T, Über die Bildung des Kakaoaromas aus seinen Vorstufen. *Fette Seifen Anstrichmittel* **73**:515–521 (1971).
19 Mohr W, Landschreiber E and Severin T, Zur Spezifität des Kakaoaromas. *Fette Seifen Anstrichmittel* **78**:88–95 (1976).
20 Ziegleder G, Untersuchungen über die Reaktion von Aromavorstufen des Kakaos. *Proceedings of the 1st Symposion on useful plants of the Tropics, Sep 22–24 1995, Berichte der Angewandten Botanik* **5**:148–151 (1995).
21 Oberparleiter S and Ziegleder G, Amadori-compounds as cocoa aroma precursors. *Nahr-Food* **41**:142–145 (1997).
22 Kirchhoff PM, Biehl B and Crone G, Peculiarity of the accumulation of free amino acids during cocoa fermentation. *Food Chem* **31**:295–311 (1989).
23 Hansen CE, del Olmo M and Burri C, Enzyme activities in cocoa beans during fermentation. *J Sci Food Agric* 77:273–281 (1998).
24 Bytof G, Biehl B, Heinrichs H and Voigt J, Specificity and stability of the carboxypeptidase activity in ripe, ungerminated seeds of *Theobroma cacao* L. *Food Chem* **54**:15–21 (1995).
25 Shutov AD, Do Ngok Lanh and Vaintraub IA, Purification and partial characterization of protease B from germinating vetch seeds. *Biokhimya* **47**:814–821 (1982).
26 Voigt J, Biehl B, Heinrichs H, Kamaruddin S, Gaim Marsoner G and Hugi A, In-vitro formation of cocoa-specific aroma precursors: aroma-related peptides generated from cocoa-seed protein by co-operation of an aspartic endoprotease and a carboxypeptidase. *Food Chem* **49**:173–180 (1994).
27 Voigt J, Kamaruddin S, Heinrichs H, Wrann D, Senyuk V, and Biehl B, Developmental stage-dependent variation of the level of globular protein and aspartic endoprotease during ripening and germination of *Theobroma cacao* L seeds. *J Plant Physiol* **145**:299–307 (1995).
28 Lerceteau E, Rogers J, Pètiard V and Crouzillat D, Evolution of cacao bean proteins during fermentation: a study by two-dimensional electrophoresis. *J Sci Food Agric* **79**:319–625 (1999).
29 Voigt J and Biehl B, Precursors of the cocoa-specific aroma components are derived from the vicilin-class (7s) globulin of the cocoa seeds by proteolytic processing. *Bot Acta* **108**:283–289 (1995).
30 Rohan TA and Stewart T, The precursors of chocolate aroma: the distribution of free amino acids in different commercial varieties of cocoa beans. *J Food Sci* **30**:416–419 (1965).
31 Rohan TA and Stewart T, The precursors of chocolate aroma: changes in the free amino acids during the roasting of cocoa beans. *J Food Sci* **31**:202–205 (1966).
32 Rohan TA and Stewart T, The precursors of chocolate aroma: production of free amino acids during fermentation of cocoa beans. *J Food Sci* **32**:395–398 (1967).
33 Voigt J, Heinrichs H, Wrann D and Biehl B, In vitro studies on the proteolytic formation of the characterisitic aroma precursors of fermented cocoa seeds: the significance of endoprotease specificity. *Food Chem* **51**:7–14 (1994).
34 Pettipher GL, The extraction and partial purification of cocoa storage proteins. *Café Cacao Thé* **34**:23–26 (1990).