Food Chemistry **47** (1993) 145–151

The major seed proteins of *Theobroma cacao* L.

Jürgen Voigt,* Böle Biehl

Botanisches Institut der Technischen Universität Braunschweig, Mendelssohnstraße 4, W-3300 Braunschweig, Germany

&

Syed Kamaruddin Syed Wazir

Malaysian Agricultural Research and Development Institute (MARDI), Cocoa and Coconut Research, Division Hilir Perak, Malaysia

(Received 1 July 1992; revised version received and accepted 25 September 1992)

Differential extractions of proteins from *Theobroma cacao* seeds have revealed the presence of an albumin fraction and a globulin fraction with proportions of 52% and 43%, respectively, of total seed proteins. In contrast to some earlier reports, we could not detect any prolamin. The 'glutelin fraction' described in the literature was found to consist of residual globulin. After fermentation, the first step in cocoa processing, the proportion of the globulin fraction is considerably reduced. The major albumin is a polypeptide with an apparent molecular weight of 19 kDa. The globulin fraction contained polypeptides with apparent molecular sizes of 47 kDa, 31 kDa, and 14·5 kDa. Globulin prepared in the absence of the aspartyl protease inhibitor pepstatin contained two additional polypeptides with apparent molecular sizes of 28 kDa and 16 kDa, respectively. The negative globulin of *Theobroma cacao* is a glycoprotein with a sedimentation coefficient of 7–8S and a molecular weight of 150 kDa. Its subunits are not cross-linked by disulphide bridges—in contrast to the legumin-like storage globulins which are predominant in the seeds of all other dicotyledons studied so far. Therefore, *Theobroma cacao* is the first plant described to date whose seeds contain a vicilin-like globulin, but apparently no legumin-class globulin.

INTRODUCTION

During their development, plant seeds accumulate large amounts of storage proteins that serve as sources of nitrogen, sulphur and carbon compounds during seed germination (Shotwell & Larkins, 1988). Storage proteins accumulate in membrane-delimited organelles, the protein bodies (Briarty *et al.*, 1970; Pernollet, 1978).

As shown by Osborne (1924), the major seed proteins can be classified on an operational basis into four solubility classes: albumin (water-soluble), globulin (salt-soluble), prolamin (alcohol-soluble) and glutelin (soluble in dilute acids or alkali). Plant seeds of different taxa contain rather different proportions of albumins, globulins, prolamins and glutelins (Higgins, 1984; Bewley & Black, 1985; Shotwell & Larkins, 1988). In most monocotyledonous seeds studied so far, prolamins are the major seed proteins (Bewley & Black, 1985), whereas in dicotyledons (Higgins, 1984; Bewley & Black, 1985) and gymnosperms (Misra & Green, 1990) globulins predominate. On the basis of their amino acid sequences, subunit compositions, and the processing of the corresponding polypeptide precursors, the globulins studies so far can be assigned to two different classes: the legumin-like and the vicilin-like globulins (Derbyshire *et al.*, 1976; Higgins, 1984; Shotwell & Larkins, 1988).

The seeds of *Theobroma cacao* have been reported to contain albumins, globulins, prolamin and glutelin, the albumin being the major seed protein fraction (Niepage, 1961; Zak & Keeney, 1976*a,b*). Recently, cDNAs encoding the major albumin and a globulin of cocoa seeds, respectively, have been cloned and sequenced (Spencer & Hodge, 1991, 1992; Tai *et al.*, 1991; McHenry & Fritz, 1992). The deduced amino acid sequence of the albumin is homologous with the Kunitz protease and α-amylase inhibitor family (Spencer & Hodge, 1991; Tai *et al.*, 1991). The amino acid sequence of the 65 kDa precursor for the cocoa globulin deduced from the nucleotide sequence of the cloned cDNA revealed considerable sequence homologies with vicilin-class globulins of some dicotyledonous plants, especially with an α-globulin of cotton seeds (McHenry & Fritz, 1992; Spencer & Hodge, 1992). Two major polypeptides of cocoa seeds with apparent molecular sizes of 47 kDa and 31 kDa, respectively, are found to

* To whom correspondence should be addressed.

Food Chemistry 0308-8146/93/$06.00

be derived from these precursors (Spencer & Hodge, 1992). Several years ago, two polypeptides with similar apparent molecular masses (44 kDa and 26 kDa) were definitely shown to be vacuolar storage proteins (Biehl *et al.*, 1982). In contrast to the major albumin and the vicilin-class globulin, there are, however, no data concerning the prolamin and the glutelin which have been found in the seeds of *Theobroma cacao* (Niepage, 1961; Zak & Keeney, 1976*a,b*). Furthermore, there are no data concerning the presence or absence of legumin-like globulins which are the predominant globulins in the seeds of other dicotyledonous plants (Higgins, 1984; Bewley & Black, 1985).

The storage proteins of *Theobroma cacao* seem to be important with respect to the formation of the cocoa-specific flavour. During fermentation, the first step in cocoa processing which is essential for flavour formation (Rohan, 1964; Ziegleder & Biehl, 1988). a selective degradation of vacuolar storage proteins has been observed (Biehl *et al.* 1982). Furthermore, it has been shown that both selectivity and degree of proteolytic processes occurring in the seeds during fermentation and flavour potential of the finally obtained raw cocoa are substantially influenced by the fermentation conditions, especially by the time course of acidification caused by the action of microorganisms on the pulp (Biehl *et al.*, 1985). Flavour precursors have been extracted from fermented cocoa seeds, fractionated and classified (Mohr *et al.*, 1971, 1976). Peptides from fermented cocoa seeds, free amino acids and reducing sugars are required to obtain cocoa-specific flavour during roasting (Mohr *et al.*, 1971, 1976). These findings indicate that the flavour-related peptides are derived from specific seed proteins (Rohan, 1964; Biehl *et al.*, 1982, 1985; Ziegleder & Biehl, 1988). Therefore, we have re-investigated the major seed proteins of *Theobroma cacao* in order to classify those seed proteins from which the cocoa-specific flavour precursors might be derived.

Fig. 1. SDS–PAGE profiles of the protein fractions obtained by solubility fractionation of proteins from unfermented cocoa seeds. The protein fractions were obtained by solubility fractionation of acetone dry powder (AcDP) from unfermented cocoa seeds, as described in Materials and Methods, and 30 μg of each fraction—with the exception of the 'prolamin fraction'—were subjected to SDS/polyacrylamide gel electrophoresis according to Laemmli (1970). The gels were stained with Coomassie brilliant blue. The molecular weight of protein standards (M_r) is indicated by the numbers on the right.

MATERIALS AND METHODS

Materials

Cocoa seeds were from ripe, genetically-undefined pods harvested at the Cocoa and Coconut Division of the Malaysian Agricultural Research and Development Institute (MARDI, Hilir Perak, Malaysia). Unfermented seeds were taken from the pods immediately after arrival (4–5 days after harvesting), shock-frozen after removal of testae and radiculae and freeze-dried. Fermented and air-dried, but unroasted cocoa seeds (high quality standard; Ivory Coast) were generously supplied by Dr Hemmerich (Schokinag, Mannheim, Germany).

Extraction of fat

The dry cotyledons were crushed and extracted repeatedly in a Soxhlet apparatus with petroleum ether (bp 40–70°C). After solvent evaporation, the material was powdered and extracted again for 8 h in the same manner. Finally, purine alkaloids were partially extracted with chloroform for 8 h in a Soxhlet apparatus.

Preparation of acetone dry powder (AcDP)

Acetone dry powder (AcDP) of cocoa seeds was prepared essentially as recently described (Kirchhoff *et al.*, 1989). To remove the polyphenols, the defatted powders from unfermented or fermented seeds were extracted three times with 80% (v/v) aqueous acetone containing—unless otherwise stated—5 mM sodium ascorbate and subsequently with 70% (v/v) aqueous

acetone. The suspensions (200 ml aqueous acetone per 10 g seed powder) were stirred for 1 h at 0°C and the extracts removed by centrifugation (15 min at 13 000*g*). After the final extraction step, efficiency of polyphenol extraction was checked by heating an aliquot of the acetone dry powder with 5 M HCl (red colour indicates the presence of residual polyphenols). After complete extraction of polyphenols, residual water was removed by extraction with 100% acetone. After final centrifugation, the sediment was evaporated under reduced pressure to remove the residual acetone. The acetone dry powder (AcDP) was stored at −20°C.

Extraction of seed proteins

Proteins were extracted and fractionated into various solubility classes following a modification of the procedure described by Hu and Esen (1981). The AcDP was extracted successively with 10 mM Tris-HCl (pH 7·5, containing 2 mM EDTA), 0·5 M NaCl (containing 2 mM EDTA and 10 mM Tris-HCl, pH 7·5), 70% (v/v) ethanol and 0·1 N NaOH, to obtain the albumin, globulin, prolamin and glutelin fractions, respectively. Unless otherwise stated, all these solvents contained 5 mM sodium ascorbate. To avoid proteolytic digestion of the seed proteins during the extraction procedure, protease inhibitors (10 μM pepstatin A and 1 mM PMFS) were added to all the different solvents. Repeated extractions with the different solvents (four times each) were found to be necessary to avoid a carry-over to the next fraction as already shown by Krochko *et al.* (1990).

Determination of protein

Aliquots of the different protein fractions were adjusted to 10% (w/v) TCA and chilled in an ice bath for 30 min to precipitate protein. Following centrifugation at 15 000*g* for 20 min, the precipitates were washed repeatedly with bidistilled water and solubilized in 1 M NaOH. Protein concentrations were then determined by the method of Lowry *et al.* (1951) using bovine serum albumin as standard.

Table 1. Proportions of the different solubility classes of seed proteins in unfermented and fermented cocoa seeds

Seed protein fraction[a]	Unfermented seeds		Fermented seeds	
	(mg/g AcDP)	(%)	(mg/g AcDP)	(%)
Albumin	236 ± 15	52 ± 3·3	172 ± 22	79 ± 10·1
Globulin	197 ± 14	43 ± 3·1	18 ± 8	8·3 ± 3·7
Prolamin	1	—	1	—
Glutelin	24 ± 5	5 ± 1·0	28 ± 8	12·8 ± 3·7

[a] The seed protein fractions were prepared as described in Materials and Methods, starting from polyphenol-free acetone dry powder in the case of unfermented seeds. The acetone dry powder prepared from fermented seeds was not polyphenol-free, due to the irreversible binding of polyphenols/quinones to proteins during fermentation and air drying. These irreversibly bound components might interfere with the determination of protein according to Lowry *et al.* (1951), although the samples were subjected to TCA precipitation prior to the analysis.

Sucrose gradients

Sedimentation equilibrium analysis of cocoa seed globulin was accomplished on 10–30% sucrose gradients (in 0·5 M NaCl–1 mM DTE–20 mM Tris-HCl, pH 8·0) spun in a Beckman SW27 rotor at 22 000*g* for 24 h at 15°C. Sunflower globulin (12S), pea vicilin (7S) and bovine liver catalase (2.2S) were used as calibration standards. Sunflower globulin and pea vicilin were prepared from the corresponding seeds as described in the literature (Derbyshire *et al.*, 1976). Bovine liver catalase was obtained from Sigma Chemie (Deisenhofen, Germany).

Polyacrylamide gel electrophoresis

One-dimensional electrophoretic separations were performed by SDS/polyacrylamide gel electrophoresis on gel slabs containing 12·5% (w/v) acrylamide according to Laemmli (1970). The protein fractions to be analysed were subjected to TCA precipitation and the precipitates repeatedly washed with bidistilled water and redissolved in urea–SDS buffer containing 8 M urea, 2% (w/v) SDS, 10 mM EDTA, 20 mM Tris-HCl (pH 7·5) and—unless otherwise stated—200 mM 2-mercaptoethanol (Voigt, 1985). After addition of 1/4 volume of sample buffer according to Laemmli (1970) containing 0·005% (w/v) bromophenol blue as tracking dye, the samples were subjected to SDS-PAGE analysis. Gels were either stained for protein with Coomassie brilliant blue R-250 or stained for glycoproteins using the periodic-acid–Schiff (PAS) technique (Zacharius *et al.*, 1969).

RESULTS

Analysis of the solubility classes of seed proteins of unfermented *Theobroma cacao* seeds revealed the presence of albumin and globulin (Table 1). The albumin was found to be the predominant protein fraction (52% of total seed proteins). For the globulin fraction, a proportion of 43% of total seed proteins was found (Table 1). In contrast to some earlier reports (Niepage, 1961; Zak & Keeney, 1976*a,b*), we did not detect any prolamin (Fig. 1, Table 1). Furthermore, the proportion of the 'glutelin' fraction was considerably lower than previously reported (Table 1; *cf.* Niepage, 1961; Zak & Keeney, 1976*a,b*) It has to be pointed out that in the case of *Theobroma cacao* seeds, a complete extraction of polyphenols was a prerequisite for a reliable fractionation of the different solubility classes of seed proteins according to Osborne (1924). Otherwise, proteins were irreversibly denatured (tanned) by quinones formed during the fractionation procedure. Furthermore, each extraction had to be repeated three times to ensure an almost quantitative extraction of the

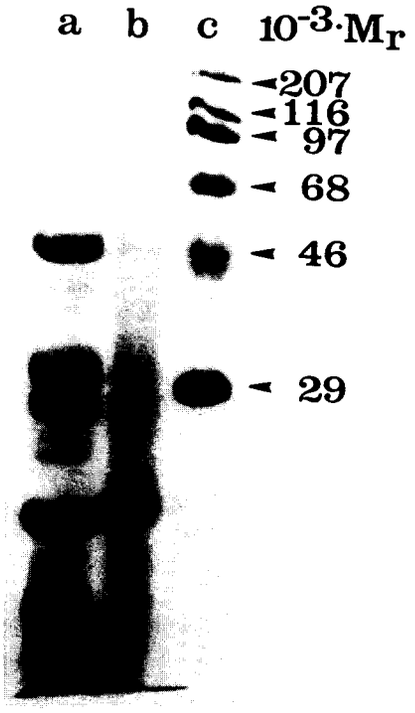

Fig. 2. Comparative SDS–PAGE analysis of total seed proteins from unfermented and fermented cocoa seeds. Acetone dry powders were prepared from unfermented and fermented cocoa seeds as described in Materials and Methods. Urea–SDS extracts (Voigt, 1985) were prepared and aliquots corresponding to 300 μg acetone dry powder were subjected to SDS/polyacrylamide gel electrophoresis according to Laemmli (1970). The gels were stained with Coomassie brilliant blue. Lane **a**: unfermented seeds; lane **b**: fermented seeds; lane **c**: protein molecular weight standards (M_r) indicated by the numbers on the right.

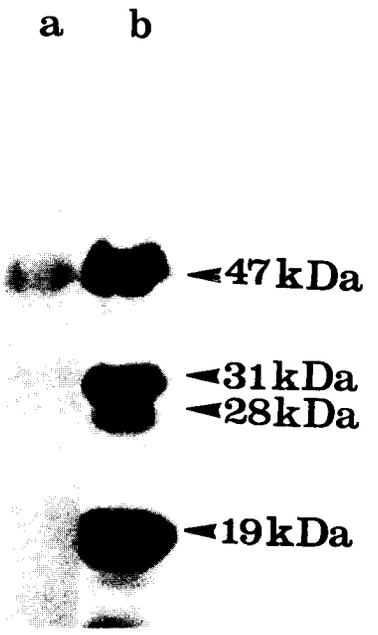

Fig. 3. Comparative analysis of the SDS-PAGE profiles of glycoproteins (lane **a**) and proteins (lane **b**) present in an unfractionated high-salt extract (0·5 M NaCl) of acetone dry powder from unfermented cocoa seeds. The gels were stained for glycoproteins using the periodic-acid–Schiff (PAS) technique (Zacharius *et al.*, 1969) or stained for proteins with Coomassie brilliant blue. As the periodic-acid–Schiff staining of glycoproteins is not very sensitive, 150 μg of seed proteins were subjected to SDS–PAGE analysis of glycoproteins. The molecular weight of protein standards (M_r) is indicated by the numbers on the right.

different protein fractions (*cf.* Krochko *et al.*, 1990). Comparative SDS–PAGE analysis of the different protein fractions (Fig. 1) revealed that the polypeptide patterns of the globulin and the 'glutelin' fractions were rather similar. Therefore, the 'glutelin' fraction of *Theobroma cacao* seeds consisted of residual globulin(s), even though the high salt extraction of globulin(s) was repeated several times. Polypeptides with apparent molecular weights of 47 kDa, 31 kDa and 14·5 kDa were the predominant components of the globulin fraction. Globulin prepared in the absence of the aspartyl protease inhibitor, pepstatin A, contained two additional polypeptides with apparent molecular sizes of 28 kDa and 16 kDa, respectively. The albumin fraction contained one predominant polypeptide with an apparent molecular size of 19 kDa (Fig. 1).

Comparative SDS-PAGE analyses of SDS-extracts from the same amounts of acetone dry powders of unfermented seeds and fermented and subsequently air-dried seeds revealed that the amount of SDS-soluble protein was reduced during fermentation and subsequent air-drying (Fig. 2). Considerable amounts of the 19 kDa albumin were still observed after fermentation and air-drying (Fig. 2, lane **b**), although to a variable extent. The globulin(s), however, were almost quantitatively lost after fermentation and air-drying (Table 1, Fig. 2). The quantitative loss of the globulin(s) is apparently due to a selective proteolytic digestion during fermentation (*cf.* Biehl *et al.*, 1982, 1985). These findings were further corroborated by SDS–PAGE analyses of the Osborne fractions from unfermented and fermented cocoa seeds (data not shown).

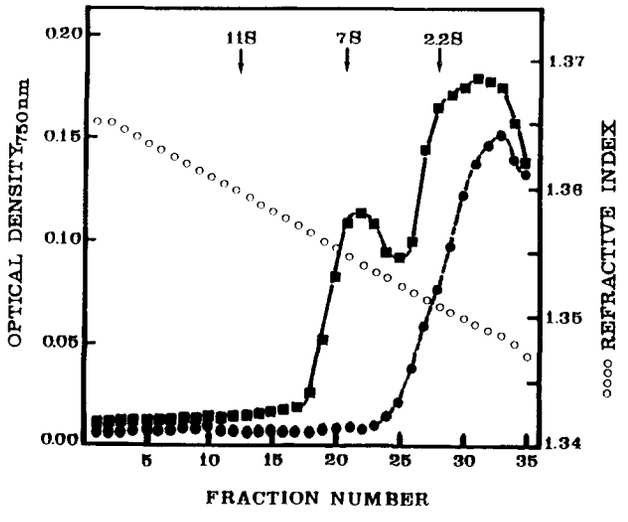

Fig. 4. Analytical sucrose gradient centrifugation of low-salt and high-salt extracts from cocoa seed AcDP. Extracts were adjusted to 0·5 M NaCl and subjected to centrifugation on 10–30% sucrose gradients. The optical densities are based on Lowry absorbance (Lowry *et al.*, 1951) of 20 μl aliquots. The calibration standards, bovine liver catalase (2.2S), pea vicilin (7S) and sunflower globulin (11S), sedimented as indicated by arrows. ●- - -● Albumin fraction; ■——■ globulin fraction.

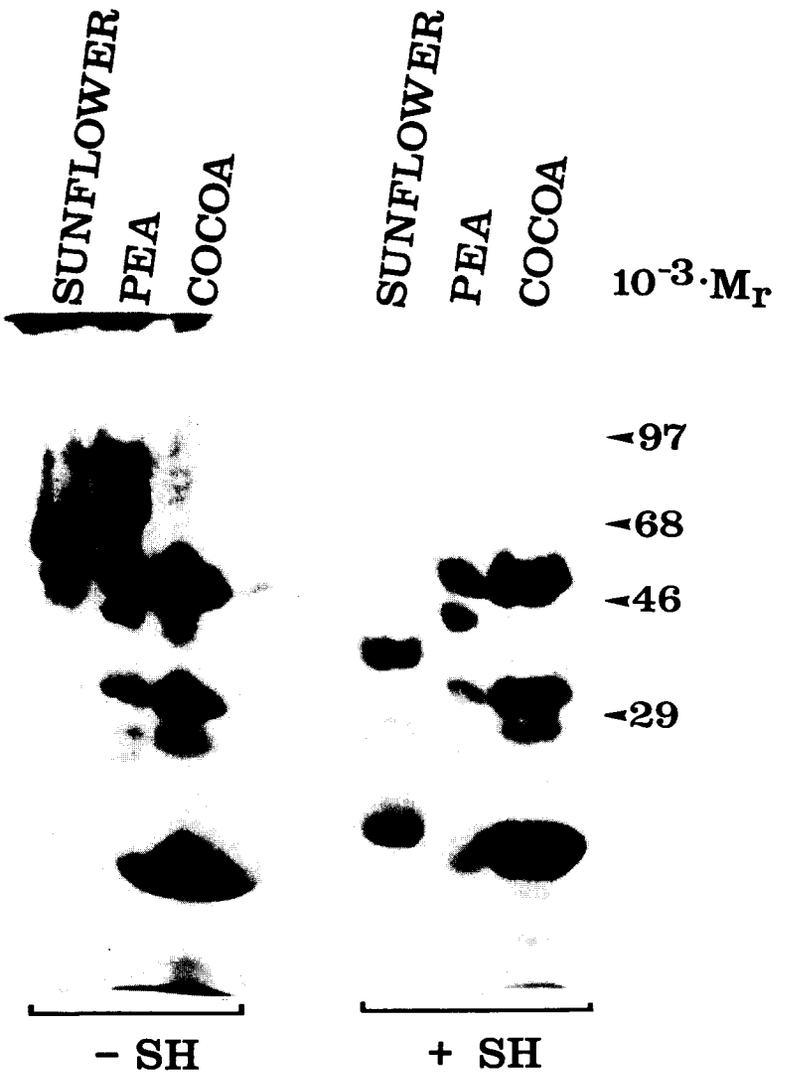

Fig. 5. Coomassie blue-stained SDS–PAGE profiles of total seed proteins from sunflower, pea and unfermented cocoa seeds run under non-reducing (−SH) and reducing (+SH) conditions. Acetone dry powders were prepared from sunflower, pea and unfermented cocoa seeds in the absence of compounds which are able to reductively cleave disulphide bonds. Urea–SDS extracts (Voigt, 1985) were prepared from the acetone dry powders in the absence (−SH) or presence (+SH) of 200 mM 2-mercaptoethanol and subjected to SDS/polyacrylamide gel electrophoresis according to Laemmli (1970). Pea seeds contain both legumin-like and vicilin-like globulins (Derbyshire *et al.*, 1976), whereas sunflower seeds only contain legumin-like globulin.

To analyse whether or not the major seed proteins of *Theobroma cacao* are glycosylated, a seed protein fraction obtained by extraction of acetone dry powder with 0·5 M NaCl was subjected to SDS/polyacrylamide gel electrophoresis and parallel sections of the gels were subsequently stained for glycoproteins using the periodic-acid–Schiff (PAS) technique (Fig. 3, lane **a**) and for protein (Fig. 3, lane **b**), respectively. The most significant stain for sugar residues was detected in the large subunit (47 kDa) of the globulin (*cf.* Fig. 1). The 19 kDa albumin was not stained for glycoprotein by the periodic-acid–Schiff technique (Fig. 3).

Sucrose density gradient analysis of the albumin and the globulin fractions of cocoa seeds revealed one peak for the albumin and two peaks for the globulin fraction, respectively (Fig. 4). The globulin peaks were determined to have S values of 6–7S and 1–2S, respectively. The predominant protein of the albumin fraction has a sedimentation coefficient of about 1S.

Legumin-like globulins contain polypeptide subunits which are cross-linked via disulphide bonds, whereas the polypeptide subunits of vicilin-like globulins are not cross-linked via disulphide bonds (Derbyshire *et al.*, 1976). Therefore, the Coomassie blue stained profiles obtained after separation of seed proteins in their non-reduced (−ME) and reduced (+ME) forms are shown in Fig. 5. As controls, pea and sunflower globulins were prepared and analysed parallel to the globulin from unfermented cocoa seeds. The gel shown in Fig. 5 clearly demonstrates that in contrast to the pea globulin fraction (consisting of a mixture of legumin and vicilin) and the sunflower globulin fraction, which only contains a 12S globulin of the legumin type, the protein profile of the cocoa globulin fraction is not influenced by the presence or absence of 2-mercaptoethanol or dithiothreitol. Therefore, the polypeptide subunits of the major cocoa globulin(s) are not cross-linked by disulphide bonds. This finding is further corroborated by two-dimensional electrophoresis: SDS–PAGE under non-reducing conditions in the first dimension,

followed by electrophoresis under reducing conditions (data not shown).

DISCUSSION

The predominant proportion of the seed proteins of *Theobroma cacao* is an albumin with an apparent molecular weight of 19 kDa. This predominant 19 kDa albumin is presumably identical with the 21 kDa albumin, whose amino acid sequence (deduced from the nucleotide sequence of cloned cDNA) was recently reported (Spencer & Hodge, 1991; Tai *et al.*, 1991). Besides the albumin, considerable amounts of globulin have been found (Fig. 1, Table 1). In contrast to some earlier reports (Niepage, 1961; Zak & Keeney, 1976*a,b*), we could not detect a prolamin fraction (Fig. 1, Table 1). Furthermore, we could not confirm the existence of glutelin(s) in cocoa seeds reported by some authors (Niepage, 1961; Zak & Keeney, 1976*a,b*).

In the case of cocoa seeds, fractionation of seed proteins according to Osborne (1924) and determinations of the relative amounts of proteins of the different solubility classes can be interfered with by three experimental artifacts:

(1) irreversible denaturation of proteins by quinones resulting from oxidation of polyphenols which are present in considerable amounts in cocoa seeds;

(2) proteolytic digestions of seed proteins which might occur during the differential extraction (in the absence of protease inhibitors and especially at low pH and high temperature), because ungerminated cocoa seeds contain exceptionally high levels of protease activities (Biehl *et al.*, 1982, 1991);

(3) carry-over of proteins to the next fraction, if the extractions are not repeated with the same solvent several times.

In particular, the complete removal of polyphenols is a prerequisite for the fractionation of cocoa seed proteins according to Osborne (1924). This point was not sufficiently considered in previous reports (Niepage, 1961; Zak & Keeney, 1976*a,b*; Timbie & Keeney, 1977; Fritz *et al.*, 1985). Therefore, the resulting polypeptide patterns were rather different from the polypeptide patterns obtained after complete removal of polyphenols (Biehl *et al.*, 1982; Pettipher, 1990; Spencer & Hodge, 1991; also the present paper). The globulin fraction contains only vicilin-type globulins (Spencer & Hodge, 1992; also the present paper), whereas all the seeds of other plants studied so far contain either legumins or both legumins and vicilins (Higgins, 1984; Bewley & Black, 1985; Shotwell & Larkins, 1988). This finding is important from a phylogenetic viewpoint and with respect to the formation of cocoa-specific flavour precursors during fermentation of cocoa seeds (Rohan, 1964; Mohr *et al.*, 1971, 1976; Biehl *et al.*, 1982, 1985; Ziegleder & Biehl, 1988).

The vicilin-like globulin was found to be selectively degraded by endogenous proteases under fermentation conditions where optimal levels of aroma precursors were obtained (Biehl *et al.*, 1982, 1985; also the present paper). Therefore, the flavour-related peptides which are formed during fermentation and are assumed to be responsible for the formation of cocoa-specific flavour components during roasting (Ziegleder & Biehl, 1988) seem to be formed by proteolytic digestion of the cocoa vicilin.

ACKNOWLEDGEMENT

This work has been supported by a grant (AIF 8427) from the Arbeitsgemeinschaft Industrieller Forschungsvereinigungen e.V./Forschungskreis der Ernährungsindustrie (FEI).

REFERENCES

Bewley, J. D. & Black, M. (1985). In *Seeds: Physiology of Development and Germination*, ed. J. D. Bewley & M. Black. Plenum Press, New York, pp. 29–83.

Biehl, B., Wewetzer, C. & Passern, D. (1982). Vacuolar (storage) proteins of cocoa seeds and their degradation during germination and fermentation. *J. Sci. Food Agric.*, **33**, 1291–304.

Biehl, B., Brunner, E., Passern, D., Quesnel, V. C. & Adomako, D. (1985). Acidification, proteolysis and flavour potential in fermenting cocoa beans. *J. Sci. Food Agric.*, **36**, 583–98.

Biehl, B., Ziegeler-Berghausen, H., Srivastava, S., Xiong, Q., Passern, D., Heinrichs, H. & Hammoor, M. (1991) Cocoa specificity of proteolytic flavour precursors: the cocoa seed proteases. *Proc. 1991 Int. Cocoa Conf.: Challenges in the 90s*, 25–27 September 1991. Kuala Lumpur, Malaysia.

Briarty, L. G., Coult, D. A. & Boulter, D. (1970). Protein bodies of germinating seeds of *Vicia faba*: Changes in fine structure and biochemistry. *J. Exp. Bot.*, **21**, 513–24.

Derbyshire, E., Wright, D. J. & Boulter, D. (1976). Legumin and vicilin, storage proteins of legume seeds. *Phytochemistry*, **15**, 3–24.

Fritz, P. J., Fritz, K. A., Kaufman, J. M., Patterson, G. R., Robertson, C. A., Stoesz, D. A. & Wilson, M. R. (1985). Cocoa seeds: changes in protein and polysomal RNA during development. *J. Food Sci.*, **50**, 946–50.

Higgins, J. V. (1984). Synthesis and regulation of major proteins in seeds. *Annu. Rev. Plant Physiol.*, **35**, 191–221.

Hu, B. & Esen, A. (1981). Heterogeneity of soybean seed proteins: one dimensional electrophoretic profiles of six different solubility fractions. *J. Agric. Food Chem.*, **29**, 497–501.

Kirchhoff, P.-M., Biehl, B. & Crone, G. (1989). Peculiarity of the accumulation of free amino acids during cocoa fermentation. *Food Chem.*, **31**, 295–311.

Krochko, J. E., Charbonneau, M. R., Coulter, K. M., Bowley, S. R. & Bewley, J. D. (1990). A comparison of seed storage proteins in subspecies and cultivars of *Medicago sativa*. *Can J. Bot.*, **68**, 940–8.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–5.

Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951). Protein measurements with the folin phenol reagent. *J. Biol. Chem.*, **193**, 257–65.

McHenry, L. & Fritz, P. J. (1992). Comparison of the structure and nucleotide sequence of vicilin genes of cocoa and cotton raise questions about vicilin evolution. *Plant. Mol. Biol.*, **18**, 1173–6.

Misra, S. & Green, M. J. (1990). Developmental gene expres-

sion in conifer embryo genesis and germination. I. Seed proteins and protein body composition of mature embryo and the megagametophyte of white spruce (*Picea glauca* (Moench) Voss.). *Plant Sci.*, **68**, 163–73.

Mohr, W., Röhrle, M. & Severin, Th. (1971). Über die Bildung des Kakaoaromas aus seinen Vorstufen. *Fette, Seifen, Anstrichmittel*, **73**, 515–21.

Mohr, W., Landschreiber, E. & Severin, Th. (1976). Zur Spezifität des Kakaoaromas. *Fette, Seifen, Anstrichmittel*, **78**, 88–95.

Niepage, N. (1961). Versuche zur Fraktionierung und Bausteinanalyse der Proteine in unfermentierten und fermentierten Kakaobohnen. *Gordian*, **1445**, 13–20.

Osborne, T. B. (1924). *The Vegetable Proteins*, 2nd edn. Longmans, Green & Co., London.

Pernollet, J.-C. (1978). Protein bodies of seeds: ultrastructure, biochemistry, biosynthesis and degradation. *Phytochemistry*, **17**, 1473–80.

Pettipher, G. L. (1990). The extraction and partial purification of cocoa storage proteins. *Café Cacao Thé*, **34**, 23–6.

Rohan, T. A. (1964). The precursors of chocolate aroma: a comparative study of fermented and unfermented cocoa beans. *J. Food Sci.* , **29**, 456–9.

Shotwell, M. A. & Larkins, B. A. (1988). The biochemistry and molecular biology of seed storage proteins. In *The Biochemistry of Plants. A Comprehensive Treatise*, Vol. 15, ed. A. Marcus. Academic Press, New York, pp. 297–345.

Spencer, M. E. & Hodge, R. (1991). Cloning and sequencing of the cDNA encoding the major albumin of *Theobroma cacao*. *Planta*, **183**, 528–35.

Spencer, M. E. & Hodge, R. (1992). Cloning and sequencing of a cDNA encoding the major storage proteins of *Theobroma cacao*. Identification of the proteins as members of the vicilin class of storage proteins. *Planta*, **186**, 567–76.

Tai, H., McHenry, L., Fritz, P. J. & Furtek, D. B. (1991). Nucleic acid sequence of a 21 kDa cocoa seed protein with homology to the soybean trypsin inhibitor (Kunitz) family of protease inhibitors. *Plant Mol. Biol.*, **16**, 913–15.

Timbie, D. J. & Keeney, P. G. (1977). Extraction, fractionation, and amino acid composition of brazilian comun cacao proteins. *J. Agric. Food Chem.*, **25**, 424–6.

Voigt, J. (1985). Extraction by lithium chloride of hydroxyproline-rich glycoproteins from intact cells of *Chlamydomonas reinhardtii*. *Planta*, **164**, 379–89.

Zacharius, R. M., Zell, T. E., Morrison, J. M. & Woodlock, J. J. (1969). Glycoprotein staining following electrophoresis on acrylamide gels. *Anal. Biochem.*, **30**, 148–52.

Zak, D. K. & Keeney, P. G. (1976*a*). Extraction and fractionation of cocoa proteins as applied to several varieties of cocoa beans. *J. Agric. Food Chem.*, **24**, 479–82.

Zak, D. K. & Keeney, P. G. (1976*b*). Changes in cocoa proteins during ripening of fruit, fermentation, and further processing of cocoa beans. *J. Agric. Food Chem.*, **24**, 483–6.

Ziegleder, G. & Biehl, B. (1988). Analysis of cocoa flavour components and flavour precursors. In *Methods in Plant Analysis*, Vol. 8, new series, ed. H. Linskens and J. F. Jackson. Springer-Verlag, Berlin, Heidelberg and New York, pp. 321–93.