Food Research International 90 (2016) 53–65

Contents lists available at ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

Biochemical fate of vicilin storage protein during fermentation and drying of cocoa beans

Neha Kumari [a], Kouame Jean Kofi [b], Sergio Grimbs [a], Roy N. D'Souza [a], Nikolai Kuhnert [a], Gino Vrancken [b], Matthias S. Ullrich [a,*]

[a] Department of Life Sciences and Chemistry, Jacobs University, Campus Ring 1, 28759 Bremen, Germany
[b] Cocoa Science Team, Barry Callebaut, Aalstersestraat 122, 9280 Lebbeke-Wieze, Belgium

ARTICLE INFO

Article history:
Received 19 August 2016
Received in revised form 17 October 2016
Accepted 23 October 2016
Available online 26 October 2016

Keywords:
Theobroma cacao
Vicilin
Peptides
Glycosylation
Phosphorylation
MALDI-MS
UHPLC-ESI-MS/MS

ABSTRACT

Key cocoa-specific aroma precursors are generated during the fermentation of cocoa beans via the proteolysis of the vicilin-like globulin. Previous studies had shown that degradation of this particular 566 amino acid-long storage protein leads to three distinct subunits with different molecular masses. Although oligopeptides generated from the proteolysis of vicilin-like globulin have been studied previously, changes occurring to vicilin at different stages of fermentation have not yet been explored in detail. The aim of this study was to investigate the fate of vicilin protein from the non-fermented stage up to the dried cocoa beans. Our results showed a remarkable shift in the electrophoretic mobility of vicilin towards higher pI during the onset of fermentation. The pI-shifted subunit was found susceptible to further degradation into a lower-molecular-weight vicilin subunit. The observed pI shift correlated with, but did not depend on protein phosphorylation. Glycosylation of some but not all vicilin subunits occurred at different stages of the fermentation process. Peptides generated from vicilin throughout fermentation were analyzed by UHPLC-ESI-MS/MS revealing an initial increase and subsequent decrease in the diversity of peptides with an increasing degree of fermentation. We furthermore describe the rate of degradation of different vicilin subunits. The detected diversity and dynamics of vicilin peptides will help to define biochemical markers of distinct steps of the fermentation process.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Before getting transformed into chocolate and other cocoa products, the content of cocoa beans undergoes a series of complex metabolic changes during processing steps involving fermentation, drying, roasting, and conching. The development of chocolate flavor requires that cocoa beans are fermented and dried, since during these processes essential flavor precursors are obtained, which are fully expressed later during the roasting and conching processes. These precursors form new types of molecules with other cocoa compounds as result of Maillard reactions and other biochemical processes (Lopez & Dimick, 2008). Cocoa fermentation starts with the prevalence of a succession of different microorganisms, which metabolize sugars from the mucilaginous pulp, produce ethanol, and subsequently release acetic and lactic acid (Ostovar & Keeney, 1973; Schwan & Wheals, 2004). The resulting acidification causes the death of the bean's embryo and induces several enzymatic processes resulting in the generation of essential chocolate-specific aroma precursors (Amin, Jinap, & Jamilah, 1997; Biehl, Brunner, Passern, Quesnel, & Adomako, 1985; Buyukpamukcu et al., 2001; de Brito et al., 2001; Voigt & Lieberei, 2014; Ziegleder & Biehl, 1988).

Cocoa storage proteins undergo the most intense modifications during fermentation, where microbiological fermentation products-triggered enzymatic reactions lead to their extensive breakdown yielding peptides and amino acids, which are important flavor precursors (Zak & Keeney, 1976). Proteins account for 10–15% of the dry weight of cocoa seeds, the second most abundant constituent after cocoa fat (Marseglia et al., 2014). Voigt and coworkers had reported that the total cocoa bean protein content consisted of 52% albumin and 43% vicilin-like globulin (Voigt et al., 1993). In contrast, Lerceteau et al. (1999), employing two-dimensional (2D) gel electrophoresis and silver staining, determined that vicilin-like globulin and the albumin storage protein represent 23 and 14%, respectively, of the total seed protein in addition to other abundant unknown polypeptides (Lerceteau et al., 1999). Albumin was identified as a 21-kDa protein, and its primary structure together with its trypsin inhibitory properties were previously reported (Kochhar, Gartenmann, & Juillerat, 2000). The vicilin-like globulin was originally described as a 66-kDa storage protein (Spencer & Hodge, 1992), giving rise to three distinct break-down units with molecular masses of 47 kDa, 31 kDa, and 15 kDa (Voigt et al., 1993). Recently, it was demonstrated that cocoa-specific aroma precursors can be obtained by in vitro proteolysis of vicilin-like globulin storage protein

* Corresponding author at: Department of Life Sciences and Chemistry, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany.
E-mail address: m.ullrich@jacobs-university.de (M.S. Ullrich).

http://dx.doi.org/10.1016/j.foodres.2016.10.033
0963-9969/© 2016 Elsevier Ltd. All rights reserved.

via addition of an aspartic endoprotease and a carboxypeptidase (Voigt, Janek, Textoris-Taube, Niewienda, & Wostemeyer, 2016). Fermentation of cocoa beans is fundamental for the activation of these two enzyme classes by microbial metabolites such as acetic acid (Schwan & Wheals, 2004). Interestingly, up to date there is no evidence that microbial enzymes might penetrate the bean's shell and create any flavor compounds (Schwan & Wheals, 2004). Furthermore, it had been shown that the formation of cocoa aroma precursors was dependent on both, the cleavage specificities of the involved proteases as well as the structure of the globular storage protein (Bytof, Biehl, Heinrichs, & Voigt, 1995; Rashidah, Jinap, Nazamid, & Jamilah, 2007; Voigt, Voigt, Heinrichs, Wrann, & Biehl, 1994a; Voigt, Wrann, Heinrichs, & Biehl, 1994b).

Although degradation of cocoa storage proteins into oligopeptides and amino acids is well-documented, the actual fate of vicilin as a main source of these compounds starting from the non-fermented cocoa beans towards different stages of fermentation, and finally towards the dried bean, remains uncharacterized. Consequently, the aim of the current work was to provide a comparative and detailed analysis of vicilin protein with respect to the molecular mass(es), isoelectric point(s) and post-translational modifications of its different subunits as well as its oligopeptide degradation products throughout the entire fermentation process.

2. Material and methods

2.1. Chemicals and reagents

Tris-HCl (Pufferan®, ≥ 99.5%), sodium dodecyl sulphate (SDS, ≥ 99.5% electrophoresis grade), HPLC-grade water (Rotisolv®), glycerol (86%), acrylamide (Rotiphorese® Gel 30:37, 5:1), acetonitrile (ACN, Rotisolv® HPLC ultra gradient grade) and 3-[(3-Cholamidopropyl)dimethyl ammonio]-1-propane sulphonate (CHAPS, Pufferan®, 98%) were purchased from Carl Roth (Karlsruhe, Germany). Dithiothreitol (DTT, Biochemica), acetone (100%, Biochemica), TEMED, ammonium persulphate (APS, analytical grade), acetic acid (100%, analytical grade) and methanol (100%, analytical grade) were purchased from Applichem (Darmstadt, Germany). Isopropanol (100%, analytical grade) and bromophenol blue sodium salt (research grade) were purchased from Serva Electrophoresis (Heidelberg, Germany). Urea (≥ 99.5%, BioScience-Grade), thiourea (ACS reagent, ≥ 99%), 2,5-Dihydroxybenzoic acid (>99%, HPLC grade), trifluoroacetic acid (TFA, 99%, ReagentPlus®), α-Cyano-4-hydroxycinnamic acid (**4-HCCA, ≥** 99%, HPLC grade), ammonium **dihydrogen phosphate ($NH_4H_2PO_4$, puriss, ≥** 99%) and formic acid (≥98% for mass spectrometry) were purchased from Sigma-Aldrich Chemie (Steinheim, Germany). Sequencing grade modified trypsin was obtained from Promega (Mannheim, Germany). 100× BioLyte® 3/10 Ampholyte and ReadyPrep™ 2-D Cleanup Kit were purchased from Bio-Rad (Munich, Germany). 2D Quant Kit was purchased from GE Healthcare Biosciences (Munich, Germany). Pierce™ Glycoprotein Staining Kit and protease inhibitor cocktails were purchased from ThermoFisher Scientific (Bremen, Germany). Coomassie® Brilliant blue G-250 (electrophoresis grade) was purchased from Merck (Darmstadt, Germany).

2.2. Cocoa samples

A series of cocoa beans derived from pods of the cocoa hybrid G11UTA402XT413 with different fermentation levels (fresh non-fermented cocoa beans, beans obtained at different fermentation time-points, and fully fermented dried cocoa beans) were harvested in April 2014 in the estate of Champ Semencier de Guessabo, Ivory Coast, and provided by Barry Callebaut, Belgium. The cocoa beans were spontaneously fermented, and the changes in pH and temperature were monitored continuously throughout the fermentation. Cocoa bean samples were collected at 24-h intervals of fermentation until 144 h and finally after traditional drying for 6 days. The samples were then immediately frozen at −20 °C. The bean material of different fermentation levels as well as non-fermented and fully fermented/dried beans was shipped to Bremen on dry ice. Further storage was done at −20 °C.

2.3. Protein extraction from cocoa beans

Mucilaginous pulp from the non-fermented beans was wiped off using Kim Wipes (Kimtech Science, Reigate, UK) and seed coats were removed by peeling them off with a sterile scalpel. The seed coats of fermented beans were removed using forceps. The beans were crushed into fine powder in a Retsch grindomix instrument (Chaotrope, Bremen, Germany). The finely ground powder (250 mg)was added in a 2-mL polypropylene microcentrifuge tube (Eppendorf, Hamburg, Germany) and suspended in 1 mL of protein extraction buffer (100 mM Tris HCl, 1% DTT and 1% SDS; pH adjusted to 8.1; protease inhibitor cocktail according to manufacturer's instruction (Thermo Scientific, Bremen, Germany)). The mixture was vortexed thoroughly for 60 s and then incubated for 60 min with gentle shaking at 4 °C, followed by centrifugation at 16,000 × *g* for 20 min and 4 °C in a bench top 5810R centrifuge (Eppendorf, Hamburg, Germany). The supernatant containing the protein fraction was gently transferred to a fresh cup and stored at −20 °C for further analysis.

2.4. Determination of protein concentration

Quantification of total protein content was performed according to the 2D Quant Kit protocol using 2 mg/mL bovine serum albumin as standard. Since the assay is based on the specific binding of copper ions to proteins, the amount of unbound copper ions measured with a colorimetric reagent at an absorbance at 480 nm is giving an inverse and linear relationship of color intensity and protein concentration within a range of 0 to 50 μg. All measurements were done in triplicates.

2.5. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

The extracted cocoa proteins were separated according to their molecular weight using SDS-PAGE. The protein samples (25 μL) were mixed with 5 μL of 6× sample buffer (Laemmli, 1970) containing bromophenol blue as tracking dye. The mixture was heated at 95 °C for 5 min and loaded on SDS–PAGE gels (83 mm × 65 mm × 1 mm) containing 12.5% (w/v) acrylamide run in the Mini-PROTEAN Tetra cell system (Bio-Rad, Munich, Germany). Electrophoresis was done at 130 V for 90 min. The resulting gel was stained with Coomassie® Blue (45% (v/v) methanol, 10% acetic acid, 2.93×10^{-3} M Coomassie® Brilliant Blue G-250) for 20 min and further treated with a destaining solution (10% (v/v) acetic acid, 5% (v/v) 2-propanol) overnight with gentle shaking.

2.6. Purification of protein samples

Before proceeding to two-dimensional gel electrophoresis, total protein extracts were purified using a Ready Prep™ 2D clean up kit as recommended by the manufacturer. This purification allowed removal of interfering substances such as ionic detergents, lipids, or phenolic compounds. After precipitation and washing in wash buffer provided with the kit, the protein sample was re-suspended in 130 μL of rehydration buffer (2 M thiourea, 6 M urea, 16.2×10^{-3} M CHAPS, 25.9×10^{-3} M DTT) and supplemented with ampholytes according to manufacturer's specification.

2.7. Two-dimensional protein gel electrophoresis

2D gel electrophoresis of cocoa proteins was performed by isoelectric focusing and subsequent SDS–PAGE. For this purpose, 80 μg of protein were applied to immobilized pH gradient (IPG) strips (7 cm, pH 3–10; Bio-Rad) by soaking for 14 h at room-temperature. Isoelectric

focusing was carried out on a Bio-Rad Protean® i12TM IEF Cell (50 V, 70 min; 150 V, 20 min; 300 V, 15 min; gradient to 600 V, 10 min; 600 V, 15 min; gradient to 1500 V, 10 min; 1500 V, 30 min; gradient to 3000 V, 20 min; 3000 V, 210 min; pause on 50 V). Next, IPG strips were equilibrated for 15 min in 6.48×10^{-2} M DTT and 0.216 M iodoacetamide solution dissolved in equilibration buffer (6 M urea, 30% (w/v) glycerol, 69.2×10^{-3} M SDS in 0.05 M Tris-HCl buffer, pH 8.8) at room-temperature. Molecular weight separation was conducted on a Bio-Rad Mini-Protean® Tetra System (50 mV, 10 min; 110 mV further on) via a 12.5% polyacrylamide gel. Molecular weight of the proteins was assessed by their visual mobilization in polyacrylamide gel and the predicted weight of the amino-acid sequence. The resulting gel was stained and destained as indicated above.

2.8. Determination of protein glycosylation

In-gel glycostaining of proteins was performed using the Pierce™ Glycoprotein Staining Kit according to the manufacturer's recommendations. The kit allows protein staining with a periodate-based oxidation reagent that leads to oxidation of cis-diol sugar groups in glycoproteins. The resulting aldehyde groups are detected by formation of Schiff-base bonds with the staining reagent that result in the production of magenta-colored signals.

2.9. In-gel protein digestion

Protein spots of interest were excised from SDS polyacrylamide gels, chopped into small pieces, and washed twice for 15 min in 100 μL of 0.05 M ammonium bicarbonate buffer, containing 50% ACN (v/v). Gel pieces were then dehydrated by the addition of 500 μL acetonitrile and incubated at room temperature (RT) for 10 min. After decanting and short air-drying, samples were supplemented with trypsin digestion buffer as previously established (Shevchenko, Tomas, Havlis, Olsen, & Mann, 2006). The tryptic digest was carried out by incubation of samples at 55 °C for 30 min. The sample was directly used for Matrix Assisted Laser Desorption Ionization-ToF Mass Spectrometry (MALDI-MS) analyses.

2.10. MALDI-MS analyses of proteins

For spectrometric identification of peptide patterns, 1 μL of protein digest solution was mixed with the same volume of 4-HCCA solution (prepared in 85% ACN, 15% H_2O, 0.1% TFA and 0.001 M $NH_4H_2PO_4$), and spotted on a MTB AnchorChip™ target with an anchor diameter of 600 μm (Bruker Daltonics, Bremen, Germany). Spots were left for drying followed by an additional spotting of 1 μL of 2, 5-dihydroxybenzoic acid, drying, and subsequent submitted to an Auto flex II ToF/ToF mass spectrometer (Bruker Daltonics, Bremen, Germany), which was used with standard parameters (acquisition range 500–4000 Da; S/N = 6, in specific cases 3; error range 50 ppm; allowed miss cleavages = 1; potential modifications 'Oxidation (M)'). Peptide masses derived from trypsin auto-digestion were used for calibration (842.50940; 1045.56370; 1713.80840; 1774.89750; 2083.00960; 2211.10400; 2283.18020 Da). Raw data were processed with Flex Analysis, version 3.0 (Bruker Daltonics, Bremen, Germany). Protein identification was carried out using the Mascot search engine (Perkins et al., 1999), using the Biotools software, version 3.1 (Bruker Daltonics, Bremen, Germany). Mass lists were searched against the NCBI database (Sayers et al., 2009). The restricting taxonomy frame for the search was set to 'other green plants' and the Mascot score probability was set for $p < 0.05$. Due to this setting, the significance threshold for a score was set between 70 and 80 (Perkins et al., 1999). Oxidation of methionine ('Oxidation (M)') was selected as variable modification. For global modifications, phosphorylation (ST) was selected. The mass error was set at 50 ppm.

2.11. Protein and peptide sample size fractionation

Vivaspin 2 concentrators (Sartorius, Göttingen, Germany) of 126 mm length, 17 mm width and an active membrane surface area of 1.2 cm^2 were used for separating medium and small-sized peptides from larger proteins. A hydrosart membrane with a molecular weight cut off of 10,000 Da was selected for peptide separation. The protein samples (2 mL) were loaded to the concentrator, which was then centrifuged at a speed of 8000 × *g* for 5 min in a swing bucket rotor centrifuge (Eppendorf, Hamburg, Germany) at 4 °C. The resulting filtrate contained peptides with molecular weights usually <10 kDa and was subjected to vacuum centrifugation until any liquid got evaporated. The resulting powder was subsequently re-suspended and separated using ultra high performance liquid chromatography electrospray ionization mass spectrometry (UHPLC-ESI-MS/MS).

2.12. UHPLC-ESI-MS/MS analysis

The peptide samples were separated by reversed phase HPLC on an Advance Bio Peptide Mapping column (2.1 × 250 mm, i.d. 2.7 μm, Agilent, Munich, Germany) equipped with an Advance Bio Peptide Mapping Fast Guard, size 2.1 × 5 mm guard column (Agilent

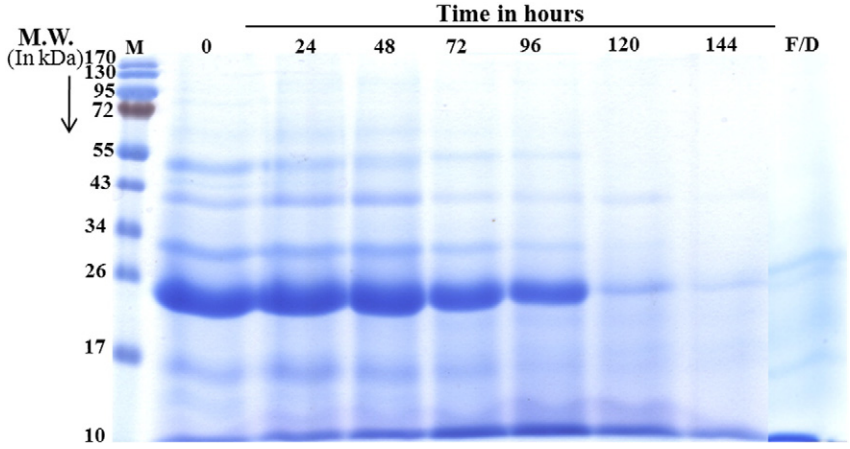

Fig. 1. Cocoa bean protein visualization by SDS-PAGE analysis. Protein fractions of cocoa bean samples at different fermentation time points including non-fermented (0 h) and fully fermented/dried sample (F/D) were run on 12.5% polyacrylamide gels and stained with Coomassie G-250. Standard protein markers were run on each gel for molecular mass determination.

Technologies, CA, USA). A gradient of eluent A (0.005% formic acid in water) and eluent B (100% acetonitrile) from 8% to 90% solvent B within 30 min was applied at a flow rate of 250 μL/min. The sample injection volume was 2 μL. Tandem MS was performed with an Impact™ HD-q-Time of Flight mass spectrometer with a ESI source (Bruker Daltonics, Bremen, Germany) (ESI-MS). The nebulizer pressure was kept at 2 bars with a dry gas flow rate of 8 L/min, and dry gas temperature of 200 °C. The mass spectrometer was calibrated with a 0.1 M sodium formate solution prior to each chromatographic run using an enhanced quadratic calibration mode. All experiments were carried out in positive ion mode over the range of 160–3000 *m/z*. Tandem mass spectra were run in Auto MS^n mode. HPLC grade water was used as blank, which was run after every five samples to check for any kind of contamination from the previous samples. The ESI-MS data obtained were then subjected to the MASCOT search software (Matrix Science, Boston, USA) as recommended previously (Perkins et al., 1999). Parameters were set to search the NCBI database for "Other green plants" with variable modifications of methionine oxidation. The mass tolerance MS was set

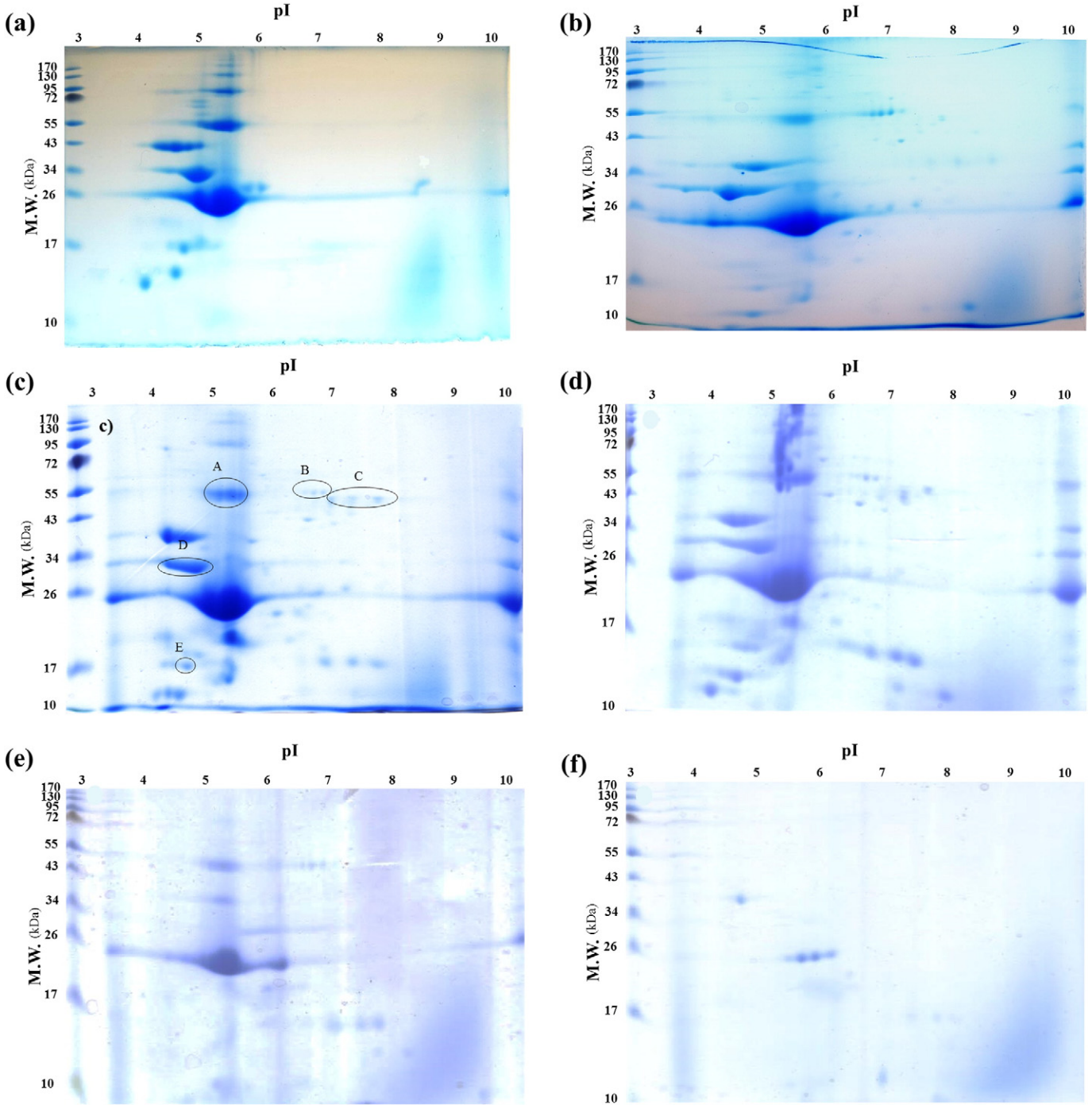

Fig. 2. Cocoa bean protein visualization by two-dimensional gel electrophoresis. Proteins were extracted from cacao bean samples at different stages of fermentation: non-fermented cacao beans (a), fermented beans at 24 h (b), 48 h (c), 72 h (d), 96 h (d), 120 h (e), and 144 h (f). 80 μg of proteins were loaded per gel except for the 144-h sample, where protein quantity was too low for quantification. Standard protein markers were run on each gel for molecular mass determination. In panel (c) different vicilin subunits were marked as: A, non-pI-shifted 47 kDa; B, pI-shifted 47 kDa; C, 44.3 kDa; D, 31 kDa; and E, 15 kDa.

to 100 ppm and MS/MS tolerance to 0.1 Da. Positive peptide hits were accepted if experimental mass errors were below 2 ppm and MS^2 spectra showed identity.

3. Results and discussion

3.1. Comparative SDS-PAGE analyses of cocoa bean samples

Cocoa bean materials from various stages of the fermentation process were subjected to protein extraction and SDS-PAGE analysis (Fig. 1). Extracts of non-fermented cocoa beans gave five major protein bands at 15, 21, 31, 43 and 47 kDa. The polypeptides with the molecular weights of 15, 31, and 47 kDa were consistent with previously reported molecular weights for the vicilin-like globulin proteins (Voigt et al., 1993). MALDI-MS analysis of tryptic digests, which will be discussed in detail below, of non-fermented cocoa beans, also confirmed this assumption. The most prominent protein band with a molecular weight of 21-kDa, likely to be albumin as previously reported in the literature (Amin et al., 1997), was indeed identified as such by tryptic digest and MALDI-MS analysis. One of the remaining major protein bands present in the non-fermented bean sample was of 43-kDa and could be identified as Kinase superfamily protein (Genbank accession number: XP_007014428). These results indicated that most of the detectable proteins were associated with the two cocoa storage proteins, albumin and vicilin.

Furthermore, our results showed continued protein degradation as the fermentation proceeded, which was in accordance with the increasing pH across fermentation (Fig. S1). This result was in agreement with previously reported data (Lerceteau et al., 1999). The protein bands at 31 and 47 kDa showed an almost complete degradation at 120 h of fermentation. At this time point, the most remarkable degradation was observed for the 21-kDa albumin protein. The bands representing proteins of 15 and 43 kDa declined continuously until 144 h. However, even after complete drying, the 21-kDa protein band was still visible suggesting that some of the albumin remained intact throughout the fermentation process. This might be due to the over-proportional presence of this particular protein among all other detectable proteins at the beginning of fermentation (Fig. 2a) or due to its low susceptibility towards degradation.

3.2. Protein analysis of fermentation by two-dimensional gel electrophoresis

In order to better understand changes in the molecular weight (M.W.) and the isoelectric point (pI) of cocoa storage proteins during the fermentation process, total proteins were extracted from cocoa bean samples and subjected to 2D gel electrophoretic analysis (Fig. 2). The results showed a decrease in the number as well as the intensity of protein spots during fermentation, which was in accordance with the results obtained from SDS-PAGE analysis. All major protein spots obtained were subjected to proteolytic digestion and MALDI-MS analysis in order to assign them to respective cocoa proteins. Additionally to the above identified cocoa storage proteins, albumin and vicilin, 17 other cocoa proteins could be identified and characterized in terms of their M.W. and pI (Table S1). The most abundant protein spot at 21-kDa and a pI of 5, representing albumin storage protein, did not undergo any M.W. changes and only minor pI changes but declined in intensity throughout the fermentation process suggesting that this storage protein is steadily and homogenously degraded without forming breakdown intermediates larger than approximately 10-kDa. In contrast, a remarkable change was noticed for the initially observed 47-kDa vicilin protein (Fig. 2a). The 24-h fermentation sample showed a clear deviation in the electrophoretic mobility for a portion of the 47-kDa vicilin visible in form of an additional chain of five 47-kDa protein spots spreading over a pI range from 6.5 to 6.9 (Fig. 2b). Thus, our result clearly showed visible variation in the two dimensional protein profile after 24 h of fermentation with respect to the non-fermented sample (Fig. 2a,b), which has not been observed previously (Lerceteau et al., 1999). All vicilin variants were labelled by a letter code in Fig. 2c from A to E. In contrast to the non-shifted vicilin variant (A), the pI-shifted 47-kDa vicilin proteins (B) were found to be more susceptible to degradation, since new 44.3-kDa vicilin subunits (C) were formed at 48 and 72-h of fermentation (Fig. 2c,d). This protein variant was represented by three major protein spots spreading over a pI range of 6.9 to 7.8. Interestingly, there were no such 44.3-kDa subunits migrating at pI of 5 indirectly suggesting that the subunits had been majorly formed from the pI-shifted vicilin variants.

In addition, a 31-kDa subunit of vicilin (D) was observed at a pI of 4.5 to 5 confirming previous results by others (Kratzer et al., 2009). The later subunit did not show any signs of a pI shift. The non-pI-shifted 31-kDa (D) and the 47-kDa (A) vicilin proteins were detectable until 96 h of fermentation and got later on completely degraded at 144 h of fermentation. However, presence of the 15-kDa vicilin subunit (E) was observed throughout the fermentation with a relative decrease in intensity. These data suggested that vicilin undergoes two major types of degradation: on the one hand, a 31-kDa and a 15-kDa subunit are formed without major alterations of the original pI range, at which these subunits are found throughout the fermentation process. On the other hand, the original 47-kDa vicilin undergoes a potential biochemical alteration shifting its pI from 5 to almost 7 followed by a breakdown

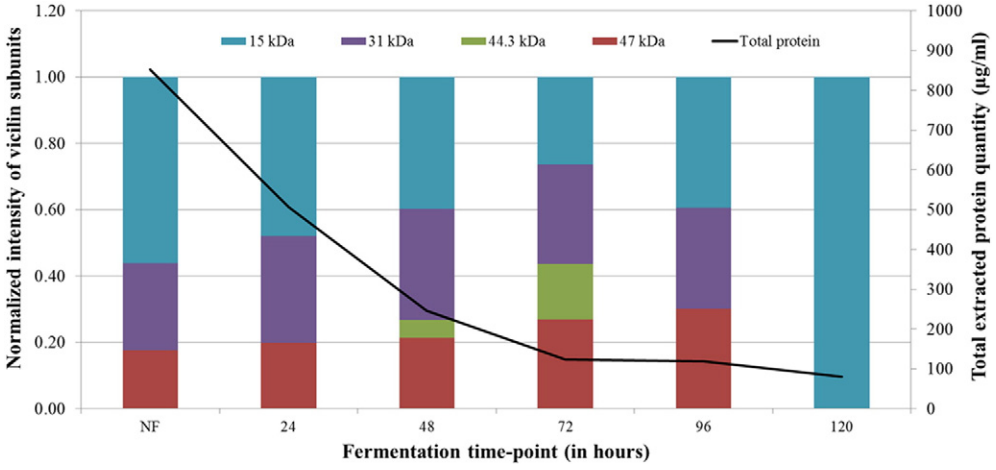

Fig. 3. Normalized intensities of different vicilin subunits during fermentation. Relative intensities of different vicilin subunits (bars) for individual fermentation stages as determined by Image J analysis and normalized to the sum of their collective intensities and relative molecular weights. The line graph represents the total extracted protein quantity at each time-point.

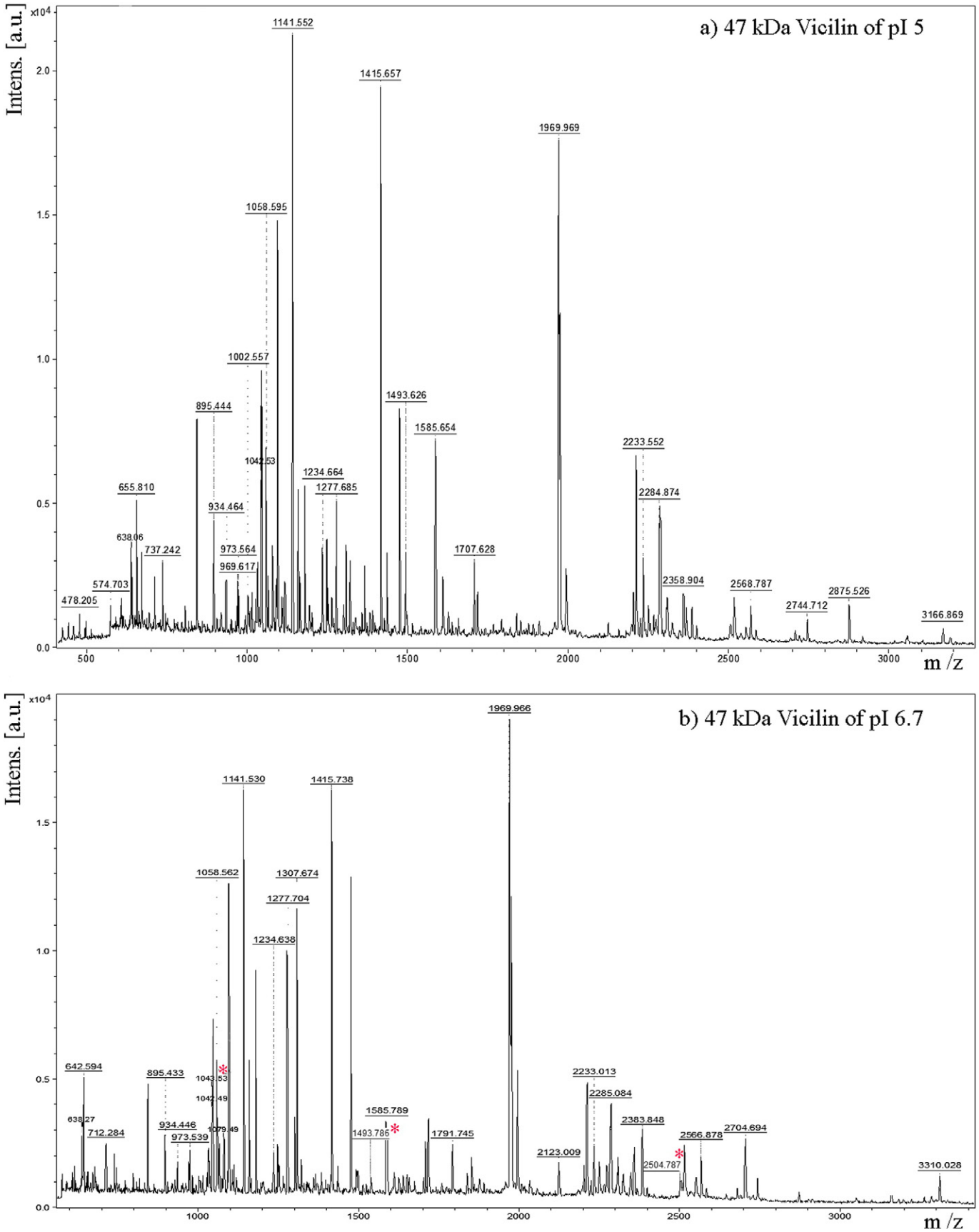

Fig. 4. Comparison of MALDI-MS spectra of the peaks obtained for the 47-kDa vicilin and its pI shifted variant. The spectra represent the prominent *m/z* ratio of the peaks obtained for the 47-kDa vicilin (a) and its pI shifted variant (b). The *m/z* ratio of the peaks marked with an asterisk (*) in the pI shifted variant were identified as being phosphorylated. All labelled peaks denote monoisotopic peptide masses $[M + H]^+$.

into a 44.3-kDa subunit. The formation of 44.3-kDa subunit was a rather unique finding and has not been reported in literature before.

To gain more insight into the temporal dynamics of vicilin degradation, the relative quantification of vicilin and each of its subunits throughout fermentation was conducted from 2D protein gels and plotted with the software Image J (www.imagej.net) (Fig. 3). The extracted total protein content at each time point has been overlaid in this figure as a proxy for the absolute protein degradation. Interestingly, our data suggest that the relative rate of degradation of the 15-kDa subunit during the major protein breakdown period (0 to 72 h) is much faster than that of the other subunits as well as of the 47-kDa vicilin itself. We attribute this to the relatively higher concentration of the 15-kDa subunit at the beginning of the fermentation. This specific degradation slowed down dramatically after 72 h of fermentation. However, the 47-kDa vicilin and its 31-kDa subunit remained relatively stable until 96 h, after which it underwent a massive breakdown (Fig. 3). In combination, these results suggested that during the first 72 h of fermentation, the 15-kDa vicilin subunit was the major target of proteolytic activity while the full size 47-kDa vicilin and its 31-kDa subunit were attacked by proteases only when the fermentation had proceeded after 96 h. The results thus obtained were found to be in agreement with previously reported data (Lerceteau et al., 1999). The 44.3-kDa vicilin subunit(s) became visible only after 48 h of fermentation and lasted until 72 h. Since no 44.3-kDa subunit was detectable at 96 h, one has to conclude a rather fast degradation within 24 h suggesting that its specific rate of degradation is significantly higher than that of all other vicilin subunits. Altogether our data suggested that after 72 h of fermentation, the process of proteolysis became apparent and concerned particularly the higher molecular weight peptides of >15 kDa. Interestingly, the pH profile of the particular fermentation series, showed a change in pH from acidic to close to neutral after 72 h (Fig. S1) suggesting a correlation between the rate of protein degradation and pH change. However, our data did not support any impact of the temperature change on the progress of storage protein proteolysis (Fig. S1).

3.3. Differences between 47-kDa vicilin and its pI shifted variant

In order to better understand the pI shift of the 47-kDa vicilin protein observed within 24 h of fermentation (Fig. 2a,b), MALDI-MS analyses of tryptically digested protein spot samples were carried out. All prominent *m/z* ratio peaks from the MALDI-MS spectra (Fig. 4), were subjected to MASCOT analysis and thus assigned to predicted peptide fragments of the full-size vicilin precursor (Genbank accession number: S22477) as summarized in Tables 1 and 2. Interestingly, all detectable peptide fragments fitted to the corresponding amino acid sequence of the previously reported 566 amino acid-long vicilin precursor from residue 132 onwards to residue 545 suggesting that the N-terminal 131 amino acid residues were either missing from the vicilin sequence at all or had been removed by proteolytic cleavage prior to the fermentation start, i.e. inside the cocoa pod during ripening. Notably, none of the previous reports had ever shown any peptide fragments obtainable from this questionable N-terminus of the vicilin precursor (Kratzer et al., 2009). In total, 20 matching peptides were obtained for the initial 47-kDa vicilin protein spot (Table 1). Except three central regions and the N-terminus, these 20 fragments covered the majority of the vicilin protein sequence. The majority of these peptide fragments could also be detected in the MALDI-MS analysis of the tryptic digest of the pI-shifted 47-kDa vicilin subunits regardless of their specific electrophoretic location (Table 2). However, four of the peptide fragments obtained for the pI-shifted vicilin were found to be phosphorylated (Table 2). Interestingly, three of these peptides were found to be identical (although non-phosphorylated) in the non-pI-shifted 47-kDa vicilin while the fourth of these phosphorylated fragments was found within a bigger peptide fragment present in the non-pI shifted variant of vicilin. These results prompted the initial assumption that phosphorylation of vicilin might lead to a pI shift as suggested for other proteins (Zhu, Zhao, Lubman, Miller, & Barder, 2005). However, this assumption remained debatable since the majority of previously reported phosphorylation-mediated pI shifts lead rather to an acidification but not a neutralization of the protein (Khaldi & Shields, 2011) as observed herein.

Since the 44.3-kDa vicilin subunit(s) detected at 48 and 72 h of fermentation (Fig. 2c,d) were found at a rather neutral pI similar to the pI-shifted 47-kDa vicilin variant, it was reasonable to assume that some or all of the phosphorylated peptide fragments found in the pI-shifted 47-kDa protein (Table 2) might also be found in the smaller subunit but rather not in the non-pI shifted 31-kDa vicilin subunit. In order to test this, tryptically digested fragments of the 44.3-kDa subunits located at a pI range of 6.9 to 7.8 as well as those derived from the 31-kDa and the 15-kDa vicilin subunits were subjected to MALDI-MS analysis (Tables 3-5). Sixteen matches were obtained for the 44.3-kDa subunit,

Table 1
Peptide fragments detected by MALDI-TOF analyses in the tryptic digest of the non-pI-shifted 47 kDa cocoa vicilin subunit.

Sequence[a]	Position[b]	Average mass (Mr)		
		Experimental	Calculated	Error (Da)
R.SEEEEGQQR.N	132–140	1090.4496	1090.4527	-0.0031
R.NNPYYFPK.R	141–148	1041.4945	1041.4920	0.0025
R.SFQTR.F	151–155	637.3164	637.3184	−0.0020
R.FRDEEGNFK.I	156–164	1140.5826	1140.5200	0.0626
R.FAENSPPLK.G	169–177	1001.5492	1001.5182	0.031
K.GTITVTHENK.E	213–223	1245.6148	1245.6354	−0.0206
K.ESYNVQR.G	224–230	894.4365	894.4195	0.017
R.GTVVSVPAGSTVYVYVVSQDNQEK.L	231–252	2264.2636	2264.1405	0.1231
K.LTIAVLALPVNSPGK.Y	253–267	1492.9615	1492.9098	0.0518
R.EKLEEILEEQR.G	300–310	1414.7946	1414.7303	0.0643
R.QQGQQGMFR.K	315–323	1078.5375	1078.4978	0.0397
R.QQGQQGMFRK.A	315–324	1207.7073	1207.6001	0.1072
K.AKPEQIR.A	325–331	840.4982	840.4817	0.0165
R.AISQQATSPR.H	332–341	1057.5778	1057.5516	0.0262
R.QSQGSQSGRQDR.R	426–437	1332.6190	1332.5457	0.0733
R.EQEEESEEETFGEFQQVK.A	439–456	2200.8707	2200.9284	−0.0577
K.APLSPGDVFVAPAGHAVTFFASK.D	457–479	2286.2651	2286.1918	0.0734
K.DQPLNAVAFGLNAQNNQR.I	480–497	1968.8489	1968.9766	−0.1277
R.QMDSEAKELSFGVPSK.L	509–524	1752.9314	1752.8473	0.0841
K.LVDNIFNNPDESYFMSFSQQR.Q	525–545	2567.1798	2567.1507	0.0291

[a] The amino acid residues located upstream and downstream of the detected peptides in the total cocoa vicilin precursor sequence (Genbank accession number S22477) are separated by points.
[b] Position in the amino acid sequence of the cocoa vicilin precursor (Spencer & Hodge, 1992).

Table 2
Peptide fragments detected by MALDI-TOF analyses in the tryptic digest of the pI-shifted 47 kDa cocoa vicilin subunit.

Sequence[a]	Position[b]	Average mass (Mr)		
		Experimental	Calculated	Error (Da)
R.SEEEEGQQR.N	132–140	1090.4496	1090.4527	−0.0031
R.NNPYYFPK.R	141–148	1041.4945	1041.4920	0.0025
R.SFQTR.F	151–155	637.3164	637.3184	−0.0020
R.FRDEEGNFK.I	156–164	1140.5826	1140.5200	0.0626
R.FAENSPPLK.G	169–177	1001.5492	1001.5182	0.031
K.GTITVTHENKESYNVQR.G	213–230	2121.9041	2122.0433	−0.1402
K.ESYNVQR.G	224–230	894.4365	894.4195	0.017
R.G**T**VV**S**VPAG**S**TVYVVSQDNQEK.L[c]	231-252	2503.9570	2503.0322	0.9248
K.LTIAVLALPVNSPGK.Y	253–267	1492.9615	1492.9098	0.0518
R.EKLEEILEEQR.G	300–310	1414.7946	1414.7303	0.0643
K.LEEILEEQR.G	302–310	1157.5673	1157.5928	−0.0255
R.QQGQQGMFR.K	315–323	1078.5375	1078.4978	0.0397
R.KAKPEQIR.A	324–331	968.5418	968.5767	−0.0349
R.AISQQATSPR.H	332–341	1057.5778	1057.5516	0.0262
R.Q**S**QG**S**QSGRQDR.R[c]	426-437	1492.6190	1492.5457	0.0733
R.EQEEE**S**EEETFGEFQQVK.A[c]	439-456	2280.8707	2280.9284	−0.0577
K.DQPLNAVAFGLNAQNNQR.I	480–497	1968.8489	1968.9766	−0.1277
K.EL**S**FGVPSK.L[c]	516-524	1042.5042	1042.4736	0.0306
K.LVDNIFNNPDESYFMSFSQQR.Q	525–545	2567.1798	2567.1507	0.0291

[a] The amino acid residues located upstream and downstream of the detected peptides in the total cocoa vicilin precursor sequence (Genbank accession number S22477) are separated by points.
[b] Position in the amino acid sequence of the cocoa vicilin precursor (McHenry & Fritz, 1992; Spencer & Hodge, 1992).
[c] Phosphorylated peptide. Amino-acid marked in bold was identified as phosphorylated.

which was localized between amino acid positions 131 and 524 of the vicilin precursor sequence (Table 3) while 13 matches could be found for the 31-kDa vicilin subunit localizing between amino acid positions 131 and 425 of the vicilin precursor sequence (Table 4). For the 15-kDa vicilin subunit, 8 fragments located between amino acid positions 314 to 434 were detected (Table 5). Our data suggested that the 44.3-kDa subunit might lack the 21 amino acid-long 2.3-kDa C-terminus of the 47-kDa vicilin variant (residues 525 to 545). Interestingly, none of the peptides obtained by tryptic digestion of the 44.3-kDa vicilin subunit were found to be phosphorylated (Table 3) suggesting that the previously observed phosphorylation of four fragments in the pI-shifted 47-kDa vicilin might not be the biochemical cause of the pI shift of the later protein variant. The two smaller subunits of vicilin located in the central part of the vicilin precursor sequence and thus were found to be in agreement with previous reports (Kratzer et al., 2009). Neither of these smaller subunits harbored any phosphorylated peptide fragments suggesting that phosphorylation is restricted to only one intermediate form of vicilin, the pI-shifted 47-kDa variant.

3.4. Protein glycosylation analysis

Total cocoa bean proteins from samples derived from time points 0, 24, and 48-h were extracted and subjected to 2D gel electrophoresis followed by glycostaining and Coomassie G-250 staining, respectively, in order to identify any potential protein glycosylation during the onset of fermentation (Fig. 5). Our results clearly revealed that several protein spots, i.e. the 31-kDa vicilin subunit throughout the tested process, showed a glycosylation signal. However, neither albumin nor any of the 47-kDa vicilin variants showed glycosylation thus ruling out any involvement of glycosylation in the pI shift of this protein. In

Table 3
Peptide fragments detected by MALDI-TOF analyses in the tryptic digest of the 44.3 kDa cocoa vicilin subunit.

Sequence[a]	Position[b]	Average mass (Mr)		
		Experimental	Calculated	Error (Da)
R.SEEEEGQQR.N	132–140	1090.4496	1090.4527	−0.0031
R.NNPYFPK.R	141–148	1041.4808	1041.4920	−0.0112
R.SFQTR.F	151–155	637.2972	637.3184	−0.0112
R.FRDEEGNFK.I	156–164	1140.5200	1140.5200	0
R.FAENSPPLK.G	169–177	1001.5021	1001.5182	−0.0161
K.GINDYR.L	178–183	736.3378	736.3504	−0.0126
K.ESYNVQR.G	224–230	894.4215	894.4195	0.002
R.EKLEEILEEQR.G	300–310	1414.7033	1414.7303	−0.027
K.LEEILEEQR.G	302–310	1157.5958	1157.5928	0.003
K.RQQGQQGMF.R	314–322	1234.5874	1234.5989	−0.0115
R.QQGQQGMFR.K	315–323	1078.4978	1078.4978	0
R.KAKPEQIR.A	324–331	968.5783	968.5767	0.0016
K.AKPEQIR.A	325–331	840.4784	840.4817	−0.0033
R.HRGGER.L	342–347	710.3456	710.3572	−0.0016
K.LNQGAIFVPHYNSK.A	389–402	1586.7692	1586.8205	−0.0513
R.QSQGSQSGR.Q	426–434	933.4268	933.4264	0.0004
R.QDRREQEEESEEETFGEFQQVK.A	435–456	2757.2236	2757.2234	0.0002
K.DQPLNAVAFGLNAQNNQR.I	480–497	1968.8489	1968.9766	0.0984
K.ELSFGVPSK.L	516–524	962.5042	962.4736	0.0306

[a] The amino acid residues located upstream and downstream of the detected peptides in the total cocoa vicilin precursor sequence (Genbank accession number S22477) are separated by points.
[b] Position in the amino acid sequence of the cocoa vicilin precursor (McHenry & Fritz, 1992; Spencer & Hodge, 1992).

Table 4
Peptide fragments detected by MALDI-TOF analyses in the tryptic digest of the 31 kDa cocoa vicilin subunit.

Sequence[a]	Position[b]	Average mass (Mr)		
		Experimental	Calculated	Error (Da)
R.SEEEEGQQR.N	132–140	1090.4496	1090.4527	−0.0031
R.NNPYYFPK.R	141–148	1041.4945	1041.4920	0.0025
R.SFQTR.F	151–155	637.3164	637.3184	−0.002
R.FRDEEGNFK.I	156–164	1140.5826	1140.5200	0.0626
R.FAENSPPLK.G	169–177	1001.5492	1001.5182	0.031
K.GINDYR.L	178–183	736.3520	736.3504	0.0016
K.GTITFVTHENKESYNVQR.G	213–230	2121.9041	2122.0433	−0.1402
K.ESYNVQR.G	224–230	894.4365	894.4195	0.017
R.EKLEEILEEQR.G	300–310	1414.7946	1414.7303	0.0643
K.RQQGQQGMFR.K	314–323	1078.5375	1078.4978	0.0397
K.AKPEQIR.A	325–331	840.4982	840.4817	0.0165
K.LNQGAIFVPHYNSK.A	389–402	1587.8905	1587.8278	0.0627
K.ATFVVFVTDGYGYAQMACPHLSR.Q	403–425	2590.3049	2590.2218	0.0831

[a] The amino acid residues located upstream and downstream of the detected peptides in the total cocoa vicilin precursor sequence (Genbank accession number S22477) are separated by points.
[b] Position in the amino acid sequence of the cocoa vicilin precursor (McHenry & Fritz, 1992; Spencer & Hodge, 1992).

contrast and interestingly, several of the spots representing the smaller 44.3-kDa vicilin subunits (pI of 6.9 to 7.8) appearing only at 48 and 72 h of fermentation clearly exhibited a glycosylation signal at 48 h of fermentation (Fig. 5e,f). This result indicated a glycosylation of vicilin only under two, apparently independent circumstances: either when a (stable) 31-kDa subunit is formed or when the larger 47-kDa variant got phosphorylated and subsequently C-terminally truncated in the course of fermentation. In support of this, the MALDI-MS analysis of the non-glycosylated 47-kDa vicilin harbored peaks at *m/z* 2564.2636, 1492.9615, and 1057.5778 (Table 1), which were attributed to amino acid positions 231–252, 253–267, and 332–341, respectively, and which were not detectable in the spectra of the tryptic digests of the glycosylated 44.3-kDa (Table 3) and 31-kDa (Table 4) vicilin subunits. The respective data are schematically summarized in Fig. 6. Noteworthy, the MALDI-MS data of the non-glycosylated 15-kDa vicilin subunit showed a peak at *m/z* 1057.5778 corresponding to amino acid residues 332–341 (Table 5) suggesting that this fragment might carry a glycosylation site just as the two other fragments mentioned above. This supposition was corroborated using the YinOYang algorithm (http://www.cbs.dtu.dk/service/YinOYang), which produces neural network-based predictions for O-linked β-*N*-Acetyl glucosamine (O-GLcNAc) in eukaryotic protein sequences (Blom, Sicheritz-Ponten, Gupta, Gammeltoft, & Brunak, 2004). This prediction revealed possible O-GLcNAc modifications at positions 193 (Thr), 235 (Ser), 338 (Thr) and 474 (Thr) of the vicilin precursor sequence (Fig. 6). Remarkably, the tryptic digest peptide patterns of glycosylated 31-kDa and 44.3-kDa vicilin subunits lacked peptides containing three out of these four predicted positions (Tables 1, 3, and 4) while three and two of these peptide fragments, respectively, could be obtained in samples of the non-glycosylated 47-kDa vicilin variants (Table 1-2). The possible candidate peptides for O-GLcNAc modification were namely R.GTVVSVPAGSTVYVVSQDNQEK.L, R.AISQQATSPR.H, and K.APLSPGDVFVAPAGHAVTFFASK.D. Confirming this, the peptide with the sequence R.AISQQATSPR.H could be detected in the tryptic digest of the non-glycosylated 15-kDa vicilin subunit (Table 5). The only potential glycosylation site not fitting the above trend was the Thr residue at position 193, for which no peptide fragment could be obtained from any of the analyzed protein spots. Fig. 6 also indicates the potential sites of protein phosphorylation. The observed glycosylation pattern in combination with the peptide analysis suggested that the full-size 47-kDa vicilin undergoes phosphorylation but no glycosylation before further breakdown. Subunits of 44.3 kDa and 31 kDa but not the 15-kDa subunit get glycosylated but show no sign of phosphorylation. The glycosylations of the 44.3-kDa and the 31-kDa subunits, albeit at the same position, must occur at different time points and are possibly not related to each other. Notably, the 15-kDa and the 31-kDa subunits must be rather derived from the full-size vicilin at the beginning of the fermentation process or even prior to it but cannot come from the 44.3-kDa subunit. Whether or not phosphorylation and glycosylation depend on each other, remains to be analyzed in future studies. Our data are thus complementing other researchers' previous findings (Kratzer et al., 2009; Lerceteau et al., 1999; Voigt et al., 1993) and provide a second-to-none resolution for the processes going on with the storage protein vicilin during fermentation. The results obtained for the amino acid sequence localization of the 31-kDa and the 15-kDa subunit on the original vicilin precursor sequence was in agreement with a previous study (Kratzer et al., 2009). However, the localization of the 47-kDa vicilin based on the peptide fragments differed from the previously reported fragmentation pattern (Kratzer et al., 2009).

Table 5
Peptide fragments detected by MALDI-TOF analyses in the tryptic digest of the 15 kDa cocoa vicilin subunit.

Sequence[a]	Position[b]	Average mass (Mr)		
		Experimental	Calculated	Error (Da)
K.RQQGQQGMF.R	314–322	1234.5874	1234.5989	−0.0115
R.QQGQQGMFR.K	315–323	1078.4978	1078.4978	0
R.KAKPEQIR.A	324–331	968.5783	968.5767	0.0016
K.AKPEQIR.A	325–331	840.4784	840.4817	−0.0033
R.AISQQATSPR.H	332–341	1057.5538	1057.5516	0.0022
R.HRGGER.L	342–347	710.3456	710.3572	−0.0016
K.LNQGAIFVPHYNSK.A	389–402	1586.7692	1586.8205	−0.0513
R.QSQGSQSGR.Q	426–434	933.4268	933.4264	0.0004

[a] The amino acid residues located upstream and downstream of the detected peptides in the total cocoa vicilin precursor sequence (Genbank accession number S22477) are separated by points.
[b] Position in the amino acid sequence of the cocoa vicilin precursor (McHenry & Fritz, 1992; Spencer & Hodge, 1992).

3.5. Analysis of vicilin breakdown peptide pool during fermentation

Cocoa bean samples at different stages of fermentation were subjected to free peptide analysis via UHPLC-ESI-MS/MS in order to complement the herein observed degradation profile of vicilin. In each of the samples and along with numerous other free peptides, various vicilin-originating oligopeptides could be detected and characterized on the basis of the high resolution determination of *m/z* ratio and their corresponding fragmentation patterns. The free vicilin-based peptides obtained from the samples at different fermentation stage are represented schematically in Fig. 7. The ESI-MS analysis indicated that the sequences of free peptides were localized to distinct zones spread throughout the entire C-terminal vicilin sequence except the N-terminal 131 amino acid residues of vicilin. This result was in agreement

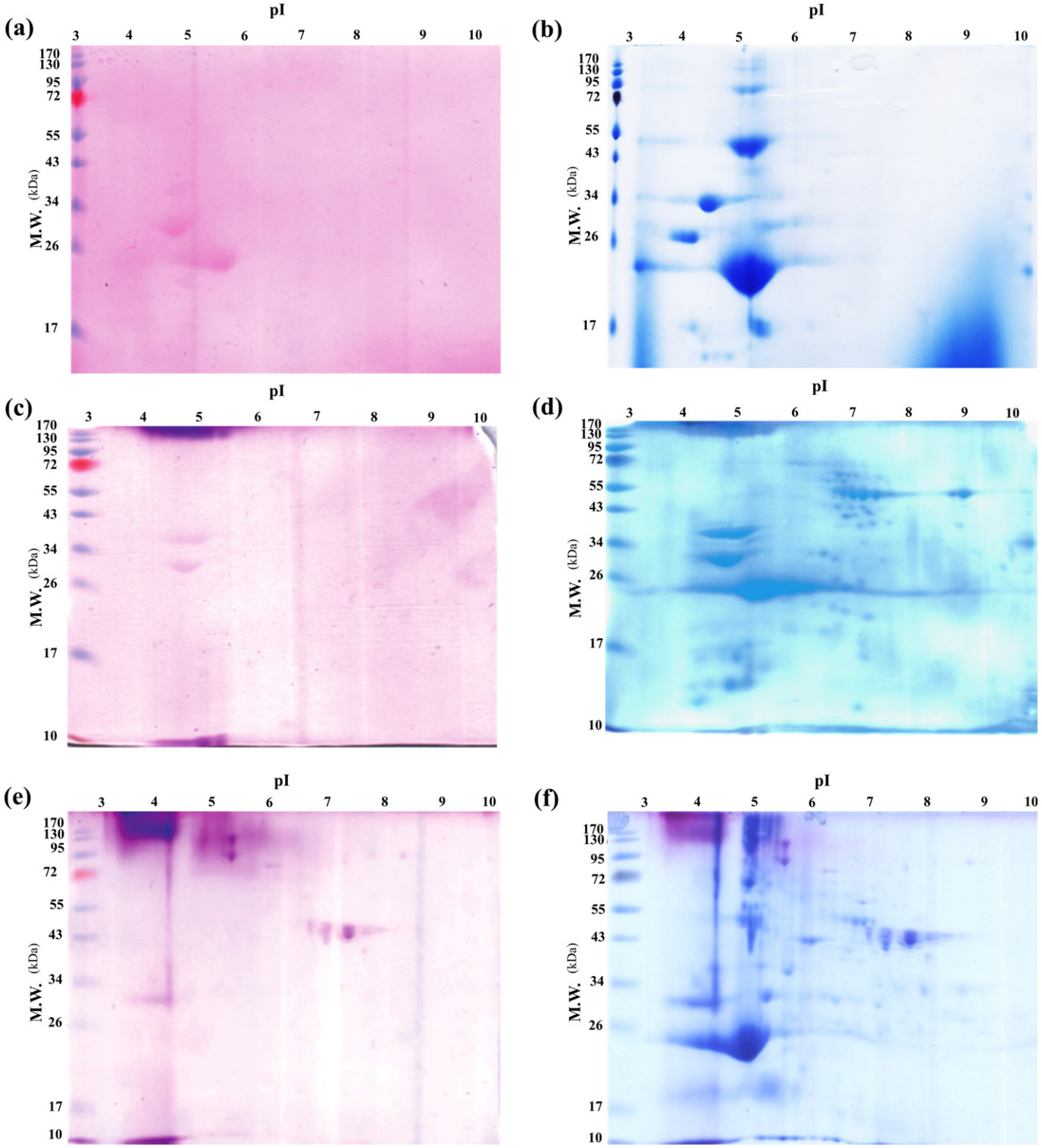

Fig. 5. Visualization of protein glycosylation using 2D gel electrophoresis. Proteins extracted from cacao bean samples from different fermentation stages were electrophoretically separated and stained with either Glycoprotein staining solution (left panels) or Coomassie Blue G250 (right panels). (a, b) non-fermented cocoa beans; (c, d) 24-h fermented cocoa beans; and (e, f) 48-h fermented cocoa beans. 120 μg of protein was loaded per gel. Standard protein markers were run on each gel for molecular mass determination.

with results of a recent study by others (Marseglia et al., 2014) and furthermore suggested that the presumed N-terminal region of the vicilin precursor did not undergo any degradation even before the start of fermentation. It had been previously suggested that this could be due to some sort of N-terminal protection (Kratzer et al., 2009). However, since this N-terminus is likewise not found in any vicilin protein variants as demonstrated above, its presence in the cocoa bean proteome remains obscure. Interestingly and albeit several attempts, no previous study had demonstrated its presence by mass spectrometry or any other biochemical test (Marseglia et al., 2014; Voigt et al., 2016). It can therefore not been excluded that the previous annotation of the N-terminus of the vicilin precursor protein (Spencer & Hodge, 1992) was inaccurate and deserves revision. Our results suggested that native proteolytic degradation of vicilin started within the first 24 h of

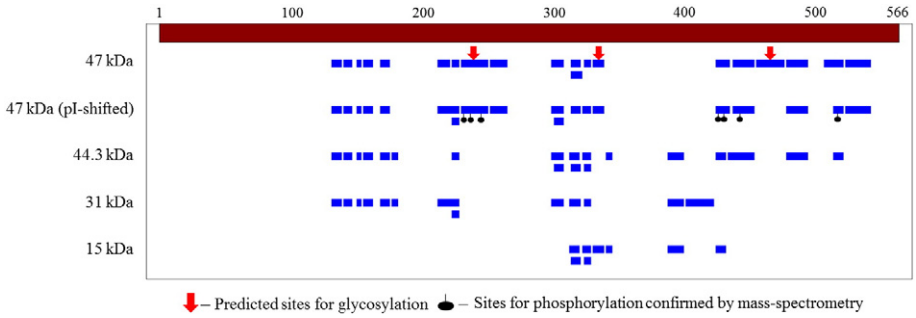

Fig. 6. Localisation of different vicilin subunits within the cocoa vicilin precursor sequence. Peptide fragments (bars) were generated by tryptic digest and identified by MALDI-TOF analysis and analyzed by MASCOT search. Size and position of each bar represents the length and position of each peptide within the vicilin precursor protein sequence (long red bar on top). No peptide hits could be found for the amino acid sequence 1–131. Positions for predicted post-translational modifications are marked with red arrows for glycosylation sites and black circles for phosphorylation sites as confirmed by mass spectrometry.

fermentation, as 65 different vicilin oligopeptides could be detected in the 24-h fermented cocoa bean sample while none were found in the non-fermented sample. Extensive generation of peptides was observed until 72 h of fermentation, after which the number of peptides decreased steadily. No free peptides were found in the last two fermentation time-points at 144 and 168 h as well as in the dried samples. A possible explanation for the gradual decrease in the number of vicilin oligopeptides could be the further breakdown of oligopeptides into free amino-acids or di/tri-peptides. The latter type of peptides could not be detected by the herein used ESI MS analysis technique.

Detailed analysis of free peptide distribution suggested that there were regions within the vicilin sequence, from which no oligopeptides could be detected in any of the samples. These regions were localized at amino acid residues 285–316, 324–354, 418–446, 500–520, and 540–566 (Fig. 7). Absence of oligopeptides derived from certain protein regions is a phenomenon also found in other fermented products, e.g., in casein, for which only peptides from specific sequence regions are usually observed in cheese (Sforza et al., 2012). In the latter case, lower accessibility of certain parts of proteins was demonstrated to hinder proteolytic cleavage. Alternatively, such regions might become insoluble upon initial cleavage and then aggregate precluding any further cleavage or technical detection (Sforza et al., 2012). Finally, some of those regions might be too large to be detected with the herein used technique (Strupat, 2005). In contrast, we identified particularly peptide-rich regions within the vicilin sequence localized between positions 131–298, 350–417, 453–499, and 518–545. The first region was richest in peptides at 48 and 72-h of fermentation suggesting a high stability of the respective breakdown products. Amino acid sequence region 131–298 yielded 26 different peptides, from which one with the sequence EANPNTFILPHHCDAEAIY (position number 188–206) was further cleaved into five lower-molecular weight peptides (Fig. 7). All five smaller peptides from this individual proteolysis had a common N-terminus but different C-terminal amino acids indicating carboxypeptidase enzyme activity (Marseglia et al., 2014). Another peptide was generated from amino acids 476–485, namely FASKDQPLNA, which was further cleaved into four lower molecular weight peptides

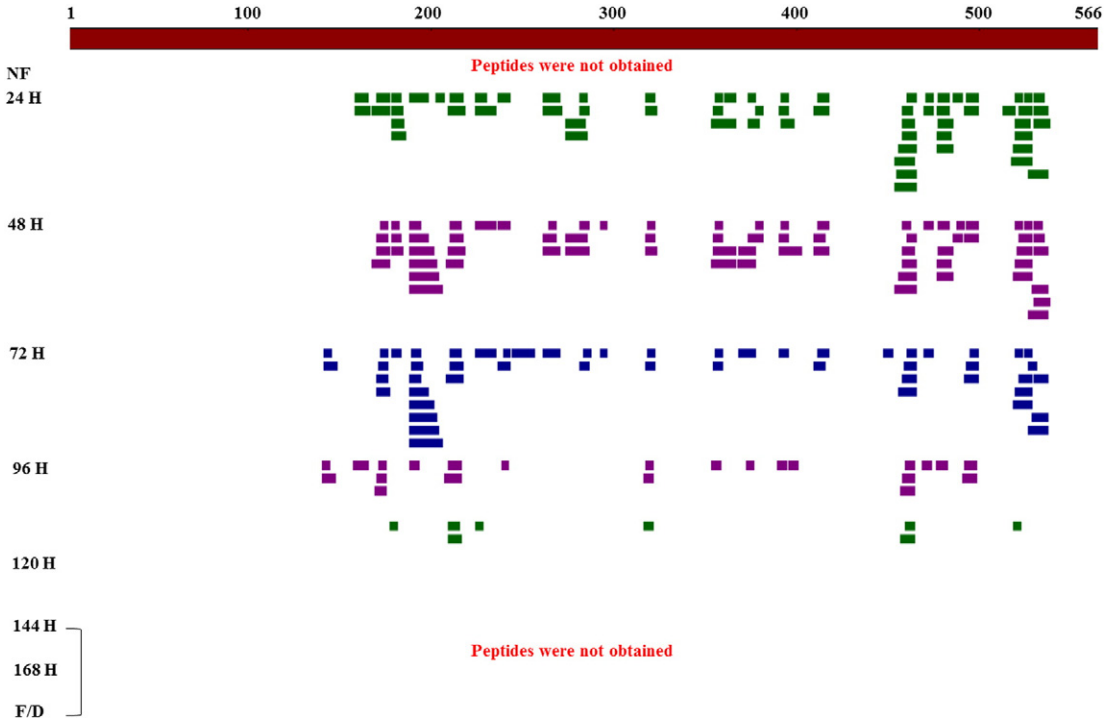

Fig. 7. Peptide repertoire of vicilin protein degraded during fermentation and resolved by UHPLC-ESI-MS/MS analysis. Peptide fractions extracted from cacao bean samples were separated by LC-MS and analyzed by MASCOT search. Size and position of each bar represents the length and position of each peptide within the vicilin precursor protein sequence (long red bar on top). No peptide hits could be found for the amino acid sequence 1–131. NF, non-fermented bean sample; samples at different time-points of fermentation; and F/D, fully fermented and dried bean samples.

all having a common N-terminus but different C-termini supporting the assumption that carboxypeptidase(s) are active.

In addition, the results suggested endoprotease activity at the vicilin sequence position 518–545, where neither common N-terminal nor C-terminal amino acids could be identified for any of the 12 peptides generated at 24 h and 48 h of fermentation. At 72 h still 9 different peptides without common termini could be found. In this context, it is interesting to note that after 72 h of fermentation literally no peptide were obtained from amino acid region 518–545 suggesting the complete breakdown of the generated peptides into di/tri-peptides or free amino acids.

Interestingly, two of the peptides, APLSPGDVF and SPGDVF, localized at amino acid sequences 457–465 and 460–465, respectively, were found to be formed early on during the fermentation and were present until 120 h of fermentation. These two peptides have been reported previously and were suggested as potential flavor precursors (Buyukpamukcu et al., 2001). Peptides APLSPGDVF and SPGDVF localized at a region of vicilin common to both, the 47-kDa and the 44.3-kDa variants. Since this C-terminal region is not part of the smaller vicilin subunits of 15 and 31 kDa, our results suggested that the larger two vicilin variants were extensively cleaved during cocoa bean fermentation, and that they might be the major source of oligopeptide precursors leading to cocoa aroma. In turn, this could imply that the smaller subunits might undergo a different type of degradation, which actually finishes earlier and which might thus contribute to the pool of free amino acids or di- and tri-peptides as potential conjugative moieties for aroma compounds such as Amadori or Steller reaction products (Granvogl, Bugan, & Schieberle, 2006).

Amino acid position 545 represented the C-terminal end-point of the 47-kDa subunit, and the flanking region was rich in peptides while no peptide was obtained derived from the amino acid sequence beyond position 545 (Fig. 7). This result suggested that the absolute C-terminus of the 566 amino-acid long vicilin precursor sequence, as published earlier must undergo a swift carboxypeptidase attack leaving no peptide(s) but just free amino acids or di- and tri-peptides behind.

Within the free peptide pool, peptides ESYNVQRGTVVS, VPAGSTVY, and SFGVPSKLVDN localized at amino-acid sequences 224–235, 236–243 and 518–528 respectively, and were found to be phosphorylated in 24, 48, and 72 h fermented samples. The phosphorylation modification was at positions 232 (Thr), 235 (Ser), 240 (Ser), and 518 (Ser), which were corresponding to the positions previously identified as phosphorylation sites by MALDI-MS analyses (Table 2). It is reasonable to assume that the corresponding peptides might have been mainly generated by proteolysis of the pI-shifted 47-kDa vicilin variant since no free phosphorylates peptides were obtained from samples later than 72 h of fermentation.

The free peptide pool did not reveal any information about glycosylation since no information about the type of glycosylation of vicilin is available making calculations of fitting m/z ratios impossible. Since neither MALDI-MS nor ESI-MS analyses yielded quantitative data, future work needs to focus on accurate quantitation of oligopeptide, di- and tri-peptide, as well as free amino acid formation during cocoa bean fermentation in order to support the herein suggested hypothetical reactions.

4. Conclusions

This study showed remarkable shifts in the electrophoretic mobility of the full-size vicilin protein towards a higher pI as the fermentation progressed. The simultaneously observed phosphorylation cannot be the cause of this pI shift since a non-phosphorylated 44.3-kDa vicilin subunit also had an increased pI. Interestingly, the 44.3-kDa vicilin subunit was identified as being glycosylated in a yet-unknown fashion. Similarly, a 31-kDa vicilin subunit was also found to be glycosylated, and speculations on candidate protein regions carrying potential glycosylation sites were made and indirectly consolidated by peptide mapping.

It can be concluded that the relative rate of degradation of different vicilin subunits varied across fermentation suggesting that the breakdown of vicilin is not homogenous. This might indicate the occurrence of differential and possibly sequential proteolytic activities. To the best of our knowledge, this is the first exhaustive investigation on the fate of cocoa's vicilin protein from non-fermented towards the dried bean sample as well as the first report on possible post-translational modifications of vicilin during fermentation.

Results of the free peptide analysis by ESI-MS suggested that cocoa proteins undergo extensive proteolysis during fermentation with a consequent increase in peptide abundance. In this context and interestingly, our study showed the permanent presence of two previously reported potential flavor precursor peptides (Buyukpamukcu et al., 2001) from early stage of fermentation (24 h) until the near-to-complete stage (120 h) allowing the conclusion that these two peptides are highly abundant and easy to detect.

The high-resolution peptide mapping from extracted proteins and the mapping of free peptides showed the suitability of combining both methods to dissect the biochemical processes proceeding during cocoa fermentation. Further work is in progress on a higher number of cocoa bean samples from different origins and different fermentation styles to map the precursor variability potentially linked to the flavor of the final product. This will help to ultimately enhance the quality and predictability of the taste of the cocoa end product.

Acknowledgments

We thank Nina Böttcher and Britta Behrends for excellent technical support in sample preparation. The authors are also grateful to Rohan L. Shah for excellent help with the peptide identification method development.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.foodres.2016.10.033.

References

Amin, I., Jinap, S., & Jamilah, B. (1997). Vicilin-class globulins and their degradation during cocoa fermentation. *Food Chemistry*, *59*(1), 1–5. http://dx.doi.org/10.1016/0308-8146(95)00188-3.

Biehl, B., Brunner, E., Passern, D., Quesnel, V. C., & Adomako, D. (1985). Acidification, proteolysis and flavor potential in fermenting cocoa beans. *Journal of the Science of Food and Agriculture*, *36*(7), 583–598. http://dx.doi.org/10.1002/jsfa.2740360710.

Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S., & Brunak, S. (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics*, *4*(6), 1633–1649. http://dx.doi.org/10.1002/pmic.200300771.

Buyukpamukcu, E., Goodall, D. M., Hansen, C. E., Keely, B. J., Kochhar, S., & Wille, H. (2001). Characterization of peptides formed during fermentation of cocoa bean. *Journal of Agricultural and Food Chemistry*, *49*(12), 5822–5827. http://dx.doi.org/10.1021/jf0104127.

Bytof, G., Biehl, B., Heinrichs, H., & Voigt, J. (1995). Specificity and stability of the carboxypeptidase activity in ripe, ungerminated seeds of theobroma-cacao l. *Food Chemistry*, *54*(1), 15–21. http://dx.doi.org/10.1016/0308-8146(95)92657-6.

de Brito, E. S., Garcia, N. H. P., Gallao, M. I., Cortelazzo, A. L., Fevereiro, P. S., & Braga, M. R. (2001). Structural and chemical changes in cocoa (*Theobroma cacao* L) during fermentation, drying and roasting. *Journal of the Science of Food and Agriculture*, *81*(2), 281–288. http://dx.doi.org/10.1002/1097-0010(20010115)81:2<281::aid-jsfa808>3.0.co;2-b.

Granvogl, M., Bugan, S., & Schieberle, P. (2006). Formation of amines and aldehydes from parent amino acids during thermal processing of cocoa and model systems: New insights into pathways of the strecker reaction. *Journal of Agricultural and Food Chemistry*, *54*(5), 1730–1739. http://dx.doi.org/10.1021/jf0525939.

Khaldi, N., & Shields, D. C. (2011). Shift in the isoelectric-point of milk proteins as a consequence of adaptive divergence between the milks of mammalian species. *Biology Direct*, *6*, 40. http://dx.doi.org/10.1186/1745-6150-6-40.

Kochhar, S., Gartenmann, K., & Juillerat, M. A. (2000). Primary structure of the abundant seed albumin of *Theobroma cacao* by mass spectrometry. *Journal of Agricultural and Food Chemistry*, *48*(11), 5593–5599. http://dx.doi.org/10.1021/jf0006067.

Kratzer, U., Frank, R., Kalbacher, H., Biehl, B., Wostemeyer, J., & Voigt, J. (2009). Subunit structure of the vicilin-like globular storage protein of cocoa seeds and the origin of cocoa- and chocolate-specific aroma precursors. *Food Chemistry*, *113*(4), 903–913. http://dx.doi.org/10.1016/j.foodchem.2008.08.017.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, *227*(5259), 680–685.
Lerceteau, et al. (1999). Evolution of cacao bean proteins during fermentation: A study by two-dimensional electrophoresis. *Journal of the Science of Food and Agriculture*, *79*(4), 619–625. http://dx.doi.org/10.1002/(Sici)1097-0010(19990315)79:4<619::Aid-Jsfa230>;2-F.
Lopez, A. S., & Dimick, P. S. (2008). *Cocoa fermentation biotechnology*. Wiley-VCH Verlag GmbH, 561–577.
Marseglia, A., Sforza, S., Faccini, A., Bencivenni, M., Palla, G., & Caligiani, A. (2014). Extraction, identification and semi-quantification of oligopeptides in cocoa beans. *Food Research International*, *63*, 382–389. http://dx.doi.org/10.1016/j.foodres.2014.03.046.
Ostovar, K., & Keeney, P. G. (1973). Isolation and characterization of microorganisms involved in the fermentation of Trinidad's cacao beans. *Journal of Food Science*, *38*(4), 611–617. http://dx.doi.org/10.1111/j.1365-2621.1973.tb02826.x.
Perkins, et al. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, *20*(18), 3551–3567. http://dx.doi.org/10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2.
Rashidah, S., Jinap, S., Nazamid, S., & Jamilah, B. (2007). Characterisation of the ability of globulins from legume seeds to produce cocoa specific aroma. *ASEAN Food Journal*, *14*(2), 103–114.
Sayers, E. W., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., ... Ye, J. (2009). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, *37*(Database issue), D5–15. http://dx.doi.org/10.1093/nar/gkn741.
Schwan, & Wheals (2004). The microbiology of cocoa fermentation and its role in chocolate quality. *Critical Reviews in Food Science and Nutrition*, *44*(4), 205–221. http://dx.doi.org/10.1080/10408690490464104.
Sforza, S., Cavatorta, V., Lambertini, F., Galaverna, G., Dossena, A., & Marchelli, R. (2012). Cheese peptidomics: A detailed study on the evolution of the oligopeptide fraction in Parmigiano-Reggiano cheese from curd to 24 months of aging. *Journal of Dairy Science*, *95*(7), 3514–3526. http://dx.doi.org/10.3168/jds.2011-5046.
Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., & Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols*, *1*(6), 2856–2860. http://dx.doi.org/10.1038/nprot.2006.468.
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*(4), 567–576. http://dx.doi.org/10.1007/BF00198037.
Strupat, K. (2005). Molecular weight determination of peptides and proteins by ESI and MALDI. *Methods in enzymology. Vol. 405*. (pp. 1–36). Academic Press.
Voigt, et al. (1993). the major seed proteins of *Theobroma-cacao* L. *Food Chemistry*, *47*(2), 145–151. http://dx.doi.org/10.1016/0308-8146(93)90236-9.
Voigt, J., Janek, K., Textoris-Taube, K., Niewienda, A., & Wostemeyer, J. (2016). Partial purification and characterisation of the peptide precursors of the cocoa-specific aroma components. *Food Chemistry*, *192*, 706–713. http://dx.doi.org/10.1016/j.foodchem.2015.07.068.
Voigt, J., & Lieberei, R. (2014). Biochemistry of cocoa fermentation. In R. F. Schwan, & G. H. Fleet (Eds.), *Cocoa and coffee fermentations*. Boca Raton: CRC Press.
Voigt, J., Voigt, G., Heinrichs, H., Wrann, D., & Biehl, B. (1994a). In-vitro studies on the proteolytic formation of the characteristic aroma precursors of fermented cocoa seeds - The significance of endoprotease specificity. *Food Chemistry*, *51*(1), 7–14. http://dx.doi.org/10.1016/0308-8146(94)90040-X.
Voigt, J., Wrann, D., Heinrichs, H., & Biehl, B. (1994b). The proteolytic formation of essential cocoaspecific 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 Chemistry*, *51*(2), 197–205. http://dx.doi.org/10.1016/0308-8146(94)90257-7.
Zak, D. L., & Keeney, P. G. (1976). Changes in cocoa proteins during ripening of fruit, fermentation, and further processing of cocoa beans. *Journal of Agricultural and Food Chemistry*, *24*(3), 483–486. http://dx.doi.org/10.1021/Jf60205a056.
Zhu, K., Zhao, J., Lubman, D. M., Miller, F. R., & Barder, T. J. (2005). Protein pI shifts due to posttranslational modifications in the separation and characterization of proteins. *Analytical Chemistry*, *77*(9), 2745–2755. http://dx.doi.org/10.1021/ac048494w.
Ziegleder, G., & Biehl, B. (1988). Analysis of cocoa flavour components and flavour precursors. In H. -F. Linskens, & J. F. Jackson (Eds.), *Analysis of nonalcoholic beverages* (pp. 321–393). Berlin, Heidelberg: Springer Berlin Heidelberg.
http://www.cbs.dtu.dk/services/YinOYang/. (Accessed 16.08.16).

Supplementary Table S1.

Proteins identified from cocoa beans by MALDI-ToF mass spectropmetry.

Serial number	Protein name	Accession number	Theoretical M.W.(kDa)	Theoretical pI
1	ß-galactosidase	XP_007020169.1	77.1	6.3
2	Threonyl-tRNA synthetase isoform 4	XP_007044585	71.0	6.7
3	Membrane trafficking VPS53 family protein	XP_007039608	63.0	6.5
4	Vicilin Precursor	S22477	62.0	5.0
5	Serine/Threonine specific protein kinase	AAL66190	58.0	6.1
6	UDP-glucosyl transferase family 1 protein	XP_007009133	54.0	6.0
7	Vicilin	CAA44494.1	53.7	6.5
8	Phosphoglycerate kinase	XP_007022578.1	51.0	7.9
9	Enolase	EOY08109.1	47.9	5.6
10	Kinase superfamily protein	XP_007014428	43.0	5.0
11	Peroxidase	XP_007019580	39.0	8.7
12	SBP family protein	XP_007036243.1	37.5	6.4
13	Albumin	1802409A	23.9	5.0
14	Calmodulin like 42	XP_007033453	22.0	4.4
15	Histidine containing phosphotransfer protein 6	XP_007023197	18.0	7.7
16	Oleosin family protein	XP_007031561	16.9	5.0
17	Stilbene synthase	AAT84858	16.2	5.7
18	Trypsin inhibitor	AAL85654.1	16.0	4.5
19	Iron-sulfur cluster biosynthesis family protein isoform 2	EOY08243.1	14.6	5.1

Supplementary Figure S1.

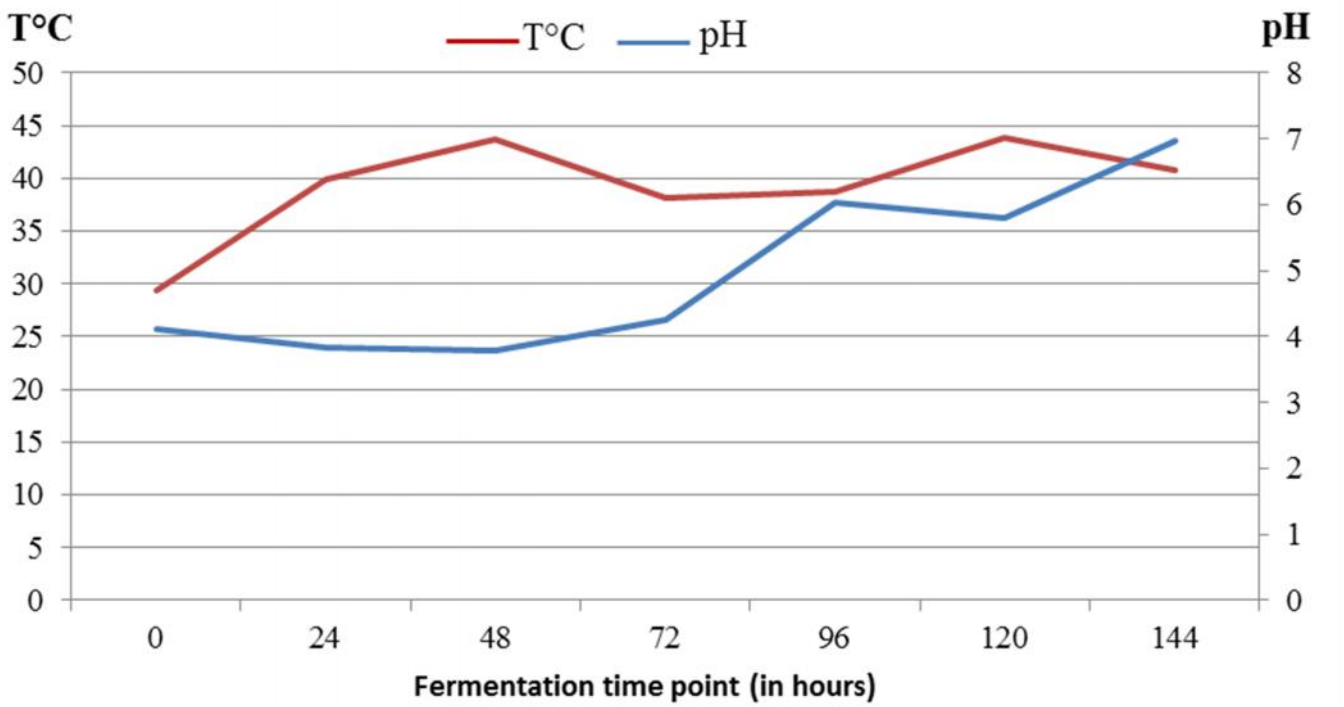

Figure S1: Temperature and pH profile of a spontaneous fermentation of cocoa hybrid G11UTA402XT413. Fermentation was conducted in April 2014 at Champ Semencier de Guessabo, Ivory Coast. Changes in pH and temperature were monitored continuously and cocoa bean samples were collected at 24-h intervals. Red graph indicates change in temperature while blue graph indicates the change in pH.