

Original article

## Determination and analyses of the N-termini of oil-body proteins, steroleosin, caleosin and oleosin

Li-Jen Lin <sup>a</sup>, Pao-Chi Liao <sup>b</sup>, Hsueh-Hui Yang <sup>a</sup>, Jason T.C. Tzen <sup>a,\*</sup>

<sup>a</sup> Graduate Institute of Biotechnology, National Chung-Hsing University, 40227 Taichung, Taiwan, ROC

<sup>b</sup> Department of Environmental and Occupational Health, College of Medicine, National Cheng-Kung University, Tainan, Taiwan, ROC

Received 16 May 2005; received in revised form 7 June 2005; accepted 26 July 2005

Available online 31 August 2005

### Abstract

Seed oil bodies comprise a triacylglycerol matrix shielded by a monolayer of phospholipids and proteins. These surface proteins include an abundant structural protein, oleosin, and at least two minor protein classes termed caleosin and steroleosin. Two steroleosin isoforms (41 and 39 kDa), one caleosin (27 kDa), and two oleosin isoforms (17 and 15 kDa) have been identified in oil bodies isolated from sesame seeds. The signal peptides responsible for targeting of these proteins to oil bodies have not been experimentally determined. Hydrophathy analyses indicate that the hydrophobic domain putatively responsible for oil-body anchoring is located in the N-terminal region of steroleosin, but in the central region of caleosin or oleosin. Direct amino acid sequencing showed that both steroleosin isoforms possessed a free methionine residue at their N-termini while caleosin and oleosin isoforms were N-terminally blocked. Mass spectrometry analyses revealed that N-termini of both caleosin and 17 kDa oleosin were acetylated after the removal of the first methionine. In addition, deamidation was observed at a glutamine residue in the N-terminal region of 17 kDa oleosin.

© 2005 Elsevier SAS. All rights reserved.

**Keywords:** Caleosin; N-terminal modification; Oil-body; Oleosin; Steroleosin

### 1. Introduction

Seeds are the part of a flowering plant in which the initial source of nutrition, such as carbohydrates and lipids, for germination and subsequent seedling growth is stored. Storage carbohydrates and lipids are assembled in discrete subcellular particles, i.e. starch granules and oil bodies [11,16,17]. Seed oil bodies comprise a triacylglycerol matrix surrounded by a monolayer of phospholipids embedded with proteins [29] and are remarkably stable both in cells and isolated preparation as a consequence of the steric hindrance and electronegative repulsion provided by their surface proteins [19]. Three classes of proteins termed oleosin, caleosin, and steroleosin have been identified in oil bodies of diverse seeds [32].

Due to the insolubility of oil-body proteins caused by their hydrophobic oil-body anchoring domains, no three-

dimensional structure derived from X-ray or NMR is available for oleosin, caleosin, or steroleosin at the present time. Putative structures of these three oil-body proteins are predicted based on sequence analyses and spectrometric determination. Oleosin is proposed to comprise three structural domains: an N-terminal amphipathic domain, a central hydrophobic oil-body anchoring domain, and a C-terminal amphipathic  $\alpha$ -helical domain [33]. Similarly, caleosin is proposed to comprise three structural domains: an N-terminal hydrophilic calcium-binding domain, a central hydrophobic oil-body anchoring domain, and a C-terminal hydrophilic phosphorylation domain [6]. In contrast, steroleosin is proposed to comprise an N-terminal oil-body anchoring domain and a soluble sterol-binding dehydrogenase domain [13].

Targeting of oleosin and caleosin to seed oil bodies has been investigated in either in vivo or in vitro systems [1,7,21,34]. A comparable proline knot motif present in the central hydrophobic oil-body anchoring domains of oleosin and caleosin has been assumed to play an important role in their targeting to oil bodies. Targeting of steroleosin to seed oil bodies has not been investigated so far. The proline knot

Abbreviations: SRP, signal-recognition particle.

\* Corresponding author. Tel.: +886 4 2284 0328; fax: +886 4 2285 3527.

E-mail address: TCTZEN@dragon.nchu.edu.tw (J.T.C. Tzen).

motif responsible for oil-body targeting of oleosin and caleosin is not present in steroleosin; instead, a proline knob motif is found in the middle of the N-terminal hydrophobic segment of steroleosin [14].

N-terminal regions of proteins are frequently responsible for the targeting of their leading polypeptides [25]. Modifications of N-termini occur commonly and may be related to the function, stability and turnover rate of proteins [2,10]. As the structural organizations of steroleosin, caleosin, and oleosin are different, we wonder if these oil-body proteins possess similar or different N-termini, which may provide a clue to investigate their targeting and biological functions. Here, we determined and analyzed the N-termini and their modifications of steroleosin, caleosin, and oleosin by direct amino acid sequencing and mass spectrum analyses.

## 2. Results

### 2.1. Direct amino acid sequencing of N-termini of oil-body proteins

Three proteins classes, i.e. steroleosin (41 and 39 kDa), caleosin (27 kDa), and oleosin (17 and 15 kDa) are present in sesame oil bodies (Fig. 1A). Hydropathy plot analyses suggest that the hydrophobic domain putatively responsible for oil-body anchoring is located in the N-terminal region of steroleosin, but in the central region of caleosin or oleosin (Fig. 1B). Direct amino acid sequencing of the first 10 residues showed that both steroleosin isoforms possessed a free methionine at their N-termini (Table 1). In comparison with sequences deduced from their corresponding cDNA clones

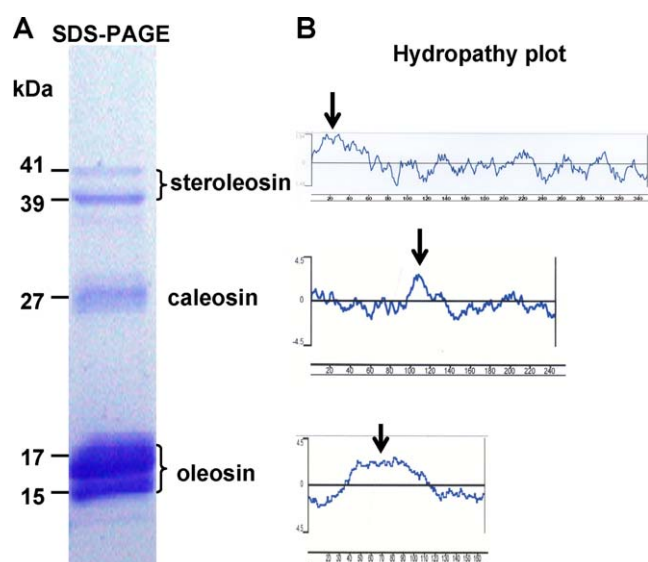


Fig. 1. (A) SDS-PAGE of proteins extracted from sesame oil bodies. Labels on the left indicate the molecular masses of the three types of oil-body proteins. (B) Hydropathy plots of sesame steroleosin, caleosin, and oleosin. Hydrophobicity scale was plotted versus amino acid sequences of sesame steroleosin (39 kDa), caleosin (27 kDa), and oleosin (17 kDa) with a window size of 19 using hydropathy index described by Kyte and Doolittle [12].

Table 1  
Direct N-terminal sequencing of sesame oil-body proteins

	$M_r$	N-terminal amino acid sequence
Steroleosin	41	MDLINSLLNF
	39	MDLIHTFLNL
Caleosin	27	Blocked
Oleosin	17	Blocked
	15	Blocked

(GenBank accession nos. **AF421889**, **AF498264**), the free N-terminal methionine of the purified steroleosin proteins matches the initial methionine residue of the theoretically translated polypeptides. Accordingly, sequence analysis by the SignalP computer program [18] predicts that both steroleosin proteins possess a non-cleavable N-terminal sequence responsible for ER targeting via the signal-recognition particle (SRP) dependent pathway. Evidently, no post-translational cleavage or modification occurs in the N-termini of mature steroleosin isoforms isolated from sesame oil bodies. In contrast, caleosin and both oleosin isoforms were resistant to Edman degradation during direct sequencing, presumably due to their N-terminal block via translational modification. Sequence analyses by the SignalP program predict that no signal sequence responsible for ER targeting is present in the N-terminal 70 residues of caleosin and oleosin isoforms.

### 2.2. Mass spectrum analyses of N-termini of caleosin and oleosin

To reveal the putative modification in the N-termini of caleosin and oleosin, intact caleosin and 17 kDa oleosin were separately subjected to in-gel digestion with trypsin, and their resulting peptides were analyzed by matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) and MALDI-tandem mass spectrometry (MS/MS). Many corresponding tryptic peptides of caleosin and 17 kDa oleosin were observed in MALDI-MS spectra (Figs. 2A and 3A). Two tryptic peptides of  $m/z$  value 1151.60 for caleosin and 2501.24 for 17 kDa oleosin (indicated by arrows) were speculated to match the amino acid residues of A2-R12 in the caleosin and A2-R21 in the 17 kDa oleosin with N-terminal acetylation (calculated monoisotopic  $m/z$  values: 1151.62 and 2500.26, respectively, the difference of +0.98 Da between the observed and calculated values in the latter was later found attributed to deamidation of the peptide; deamidation mass shift: 17.0027 Da [OH] – 16.0187 Da [NH<sub>2</sub>] = +0.984 Da), and therefore subjected to MS/MS analyses. The  $m/z$  values of the product ions generated from low-energy collision-induced dissociation were compared with those expected from amino acid residues of A2-R12 in caleosin and A2-R21 in 17 kDa oleosin with indicated modification. Obviously, the first Met in the N-termini of

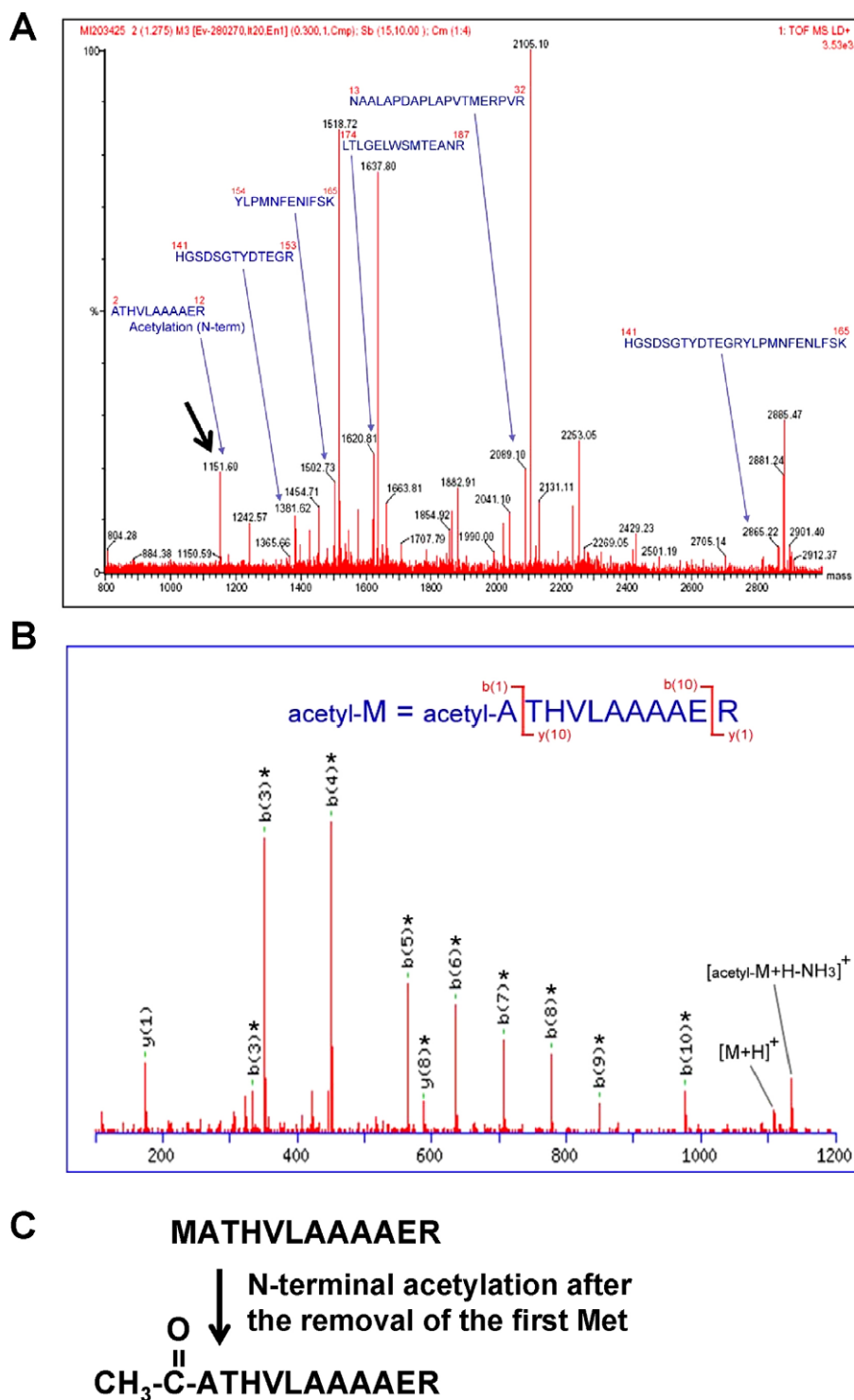


Fig. 2. Mass spectrometry analyses of caleosin. Caleosin was subjected to in-gel digestion with trypsin. (A) Tryptic peptides were analyzed by MALDI-MS. One peptide fragment (indicated by arrow) with an  $m/z$  value of 1151.6, matching the amino acids of A2-R12 with N-terminal acetylation. (B) The sequence of the peptide was subsequently confirmed by MALDI-MS-MS analyses. All matched b and y ions were labeled. These b ions labeled with (\*) confirmed the N-terminal acetylation. (C) The MS/MS data were also subjected to search algorithms against Swiss-Prot protein sequence database and the sequence with indicated modification was matched. The structure of N-terminus of caleosin was proposed according to the mass analyses.

nascent caleosin and 17 kDa oleosin was removed and the following Ala was acetylated in their mature proteins (Figs. 2C and 3C). Peptide fragments with Met at the N-terminus were not observed in the MALDI-MS data in Fig. 3A, B. In Fig. 2A, neither of these two possible Met-starting tryptic peptides

were observed: MATHVLAAAER(M),  $[M + H] m/z$  1240.6; acetylated-M,  $[M + H] m/z$  1282.7. In Fig. 3A, none of the following Met-starting tryptic peptides were observed: MADRDRPHPHQIQVHPQHPHR(M),  $[M + H] m/z$  2590.9; acetylated-M,  $[M + H] m/z$  2631.3; deamidated-M,  $[M + H]$



$m/z$  2591.9; acetylated and deamidated-M,  $[M + H]^+$   $m/z$  2632.3. In addition, deamidation was observed at Q13 of 17 kDa oleosin.

### 3. Discussion

N-terminal acetylation is a co-translational modification found in 50–80% of cytosolic eukaryotic proteins [5]. It may affect biological activities of the modified proteins, such as regulating enzymatic activities [20,23,24], oxygen affinity of hemoglobin [15,26], and microfilament assembly of cytoplasmic actin [4]. Moreover, N-terminal acetylation generally enhances protein thermal stability as the reactive amino group is eliminated [9,22] and commonly impedes protein turnover rate mediated by the ubiquitin-dependent degradation system where proteins with a free N-terminal amino group are preferentially degraded [2,10]. Oil bodies are specialized organelles designed for a long-term storage of lipid fuels as the energy source for seed germination and seedling growth. Since oleosin and caleosin, but not steroleosin, contribute to the structural stability of seed oil bodies [8], we speculate that N-terminal acetylation elevates structural stability and prevents ubiquitinated degradation of these two oil-body structural proteins to fulfill the biological function of long-term protection of the organelles.

Probably the most common post-translational modification occurring in living systems involves non-enzymatic deamidation of glutamine and asparagine residues. This process occurs during the differentiation and/or aging of cells, resulting in the accumulation of glutamate and aspartate, respectively. Because deamidation results in the introduction of negative charges, it is possible that deamidation may play a major role in protein structural changes [27]. Furthermore, it has been hypothesized that non-enzymatic deamidation of glutamine and asparagine residues may act as a “biological clock” to program the turnover of proteins [28]. In this study, deamidation of a glutamine residue was observed in 17 kDa oleosin, and thus contributed more negative charges to the surface of oil bodies. Since the maintenance of seed oil bodies as individual entities without aggregation in the physiological condition is resulted from their surface electronegative repulsion [30], the introduction of more negative charges by deamidation of oleosins may reinforce the stability of these lipid storage organelles. The enhancement of electronegative repulsion on the surface of oil bodies by deamidation of oleosins seems to be tremendous as approximately 1 million oleosin molecules are present on the surface of an oil-body [29].

Steroleosin possesses a non-cleavable N-terminal sequence putatively responsible for ER targeting via the SRP dependent pathway and anchors to oil bodies mainly by this N-terminal hydrophobic domain [14] whereas caleosin and oleosin lacking of an N-terminal signal sequence target/anchor to oil bodies via their central hydrophobic domains [3]. It is possible that steroleosin and caleosin/oleosin may be assembled to maturing oil bodies through various locations of

ER membrane via different auxiliary targeting machinery. In this current study, steroleosin isoforms were determined to possess a free methionine at their N-termini while caleosin and oleosin were found to be N-terminally blocked by acetylation after the removal of the first methionine residue. Presumably, the N-terminus of steroleosin is protected by the SRP complex and/or embedded in the ER membrane during its targeting while the N-termini of caleosin and oleosin are freely exposed to cytosol during their targeting to maturing oil bodies via central hydrophobic domains.

### 4. Methods

#### 4.1. Purification of sesame oil bodies and analysis of oil-body proteins

Mature sesame (*Sesamum indicum* L.) seeds were gifts from the Crop Improvement Department, Tainan District Agricultural Improvement Station. Isolated sesame oil bodies were subjected to further purification including two-layer flotation by centrifugation, detergent washing, ionic elution, treatment of chaotropic agent, and integrity testing with hexane [31]. Proteins extracted from purified oil bodies were resolved by SDS-PAGE using 12.5% polyacrylamide in the separating gel and 4.75% polyacrylamide in the stacking gel. After electrophoresis, the gel was stained with Coomassie Blue R-250 and destained. Hydrophathy profile was plotted with a window size of 19 using hydrophathy index described by Kyte and Doolittle [12]. N-terminal signal sequence responsible for ER targeting via the SRP pathway was predicted using the SignalP program in the World Wide Web Prediction Server Center for Biological Sequence Analysis [18].

#### 4.2. N-terminal amino acid sequencing

Steroleosin, caleosin, and oleosin resolved in an SDS-PAGE gel were transferred onto a piece of PVDF membrane at a current of 0.5 A for 30 min at 4 °C in a blotting buffer of 10% methanol and 10 mM Caps-NaOH, pH 11. After blotting, the PVDF membrane was stained with Coomassie Blue for 5 min, destained for 5 min, rinsed with water three times, and then left to dry in the air. N-terminal sequencing was executed using the Applied Biosystems 476A Protein Sequencer in Chung-Hsing University, Taiwan.

#### 4.3. In-gel digestion of caleosin and oleosin

The protein bands of caleosin and oleosin were manually excised from the gel and ground into pieces. All in-gel digestions of proteins were performed manually with trypsin in a laminar flow hood to reduce the amount of keratin contamination. The gel pieces were washed twice with 50% acetonitrile and 50% acetonitrile/25 mM ammonium bicarbonate. The proteins in-gel were then reduced and alkylated at 56 °C for 45 min in 10 mM dithiothreitol (DTT) and 55 mM iodoac-

etamide in 25 mM ammonium bicarbonate, followed by in-gel digestion with 0.1 µg of TPCK-treated modified porcine trypsin (Promega, Madison, WI, USA) in the same buffer with overnight incubation at 37 °C. The supernatant containing resulting tryptic peptides was combined with those extracted twice from the gel pieces by 50% acetonitrile/5% formic acid and subjected to mass spectrometry analysis.

#### 4.4. MALDI-MS and MALDI-MS/MS analyses of caleosin and oleosin

Tryptic peptides derived from in-gel digestion of caleosin and oleosin were analyzed by MALDI-MS and MALDI-MS/MS. All data were acquired by quadrupole-time-of-flight (Q-TOF) hybrid mass spectrometers (Micromass Q-ToF Ultima, Manchester, UK, and Applied Biosystems QSTAR, Foster City, CA, USA). The matrix used was α-cyano-4-hydroxycinnamic acid. The low-energy collision-induced dissociation MS/MS product ion spectra acquired from Q-ToF Ultima and QSTAR were analyzed using the Micromass ProteinLynx™ Global Server 2.0 and Applied Biosystems BioAnalyst™ data processing software, respectively. The MS/MS data were also subjected to search algorithms against Swiss-Prot protein sequence database using Mascot software (Matrix Science Ltd., London, UK).

#### Acknowledgments

The authors thank the Proteomics Research Core Laboratory at National Cheng-Kung University for the assistance in mass spectrometry analysis and Dr. Tien-Joung Yiu of the Crop Improvement Department, Tainan District Agricultural Improvement Station for supplying sesame seeds. The work was supported by a grant from the National Science Council, Taiwan, ROC (93-2313-B-005-054 to J.T.C. Tzen).

#### References

- [1] B.M. Abell, L.A. Holbrook, M. Abenes, D.J. Murphy, M.J. Hills, M.M. Moloney, Role of the proline knot motif in oleosin endoplasmic reticulum topology and oil body targeting, *Plant Cell* 9 (1997) 1481–1493.
- [2] A. Bachmair, D. Finley, A. Varshavsky, In vivo half-life of a protein is a function of its amino-terminal residue, *Science* 234 (1986) 179–186.
- [3] F. Beaudoin, J.A. Napier, Targeting and membrane-insertion of a sunflower oleosin in vitro and in *Saccharomyces cerevisiae*: the central hydrophobic domain contains more than one signal sequence, and directs oleosin insertion into the endoplasmic reticulum membrane using a signal anchor sequence mechanism, *Planta* 215 (2002) 293–303.
- [4] E.M. Berger, G. Cox, L. Weber, J.L. Kenney, Actin acetylation in *Drosophila* tissue culture cells, *Biochem. Genet.* 19 (1981) 321–331.
- [5] J.L. Brown, W.K. Roberts, Evidence that approximately eighty per cent of the soluble proteins from *Ehrlich ascites* cells are Nalpha-acetylated, *J. Biol. Chem.* 251 (1976) 1009–1014.
- [6] J.C.F. Chen, C.C.Y. Tsai, J.T.C. Tzen, Cloning and secondary structure analysis of caleosin, a unique calcium-binding protein in oil bodies of plant seeds, *Plant Cell Physiol.* 40 (1999) 1079–1086.
- [7] J.C.F. Chen, J.T.C. Tzen, An in vitro system to examine the effective phospholipids and structural domain for protein targeting to seed oil bodies, *Plant Cell Physiol.* 42 (2001) 1245–1252.
- [8] M.C.M. Chen, C.L. Chyan, T.T.T. Lee, S.H. Huang, J.T.C. Tzen, Constitution of stable artificial oil bodies with triacylglycerol, phospholipid, and caleosin, *J. Agric. Food Chem.* 52 (2004) 3982–3987.
- [9] A.G. Cumberlidge, K. Isono, Ribosomal protein modification in *Escherichia coli*. I. A mutant lacking the N-terminal acetylation of protein S5 exhibits thermosensitivity, *J. Mol. Biol.* 131 (1979) 169–189.
- [10] A. Hershko, H. Heller, E. Eytan, G. Kaklii, I.A. Rose, Role of the alpha-amino group of protein in ubiquitin-mediated protein breakdown, *Proc. Natl. Acad. Sci. USA* 81 (1984) 7021–7025.
- [11] A.H.C. Huang, Oleosin and oil bodies in seeds and other organs, *Plant Physiol.* 110 (1996) 1055–1061.
- [12] J. Kyte, R.F. Doolittle, A simple method for displaying the hydrophobic character of a protein, *J. Mol. Biol.* 157 (1982) 105–132.
- [13] L.J. Lin, S.S.K. Sorgan, C.C. Peng, J.T.C. Tzen, Steroleosin, a sterol-binding dehydrogenase in seed oil bodies, *Plant Physiol.* 128 (2002) 1200–1211.
- [14] L.J. Lin, J.T.C. Tzen, Two distinct steroleosins are present in seed oil bodies, *Plant Physiol. Biochem.* 42 (2004) 601–608.
- [15] W.F. Moo-Penn, K.C. Bechtel, R.M. Schmidt, M.H. Johnson, D.L. Jue, D.E.J. Schmidt, W.M. Dunlap, S.J. Opella, J. Bonaventura, C. Bonaventura, Hemoglobin Raleigh (beta1 valine replaced by acetylalanine). Structural and functional characterization, *Biochem.* 16 (1977) 4872–4879.
- [16] D.J. Murphy, Structure, function and biogenesis of storage lipid bodies and oleosins in plants, *Prog. Lipid Res.* 32 (1993) 247–280.
- [17] J.A. Napier, A.K. Stobart, P.R. Shewry, The structure and biogenesis of plant oil bodies: the role of the ER membrane and the oleosin class of proteins, *Plant Mol. Biol.* 31 (1996) 945–956.
- [18] H. Nielsen, J. Engelbrecht, S. Brunak, H.G. Von, Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites, *Protein Eng.* 10 (1997) 1–6.
- [19] C.C. Peng, I.P. Lin, C.K. Lin, J.T.C. Tzen, Size and stability of reconstituted sesame oil bodies, *Biotechnol. Prog.* 19 (2003) 1623–1626.
- [20] J. Ramachandran, C.H. Lin, Structure–activity relationships of the adrenocorticotropins and melanotropins: the synthetic approach. Structure–activity relationships of the adrenocorticotropins and melanotropins: the synthetic approach, *Adv. Enzymol.* 29 (1967) 391–477.
- [21] C. Sarmiento, J.H. Ross, E. Herman, D.J. Murphy, Expression and subcellular targeting of a soybean oleosin in transgenic rapeseed. Implications for the mechanism of oil-body formation in seeds, *Plant J.* 11 (1997) 783–796.
- [22] M.A. Siddig, J.A. Kinsey, J.R. Fincham, M. Keighren, Frameshift mutations affecting the N-terminal sequence of *Neurospora* NADP-specific glutamate dehydrogenase, *J. Mol. Biol.* 137 (1980) 125–135.
- [23] D.G. Smyth, D.E. Massey, S. Zakarian, M.D.A. Finnie, Endorphins are stored in biologically active and inactive forms: isolation of alpha-N-acetyl peptides, *Nature* 279 (1979) 252–254.
- [24] D.G. Smyth, S. Zakarian, Selective processing of beta-endorphin in regions of porcine pituitary, *Nature* 288 (1980) 613–615.
- [25] R.M. Stroud, P. Walter, Signal sequence recognition and protein targeting, *Curr. Opin. Struct. Biol.* 9 (1999) 754–759.
- [26] F. Taketa, A.G. Mauk, J.L. Lessard, β Chain amino termini of the cat hemoglobins and the response to 2,3-diphosphoglycerate and adenosine triphosphate, *J. Biol. Chem.* 246 (1971) 4471–4476.
- [27] L. Takemoto, D. Boyle, Specific glutamine and asparagine residues of gamma-S crystallin are resistant to in vivo deamidation, *J. Biol. Chem.* 275 (2000) 26109–26112.
- [28] L. Takemoto, D. Boyle, Deamidation of specific glutamine residues from alpha-A crystallin during aging of the human lens, *Biochemistry* 37 (1998) 13681–13685.

- [29] J.T.C. Tzen, A.H.C. Huang, Surface structure and properties of plant seed oil bodies, *J. Cell Biol.* 117 (1992) 327–335.
- [30] J.T.C. Tzen, G.C. Lie, A.H.C. Huang, Characterization of the charged components and their topology on the surface of plant seed oil bodies, *J. Biol. Chem.* 267 (1992) 15626–15634.
- [31] J.T.C. Tzen, C.C. Peng, D.J. Cheng, E.C.F. Chen, J.M.H. Chiu, A new method for seed oil body purification and examination of oil body integrity following germination, *J. Biochem. (Tokyo)* 121 (1997) 762–768.
- [32] J.T.C. Tzen, M.M. Wang, J.C.F. Chen, L.J. Lin, M.C.M. Chen, Seed oil body proteins: oleosin, caleosin, and steroleosin, *Current Topic in Biochem. Res.* 5 (2003) 133–139.
- [33] V.B. Vance, A.H.C. Huang, The major protein from lipid bodies of maize. Characterization and structure based on cDNA cloning, *J. Biol. Chem.* 262 (1987) 11275–11279.
- [34] T. Wahlroos, J. Soukka, A. Denesyuk, R. Wahlroos, T. Korpela, N.J. Kilby, Oleosin expression and trafficking during oil body biogenesis in tobacco leaf cells, *Genesis* 35 (2003) 125–132.