analytical chemistry

ARTICLE

pubs.acs.org/ac

# LC-MS/MS Methods for Absolute Quantification and Identification of Proteins Associated with Chimeric Plant Oil Bodies

Floriana Capuano,[†,‡,§] Nicholas J. Bond,[‡] Laurent Gatto,[‡] Frédéric Beaudoin,[§] Johnathan A. Napier,[§] Eugenio Benvenuto,[†] Kathryn S. Lilley,[‡] and Selene Baschieri*[,†]

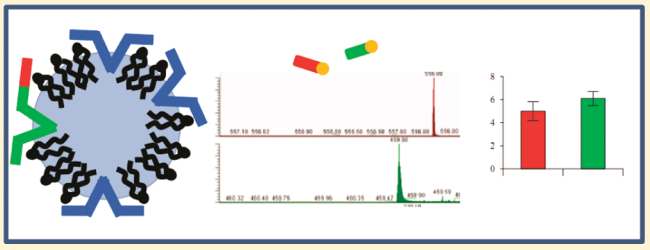
[†]Laboratorio Biotecnologie Unità Tecnica BIORAD, ENEA C.R. Casaccia, 00123 Roma, Italy

[‡]Department of Biochemistry, University of Cambridge, Building O, Downing Site, Cambridge CB2 1QW, United Kingdom

[§]Department of Biological Chemistry, Rothamsted Research, Harpenden, Herts, AL5 2JQ, United Kingdom

Supporting Information

**ABSTRACT:** Oil bodies (OBs) are plant cell organelles that consist of a lipid core surrounded by a phospholipid monolayer embedded with specialized proteins such as oleosins. Recombinant proteins expressed in plants can be targeted to OBs as fusions with oleosin. This expression strategy is attractive because OBs are easily enriched and purified from other cellular components, based on their unique physicochemical properties. For recombinant OBs to be a potential therapeutic agent in biomedical applications, it is necessary to comprehensively analyze and quantify both endogenous and heterologously expressed OB proteins. In this study, a mass spectrometry (MS)-based method was developed to accurately quantify an OB-targeted heterologously expressed fusion protein that has potential as a therapeutic agent. The effect of the chimeric oleosin expression upon the OB proteome in transgenic plants was also investigated, and the identification of new potential OB residents was pursued through a variety of liquid chromatography (LC)-MS/MS approaches. The results showed that the accumulation of the fusion protein on OBs was low. Moreover, no significant differences in the accumulation of OB proteins were revealed between transgenic and wild-type seeds. The identification of five new putative components of OB proteome was also reported.



Oil bodies (OBs) are plant cell organelles with a diameter of 0.5−2 $\mu$m, mainly found in the endosperm and embryo cells of seeds.[1] They consist of a hydrophobic triacylglycerol (TAG) core surrounded by a half-unit phospholipid membrane containing specialized proteins, of which oleosins are the most abundant.[2] Targeting of recombinant proteins to OBs can be achieved by fusion of the recombinant protein to an oleosin to form a chimeric protein, leading to the accumulation in seeds of chimeric OBs. The unique physicochemical features of OBs have enabled the development of methods that allow for easy separation from other cell components[3] of the recombinant product, heterologously expressed in plants and targeted to OBs. Chimeric OBs have the potential to be utilized in several biomedical applications, including use as vaccine carriers, where the OB itself is used for vaccine delivery without further purification of the fusion protein.[4] For chimeric OBs to be employed as biotherapeutic agents, it is of paramount importance to accurately quantify the amount of heterologous polypeptides associated with OBs. It is also necessary to have an exhaustive catalog of the organelle proteome and ensure that the protein content of the chimeric OB is not compromised as a consequence of the heterologous protein accumulation.

In this study, an Absolute QUAntitation (AQUA) mass spectrometry (MS)-based method was developed to measure the absolute abundance of a human immunodeficiency virus Type 1 (HIV-1) derived polypeptide[5] expressed as a C-terminal fusion to a sunflower oleosin to promote targeting to *Arabidopsis thaliana* OBs. Indeed, quantification of the polypeptide using gel-based densitometry was not possible as the chimeric oleosin co-migrated with the endogenous *Arabidopsis* oleosins. Furthermore, antibodies for direct detection of the HIV-1 derived polypeptide are currently not available. The MS-based AQUA method[6] was selected because it is capable of accurately quantifying proteins or protein isoforms by carrying out measurements on peptide surrogates produced upon proteolytic digestion of the protein of interest.[7] This approach relies on the comparison of mass spectrometric measurements of ions derived from peptide surrogates with peptide standards (which are typically isotopic variants of a peptide surrogate) spiked into the experiment in known amounts.

MS-based experimental approaches were also applied to establish the effect of the heterologous polypeptide expression on OB protein content in transgenic *A. thaliana* OB preparations and to identify new potential OB residents.

**Received:** July 5, 2011
**Accepted:** October 21, 2011
**Published:** October 21, 2011

## EXPERIMENTAL SECTION

**Construct Preparation, Plant Transformation, and Selection.** The double-stranded DNA encoding for a polypeptide ($Nef_{125-151}$; amino acid sequence: QNYTPGPGIRYPLTFGWCYKLVPVEPE) derived from the HIV-1 protein Nef (Genbank Accession No. NC_001802) flanked at both ends by the AvrII restriction enzyme target sequence was obtained by annealing the synthetic oligonucleotides *in vitro*:

5′-Nef125

(5′CTAGGCAAAATTATACTCCTGGACCTGGAATTCGT

TATCCTCTTACTTTTGGATGGTGTTATAAGCTTGTTC

CTGTTGAACCTGAATAACTCGAGC3′)

3′-Nef125

(5′CTAGGCTCGAGTTATTCAGGTTCAACAGGAACAAG

CTTATAACACCATCCAAAAGTAAGAGGATAACGAATTC

CGGATCCAGGAGTATAATTTTGC3′)

The DNA fragment was cloned into the pET3d (Novagen)-derived plasmid pOBP/AvrII, as a fusion to the 3′-end of the cDNA coding for the *Helianthus annuus* (sunflower) 19K Da oleosin (Genebank Accession No. X78679).[8] The cassette was inserted into the vector pKSM2 (kindly provided by Dr. E. Cahoon, University of Nebraska) between the seed-specific *Glycine max cv* Century p24 oleosin promoter and terminator sequences and then transferred to the binary vector pBIN19 (pBIN19-GM-OBPNef125).

*A. thaliana* plants *cv* Columbia were transformed using the method of Clough and Bent[9] by dipping the floral buds into cultures of *Agrobacterium tumefaciens* carrying pBIN19-GM-OBPNef125. Transformant plants were selected onto a kanamycin-containing medium and grown onto $T_3$ and homozygosity.

**OB Purification.** *A. thaliana* seeds were processed for OB purification, as described by Tzen et al.[10] Briefly, mature dry seeds were ground in 0.6 M sucrose, 1 mM EGTA, 1 mM PMSF, 10 mM sodium phosphate buffer pH 7.5 at 4 °C. After cell debris removal, the extract (total seed extract, step 1) was layered under a flotation medium (0.4 M sucrose, 10 mM sodium phosphate buffer, pH 7.5) and spun at 10000 *g* for 20 min in a swinging-bucket rotor. OBs were collected as a floating fat pad and then washed with 0.1% Tween 20 (step 2), 2 M NaCl (step 3), and 8 M urea, to remove seed contaminants, and then they were finally suspended in 10 mM sodium phosphate buffer, pH 7.5, before storage at −80 °C (purified OBs, step 4).

**Preparation of Protein Samples.** Proteins from OBs were precipitated by incubation at −20 °C for 3 h by adding 100% cold acetone after lipid removal, as described by Wang et al.[11] Samples were centrifuged at 15000 *g* for 10 min and pellet-washed with 80% acetone.

**Protein Separation through SDS-PAGE.** Protein pellets were suspended in a buffer containing 6 M urea, 2% SDS, pH 8.0 and separated based on their molecular weight on 12% polyacrylamide denaturing gels, using the glycine buffer system (SDS-PAGE).[12] Previous to SDS-PAGE, protein denaturation was performed by heating at 100 °C for 5 min. Proteins were visualized by staining with Colloidal Coomassie Brilliant Blue G-250.

**Sample Preparation for Liquid Chromatography (LC)-MS/MS Analysis.** For in-gel digests followed by LC-MS/MS analysis, Coomassie-stained SDS-PAGE pieces were excised and washed in 25 mM $NH_4HCO_3$, pH 8.0, 50% acetonitrile. Proteins were first reduced by incubation with 10 mM DTT at 37 °C for 1 h and then alkylated with 10 mM iodoacetamide for 15 min at room temperature (RT) in darkness. After dehydration, 50 μL of a trypsin (Promega) solution in 25 mM $NH_4HCO_3$, pH 8.0 (15 ng/mL), were added to the gel pieces and incubated overnight at 37 °C. Samples were desalted using ZipTips (Millipore) composed of a plug of $C_{18}$ packing material prior to being subjected to LC-MS/MS analysis.

For in-solution digest followed by LC-MS/MS analysis, OB protein pellets were resuspended in 50 mM $NH_4HCO_3$ pH 8.0, 60% methanol, to a final concentration of 0.1−0.5 μg/μL and subjected to reduction (25 mM DTT) and alkylation (12.5 mM iodoacetamide). Trypsin was added to the sample with a substrate-to-enzyme ratio of 20:1 (w/w), and the reaction was incubated overnight at 37 °C. Formic acid was added to the sample to a final concentration of 0.1%.

**Mass Spectrometry, Protein Identification, and Absolute Quantification.** All LC-MS/MS experiments designed for protein identification were performed using an Eksigent NanoLC-1D Plus (Eksigent Technologies, Dublin, CA) high-performance liquid chromatography (HPLC) system coupled to an linear trap quadropole (LTQ) Orbitrap mass spectrometer (ThermoFisher, San Jose, CA). Separation of peptides was performed using reverse-phase chromatography, operated at a flow rate of 300 nL/min, and an LC-Packings (Dionex, Sunnyvale) PepMap 100 column (C18, 75 μm i.d. × 150 mm, 3 μm particle size). Peptides were loaded onto a precolumn (Dionex Acclaim PepMap 100 C18, 5 μm particle size, 100 Å, 300 μm i.d. × 5 mm) and washed with 0.1% formic acid for 5 min at a flow rate of 10 μL/min. After this period, peptides were eluted from the precolumn and resolved over the analytical column via a linear increase in Solvent B (0.1% formic acid in acetonitrile) from 5% to 50% over a period of 50 min. Solvent A was 0.1% formic acid in water. The LC eluant was sprayed into the mass spectrometer by means of a New Objective nanospray source. All $m/z$ values of eluting ions were measured in the Orbitrap mass analyzer, set at a resolution of 7500. Doubly and triply charged peptides were then isolated and fragmented in the LTQ linear ion trap by collision-induced dissociation, and MS/MS spectra were acquired.

MS data were analyzed using the Mascot search engine (Version 2.2) and searched against an *A. thaliana* FASTA database downloaded from the Arabidopsis Information Resource (TAIR, Version 9.0) website (www.arabidopsis.org/), in which the unmodified and the chimeric sunflower oleosins sequences also were included. Proteins identified by at least two unique peptides with an ion score of >30 were considered to be present, which represented a 5% probability of a random match.

The exponentially modified Protein Abundance Index (emPAI), equal to $10^{PAI} - 1$, proportional to the abundance of a given protein, was displayed within the Mascot result report for each protein, whenever the MS/MS search reported at least 100 spectra. The emPAI values represent the number of experimentally observed tryptic peptides over the calculated number of observable tryptic peptides for each protein.[13] The count of observed peptides was performed by the Mascot search engine by

| Protein Name | Peptide Sequence | Z | Q1 | Q3 (m/z) | | |
|---|---|---|---|---|---|---|
| | | | | y6 | y8 | y9 |
| Sunflower oleosin 19K (Q39952) | HHVTTTQPQYR | 2 | 689.84 | 802.4082 | 1004.5035 | 1103.5719 |
| | HHVTTTQPQYR | 2 | 684.84 | 792.4082 | 994.5035 | 1093.5719 |
| | | | | y4 | y5 | |
| | HHVTTTQPQYR | 3 | 459.89 | 573.3019 | 701.3605 | |
| | HHVTTTQPQYR | 3 | 456.56 | 563.3019 | 691.3605 | |
| | | | | y6 | y7 | y8 |
| OBPNef125 | QNYTPGPGIR | 2 | 556.6 | 606.3597 | 707.4074 | 870.4708 |
| | QNYTPGPGIR | 2 | 551.6 | 596.3597 | 697.4074 | 860.4708 |

**Figure 1.** Transitions acquired using multiple reaction monitoring (MRM) protocol for absolute quantification of the OBPNef125 chimeric oleosin. (Legend: Z, peptide charge; Q1, $m/z$ value of parent ions; Q3 ($m/z$), $y$-ion transition.)

only including peptide matches with scores at or above the significance threshold of 0.05.

Absolute quantification of the chimeric sunflower oleosin fused to the polypeptide Nef125 (OBPNef125) was achieved using an isotope dilution approach, coupled with multiple reaction monitoring (MRM).[14] Briefly, the most suitable proteotypic peptide candidates were chosen as internal standards (AQUA peptides) by filtering peptides with supporting MS/MS spectra according to their uniqueness, size, residue composition, and quality of mass spectra. The HHVTTTQPQYR and QNYTPGPGIR amino acid sequences, unique to the sunflower oleosin and Nef polypeptide sequences, respectively, were used as templates for the synthesis of AQUA peptides (Anaspec), incorporating a single $^{13}C_6$ $^{15}N_4$ arginine at the C-terminus and each independently quantified by amino acid analysis.[15] Thus, synthetic stable isotope labeled peptides had an increase in mass of 10 Da and had identical chromatographic separation, ionization efficiency, and fragment ion distribution as their unlabeled counterparts. Their linearity range was assessed to be over 2 orders of magnitude on an LTQ-Orbitrap. To quantify the chimeric oleosin in OB preparations, 12 $\mu$g OB proteins were spiked with a mixture containing 100 fmol of each standard peptide and digested overnight by incubation with trypsin (1:20 w/w, trypsin/protein) in a buffer containing 60% methanol, 50 mM $NH_4HCO_3$. The efficiency of trypsin digestion was assessed by monitoring the ratio between the doubly charged labeled and unlabeled parent ions over time and determined to be complete after 16 h. Protein digests were dried to remove methanol and suspended in 30 $\mu$L of HPLC water. Quantitative analysis was conducted using a nanoAcquity ultraperformance liquid chromatography (UPLC) coupled to a Quattro Premier XE (Waters, Milford, MA) triple quadrupole mass spectrometer. Protein digests were diluted 4-fold in 0.1% formic acid and 3 $\mu$L of the sample loaded on to a precolumn (Symmetry C18, 180 $\mu$m i.d. $\times$ 20 mm, 5 $\mu$m p.d., 300 Å p.s.) and washed with 0.1% formic acid in a flow of 5 $\mu$L/min for 5 min. The exhaust flow was then blocked, preventing solvent flowing to waste; instead, flow continued over the BEH300 C-18 analytical column (1.7 $\mu$m, 100 $\mu$m $\times$ 100 mm, 300 Å p.s.) at 300 nL/min. Peptides were separated via reverse-phase chromatography, using gradients of linearly increasing Solvent B concentration (0% to 45%) over 40 min. An MRM experiment was configured to quantify each peptide ion by monitoring transitions designed toward the two or three most intense fragment ions, each with a dwell time of 50 ms. Signals derived from internal standards and analytes were mean-smoothed (window of 2 with one iteration) and integrated using QuantLynx software (Waters). The ratio of standard and analyte peak integrals was used to interpolate the analyte concentration. Here, only precursor ions corresponding to the peptides to be measured were selected for fragmentation and a set of fragment ions for each precursor selected in the final quadrupole for measurement. These fragment ions or transitions, summarized in Figure 1, were chosen based on the fact that they were unique and unlikely to form as the result of the fragmentation of other precursors.

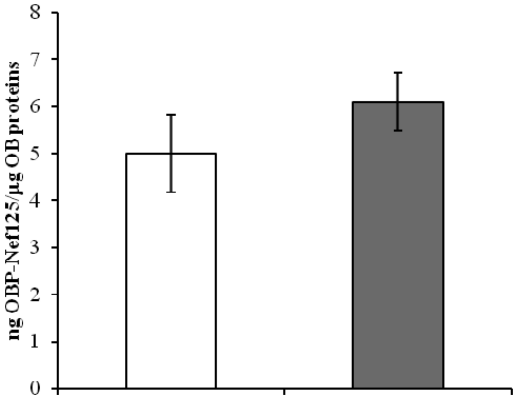



**Figure 2.** Absolute quantification of the OBPNef125 chimeric oleosin using the MRM transitions to monitor QNYTPGPGIR ($Nef_{125-151}$; white bar) and HHVTTTQPQYR (sunflower oleosin; gray bar) AQUA peptides mapping in the HIV-derived and sunflower oleosin sequence, respectively. The MRM analysis was performed in duplicate on trypsin-digested OB proteins spiked with 100 fmol of each standard peptide. The amount of the chimeric protein was 5 $\pm$ 0.8 ng/$\mu$g of total protein, when the quantification was performed with the QNYTPGPGIR peptide, and 6.1 $\pm$ 0.6 ng/$\mu$g of total protein, when performed with the HHVTTTQPQYR peptide. Error bars represent the standard deviation calculated over two technical replicates.


dx.doi.org/10.1021/ac201733m |*Anal. Chem.* 2011, 83, 9267–9272

## RESULTS AND DISCUSSION

**Quantification of a Heterologous Polypeptide Expressed in Association to *A. thaliana* Chimeric OBs.** *A. thaliana* plants expressing the OBPNef125 polypeptide were produced by *A. tumefaciens*-mediated stable transformation, using the binary vector pBIN19-GM-OBPNef125. OBs were prepared from seeds of the highest expressing homozygous line ($T_3$). Oleosin fusion expression was verified by Western blot, using an antisunflower oleosin polyclonal antibody (data not shown), and by in-gel and in-solution MS-based approaches (see below).

Absolute quantification of the chimeric sunflower oleosin fused to the polypeptide Nef125 using, as internal standards, the synthetic isotope-labeled HHVTTTQPQYR or QNYTPGPGIR AQUA peptides, unique to the sunflower oleosin and Nef polypeptide sequences, respectively, gave similar results (see Figure 2).

The chimeric oleosin was estimated to be 5 ng/$\mu$g of total proteins in the OBs purified from transgenic seeds, hence, 17.5 $\mu$g/g seeds. Previous studies showed that effective immune responses were elicited in mice by injecting microgram levels of peptide.[16,17] Thus, the amount of antigen that could potentially be delivered by immunization with OB preparations might be too low for vaccine purposes. Little or no degradation occurred at the C-terminus of the OBPNef125 fusion, as confirmed by the similar values obtained when the quantification of the target protein was undertaken using the two different peptides. Results also indicated that there was no interference by the flanking sequences on the efficiency of trypsin-induced release of the two peptides.

**Evaluation of the Effects of OBPNef125 Expression upon OB Protein Composition by In-Solution LC-MS/MS.** Three OB samples were independently purified from seeds of both wild-type (WT1, WT2, WT3) and transgenic (TG1, TG2, TG3) *A. thaliana* plants to comprehensively characterize the protein content and to establish the effect of the heterologous polypeptide expression upon the OB proteome by in-solution LC-MS/MS. Protein pellets were digested and analyzed using an LTQ-Orbitrap. Peptides were identified with Mascot using the *A. thaliana* FASTA database downloaded from the Arabidopsis Information Resource (TAIR, Version 9.0) website (www.arabidopsis.org/), including the OBPNef125 sequence. Using this approach, 17 (WT1), 24 (WT2), 20 (WT3), 19 (TG1), 26 (TG2), and 29 (TG3) proteins were identified in each sample (see Tables S-1 to S-6 in the Supporting Information). As expected, the presence of the OBPNef125 protein was verified exclusively in transgenic samples. However, tryptic peptides mapping on the Nef polypeptide were only detected in TG1 and TG3. The peptides mapping in the sunflower oleosin moiety of the oleosin chimera were, instead, detected for all of the TG samples. The MS data obtained for the TG samples, and the list of peptides identified, were deposited into the PRIDE database (PRoteomics IDEntifications, http://www.ebi.ac.uk/pride/),[21] under Accession Nos. 19642−19644, after being converted by the PRIDE Converter tool (http://code.google.com/p/pride-converter).[22]

To evaluate the effects of OBPNef125 expression upon OB protein composition, a statistical analysis of the shared proteins was performed. Twelve (12) proteins were in common, and by allowing one missing value per group, this number was increased to 17 (see Table S-7 in the Supporting Information). Although less powerful, this assumption enabled the monitoring of the variation in abundance of a larger number of proteins. The emPAI values[13] of these proteins were $\log_2$ transformed and the differential expression was statistically assessed, using linear models and the empirical Bayesian model, as implemented in the Bioconductor[23] package limma.[24] After correction for multiple testing,[25] no statistically significant difference in OB protein accumulation was detected between wild-type and transgenic seeds.

Among the 12 proteins identified in all of the samples, the most abundant oleosin isoforms (18.5, 19.7, 20.3, and 21.2 kDa) and the 11-$\beta$-hydroxysteroid dehydrogenase-like protein, previously described as OB residents,[26] were identified. Proteins known to be associated with seed storage reserves (AT2S3, At4g27160; PAP85, At3g22640; Cruciferin A (CRA), At5g44120; Cruciferin B (CRB), At1g03880 and Cruciferin C (CRC), At4g28520) or enzymatic activity (chloroplast aldehyde reductase (ChlADR), At1g54870) were also found, as well as the embryo specific protein 3 ATS3 (At5g07190).

AT2S3 is described by the Gene Ontology (GO) annotations in the UniProt Knowledgebase (UniProtKB) (http://geneontology.org/) as a protein with lipid binding and nutrient reservoir functions. In a yeast two-hybrid screen, the interaction between this protein and the cytokinin receptor AHK2 has been demonstrated.[27] Different cytokinins exert a specific action on reserve metabolism and, in particular, isoprenoid cytokinins are active in lipid mobilization.[28] Therefore, these data may suggest the involvement of AT2S3 in the metabolic pathway activated by cytokinins.

The presence of the other storage proteins may be either a consequence of the mechanisms of OB biogenesis or a result of OB interaction with protein storage vacuoles.[29] Storage proteins are actively synthesized on the rough ER,[30] where OB budding is proposed to occur.

ChlADR is a member of the short-chain dehydrogenase/reductase (SDR) family and shows 31% identity and 54% similarity to the OB resident 11-$\beta$-hydroxysteroid dehydrogenase-like protein. The SDR family of proteins includes enzymes that catalyze a wide range of activities including the metabolism of steroids, cofactors, carbohydrates, lipids, aromatic compounds, and amino acids and act in redox sensing. The association to OBs of this enzyme is proposed herein, where it could play a role in lipid metabolism.

ATS3 was also identified, confirming recently published data.[31] This protein is described by the GO as a membrane component and is involved in the embryo development process. ATS3 has a PLAT/LH2 domain, commonly found in lipase and lipoxygenase, which is responsible for interactions with lipid matrixes. Although the function of ATS3 is still unknown, the presence of this domain may promote ATS3-specific targeting to OBs.

Two further proteins (the Saposin-like aspartyl protease family protein, At1g62290, and GTP binding elongation factor, At1g07920) were detected in five samples out of six. The Saposin-like aspartyl protease family protein is described by GO as a protein associated with vacuoles. Its presence in the purified OB preparation could be a consequence of the interaction with caleosins. This protein could favor the mobilization of storage lipids by vacuole association.[29] The identification of the GTP binding elongation factor that associates with the ribosome during polypeptide chain elongation may result from its presence in the preparation of immature OBs carrying the translational machinery associated with proteins synthesized on the rough ER.

It is plausible that not all the identified proteins (seed storage proteins, ChlADR, vacuolar proteins) specifically target to

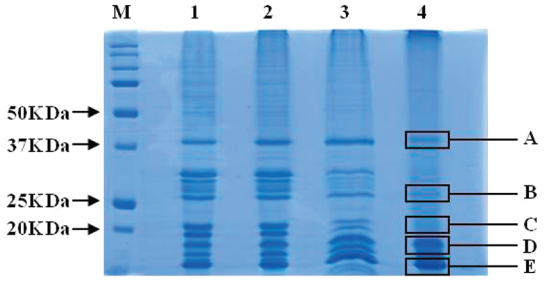



**Figure 3.** SDS-PAGE of samples collected at each step of the OB purification procedure and LC-MS/MS analysis of proteins in purified OBs. Proteins from total seed extract (50 $\mu$g, lane 1) and proteins from OBs washed with 0.1% Tween 20 (50 $\mu$g, lane 2), then with 2 M NaCl (50 $\mu$g, lane 3) and finally with 8 M urea (50 $\mu$g, lane 4) have been separated on 12% polyacrylamide gels. ("M" denotes data for the Precision Plus All Blue Protein Standard (Bio-Rad).) Legend of proteins identified in Coomassie visible bands (A−E) of purified OBs by LC-MS/MS: A, 11-$\beta$-hydroxysteroid dehydrogenase-like ((AtHSD1, Steroleosin*; At5g50600), hydroxysteroid dehydrogenase 5 (AtHSD5, At4g10020); B, glycosylphosphatidylinositol (GPI)-anchored protein (At1g54870), $\beta$-TIP ($\beta$-tonoplast intrinsic protein, At1g17810); C, AtClo1 (*Arabidopsis thaliana* Caleosin1*, At4g26740), AtClo2 (*Arabidopsis thaliana* Caleosin2*, At5g55240), ATS3 (embryo-specific protein 3, At5g07190); D, 21.2 kDa oleosin (S4*, At5g40420), 20.3 kDa oleosin (S2*, At3g27660), 19.7 kDa oleosin (S1*; At3g01570), AWPM-19-like protein (At1g04560); E, 18.5 kDa oleosin (S3*, At4g25140), 18.4 kDa oleosin (SM1*, At1g48990), 18.1 kDa oleosin (SM2*, At3g18570), and 14.8 kDa oleosin (S5*, At5g51210). [Proteins denoted with an asterisk (*) indicate nomenclature used by Kim et al.[18] for oleosins, Naested et al.[19] for caleosins, and D'Andrea et al.[20] for steroleosins.]

OBs as residents, but that they may traffic to this organelle as a consequence of the mechanisms underlying their biogenesis or as a consequence of OB isolation from seeds. Nonetheless, in view of the possible biomedical applications of OBs, in our opinion, all these proteins must be considered as real components of OB preparations, because of their consistent presence in purified samples. However, for ATS3, we suggest an organelle-specific role in OB formation.

Some of the proteins that were previously described as OB residents by in-gel MS analysis[26] have not been identified as such by the described in-solution approach. This is true for the predicted GPI-anchored protein (At1g54870), which was not detected in any samples. Equally, the caleosin AtClo1 (At4g26740) and the $\beta$-tonoplast intrinsic protein (At1g17810) were identified but their presence was not statistically validated, because the detection occurred in a subset of samples. It is possible that the bulk presence of other proteins such as seed storage proteins may mask the detection of low-abundance proteins and hamper their detection when an in-solution LC-MS/MS approach is applied. Alternatively, the variability in their detection may reflect differences in organelle purification.

**Comparative Analysis of wt OB Protein Content Using In-Solution and In-Gel LC-MS/MS.** In order to reduce the complexity of the sample and enable a better detection of low-abundance OB proteins (and, hence, broaden the organelle protein catalog), samples were also analyzed using an in-gel based MS approach. A comparison between the results obtained using in-solution and in-gel methods was performed to evaluate their potentials in OB proteome characterization.

Wild-type OB protein enrichment was monitored across the four steps of the purification procedure by SDS-PAGE. Coomassie staining of the gel showed that the purification leads to enrichment of the proteins migrating with an apparent molecular weight between 14 and 22 kDa and a reduction of those migrating between 21 and 50 kDa. This occurred mainly after the final 8 M urea treatment (see Figure 3), which is very effective at removing most of the proteins present in the total seed extract without affecting the integrity of the OBs.[10] In the final purified OB sample, five main bands were present. However, in order to obtain a comprehensive description of all of the proteins present in OB preparations, the LC-MS/MS analysis was carried out on 1-mm gel slices excised across the entirity of lane 4 (purified OBs) for a total of 13 slices. This approach resulted in the identification of 49 proteins that were present at concentrations lower than the detection limit for Coomassie staining (see Table S-8 in the Supporting Information). Several new proteins were identified by this approach, as well as confirmation of the previously described oleosins and caleosins, with structural function, and steroleosins with enzymatic activity.[2,26] ChlADR and the Saposin-like aspartyl protease family protein were also identified as by the in-solution approach.

Among the proteins newly identified in OB preparation by the in-gel approach, some were already reported as OB residents from previous studies or by sequence homology with known OB proteins. These proteins included two low-abundance oleosin isoforms expressed both in seed and in the floral tapetum (18.1 kDa, At3g18570; 18.4 kDa, At1g48990)[18] and the hydroxysteroid dehydrogenase 5 (AtHSD5, At4g10020) (sharing 60% sequence identity with the OB resident 11-$\beta$-hydroxysteroid dehydrogenase-like protein). The presence of the 14.8 kDa oleosin isoform, ATS3, AtClo2 and AWPM-19 recently identified as associated with OBs via MS after a combined chymotrypsin/trypsin digestion approach by Vermachova[31] was also confirmed.

Most of the remaining identified proteins according to the GO annotations were assigned to membranes, the endomembrane system, and ribosomes (see Table S-8 in the Supporting Information). Again, their detection could occur because the purification protocol favors the removal of masking-soluble proteins, facilitating the detection of those with hydrophobic domains that are more difficult to remove under these conditions. A single digestion approach was sufficient to obtain these results.

## CONCLUSIONS

The absolute quantification of a heterologous polypeptide targeted to the surface of oil bodies (OBs) was performed in this study, using a mass spectroscopy (MS)-based method that combined multiple reaction monitoring (MRM) with isotope dilution. The results obtained suggest that the amount of chimeric oleosin present in OB preparations from transgenic seeds was only 0.5% of total proteins and, as a consequence, alternative strategies must be considered to promote greater accumulation of heterologous proteins on the OB surface. For example, the use of mutant plants with ablated endogenous oleosin genes (such as described by Siloto[32]) might facilitate increased accumulation of the chimeric form of the oleosin.

The data obtained by the in-solution liquid chromatography (LC)-MS/MS method did not reveal any significant differences in the accumulation of OB proteins in transgenic seeds. LC-MS/MS analysis allowed the identification of novel putative OB residents. The in-gel analysis carried out across the entire lane separating purified OB proteins showed a higher sensitivity, compared to the in-solution approach. In fact, this strategy reduced the complexity of the samples and enabled the identification of

low-abundance proteins. Overall, the results obtained indicate that the use of MS-based techniques to characterize plant OB composition provides unparalleled insights into the proteome of this organelle, with specific relevance to the growing field of transgenic plant molecular farming.[33]

## ■ ASSOCIATED CONTENT

**Supporting Information.** This material is available free of charge via the Internet at http://pubs.acs.org.

## ■ AUTHOR INFORMATION

### Corresponding Author

*Tel.: +390630486612. Fax: +390630486545. E-mail: selene.baschieri@enea.it.

## ■ ACKNOWLEDGMENT

We would like to thank Dr. M. Deery and S. Hester for technical assistance. F.C. was a recipient of a Ph.D. fellowship granted by the EU FP6 "PHARMAPLANTA" Integrated Project No. LSHB-CT-2003-503565. Rothamsted Research receives grant-aided support from the BBSRC (UK). Data obtained from this study were deposited into the PRIDE database (PRoteomics IDEntifications, http://www.ebi.ac.uk/pride/),[21,22] under Accession Nos. 19642−19644.

## ■ REFERENCES

(1) Murphy, D. J. *Prog. Lipid Res.* **1991**, *29*, 299–324.
(2) Purkrtova, Z.; Jolivet, P.; Miquel, M.; Chardot, T. *C. R. Biol.* **2008**, *331*, 746–754.
(3) Roberts, N. J.; Scott, R. W.; Tzen, J. T. C. *Open Biotechnol. J.* **2008**, *2*, 13–21.
(4) Capuano, F.; Beaudoin, F. *Biotechnol. Adv.* **2007**, *25*, 203–206.
(5) Billaut-Mulot, O.; Idziorek, T. *Vaccine* **2001**, *19*, 2803–2811.
(6) Mayya, V.; Rezual, K. *Mol. Cell. Proteomics* **2006**, *5*, 1146–57.
(7) Brun, V.; Masselon, C. J. *Proteomics* **2009**, *72*, 740–749.
(8) Thoyts, P. J. E.; Millichip, M. I. *Plant Mol. Biol.* **1995**, *29*, 403–410.
(9) Clough, S. J.; Bent, A. F. *Plant J.* **1998**, *16*, 735–743.
(10) Tzen, J. T. C.; Peng, C. *J. Biochem.* **1997**, *121*, 762–768.
(11) Wang, W.; Vignani, R. *Anal. Biochem.* **2004**, *329*, 139–141.
(12) Laemmli, U. K. *Nature* **1970**, *227*, 680–685.
(13) Ishihama, Y.; Oda, Y. *Mol. Cell. Proteomics* **2005**, *4*, 1265–1272.
(14) Anderson, L.; Hunter, C. L. *Mol. Cell. Proteomics* **2006**, *5*, 573–588.
(15) Spackman, D. H.; Stein, W. H.; Moore, S. *Anal. Chem.* **1958**, *30*, 1190–1206.
(16) Matsui, M.; Kohyama, S.; Suda, T.; Yokoyama, S.; Mori, M.; Kobayashi, A.; Taneichi, M.; Uchida, T. *Biochem. Biophys. Res. Commun.* **2010**, *391*, 1494–1499.
(17) Ohnoa, S.; Kohyama, S.; Taneichi, M.; Moriya, O.; Hayashi, H.; Odad, H.; Morid, M.; Kobayashi, A.; Akatsuka, T.; Uchida, T.; Matsui, M. *Vaccine* **2009**, *27*, 3912–3920.
(18) Kim, H. U.; Hsieh, K. *J. Biol. Chem.* **2002**, *277*, 22677–22684.
(19) Naested, H.; Frandsen, G. I. *Plant Mol. Biol.* **2000**, *44*, 463–476.
(20) D'Andréa, S.; Jolivet, P. *J. Agric. Food Chem.* **2007**, *55*, 10008–10015.
(21) Vizcaino, J. A.; Cote, R.; Reisinger, F.; Barsnes, H.; Foster, J. M.; Rameseder, J.; Hermjakob, H.; Martens, L. *Nucleic Acids Res.* **2010**, *38*, D736–D742.
(22) Barsnes, H.; Vizcaino, J. A.; Eidhammer, I.; Martens, L. *Nat. Biotechnol.* **2009**, *27*, 598–599.
(23) Gentleman, R. C.; Carey, V. J. *Genome Biol.* **2004**, *5*, R80.
(24) Smyth, G. K. *Stat. Appl. Genet. Mol. Biol.* **2004**, *3*, Article3.
(25) Benjamini, Y.; Hochberg, Y. *J. R. Stat. Soc. B* **1995**, *57*, 289–300.
(26) Jolivet, P.; Roux, E. *Plant Physiol. Biochem.* **2004**, *42*, 501–509.
(27) Dortay, H.; Gruhn, N.; Pfeifer, A.; Schwerdtner, M.; Schmülling, T.; Heyl, A. *J. Proteome Res.* **2008**, *7*, 3649–3660.
(28) Stirk, W. A.; Gold, J. D.; Novák, O.; Strnad, M.; van Staden, J. *Plant Growth Regul.* **2005**, *47*, 1–7.
(29) Poxleitner, M.; Rogers, S. W. *Plant J.* **2006**, *47*, 917–933.
(30) Vitale, A.; Galili, G. *Plant Physiol.* **2001**, *125*, 115–118.
(31) Vermachova, M.; Purkrtova, Z.; Santrucek, J.; Jolivet, P.; Chardot, T.; Kodicek, M. *Proteomics* **2011**, *11*, 3430–3434.
(32) Siloto, R. M. P.; Findlay, K.; Lopez-Villalobos, A.; Yeung, E. C.; Nykiforuk, C. L.; Moloney, M. M. *Plant Cell* **2006**, *18*, 1961–1974.
(33) Boothe, J.; Nykiforuk, C.; Shen, Y.; Zaplachinski, S.; Szarka, S.; Kuhlman, P.; Murray, E.; Morck, D.; Moloney, M. M. *Plant Biotechnol. J.* **2010**, *8*, 588–606.