

Vegetable Proteomics: The Detection of Ole e 1 Isoallergens by Peptide Matching of MALDI MS/MS Spectra of Underivatized and Dansylated Glycopeptides

Anna Napoli,* Donatella Aiello, Leonardo Di Donna, Petros Moschidis, and Giovanni Sindona

Dipartimento di Chimica, Università della Calabria, Arcavacata di Rende, Italy

Received December 17, 2007

Ole e 1 (NCBI entry gi|14424429) is the major allergen of Oleaceae family. Multiple isoforms and variants are present in varying degrees of distribution. In this report, we present a new approach to the resolution of multiple forms of Ole e 1 from whole antigen extracts, based on a preliminary chemical fractionation procedure followed by MALDI MS and MS/MS measurements. The characterization of Ole e 1 isoallergens was accomplished through the identification of the amino acid sequence including the glycosylation site and the structure of the glycan moieties. The structure feature of the identified *Ole e 1.0102* (gi|2465127), *main olive allergen [Olea europaea]* (gi|13195753), *Major pollen allergen Ole e 1* (gi|33329740) and *Ole e 1c* (gi|1362131) is represented by the point mutation $K_{106} \rightarrow I$ and by the presence of a glycan moiety. Two other variants *Major pollen allergen (Allergen Ole e1) (Ole e 1)* (gi|14424429) and *Ole e 1.0103 protein [Olea europea]* (gi|2465129) were identified as nonglycosylated species. These results, partially in disagreement with Swiss-Prot annotation, were validated by matching the MALDI MS/MS spectra of the natural tryptic mixture with those obtained after deglycosylation.

Keywords: Olive pollen (*Olea europaea*) • allergens • Ole e 1 • MALDI mass spectrometry

Introduction

Ole e 1 (gi|14424429) is the major allergen of *Olea europaea* and one of the best-known allergens of the Oleaceae family. This is a pollen-specific protein which plays an important role in pollen hydration and/or germination.¹ The primary structure of Ole e 1 has been determined through amino acid² and cDNA sequencing.³ It is a polymorphic glycoprotein of 145 amino acid residues and shows two main variants differing from each other in their degree of glycosylation. Both glycosylated (High Molecular Weight, HMW) and nonglycosylated (Low Molecular Weight, LMW) variants are recognized by the sera of hypersensitive individuals.^{4,5} The high molecular weight form (HMW), according to Swiss-Prot annotation, displays Asn₁₁₁ as a potential N-glycosylation site, only.

The structure of the relevant N-glycans present in samples of Ole e 1 from Spanish Cultivars, obtained by matching HPAEC-PAD retention times and ¹H NMR frequencies, suggests that Ole e 1 mainly contains high mannose (Man₇GlcNAc₂) and xylose complex (GlcNAcMan₃XylGlcNAc₂) polysaccharides.⁶ The latter were also recognized in crude olive pollen extracts.⁷

The single peptide chain Ole e 1 is characterized by a considerable microheterogeneity in at least 37 out of the 145 amino acids position.⁸ This microheterogeneity arises from the presence of both genetic variants and post-translational modifications.^{3,9–11} However, the classical approaches for the

identification of seven, at least, isoallergenic forms of Ole e 1 failed in providing the resolution of each individual variant. Recent achievements of mass spectrometric methodologies¹² provide a rapid and sensitive tool for characterizing the primary structure of proteins and their post-translational modifications including glycosylation, which plays an important role in the biological properties of many proteins.¹³ The conventional method for proteome analysis involves two-dimensional gel (2D) as the final purification step¹⁴ followed by overnight in-gel enzyme digestion and mass spectrometric analysis.¹⁵ The limited separation capability of 2D gels to resolve certain classes of proteins has been well-documented,¹⁶ and this drawback becomes increasingly apparent when modified peptides are analyzed.

Since modern approaches, such as immunoprecipitation, protein complexes,¹⁷ and protein profiling experiments, lead to a protein mixture, recent trends in proteomics suggest the direct analysis of protein mixtures, with less emphasis on preliminary protein purification steps. For these reasons, we have developed simple procedures of chemical fractionation^{18,19} of the whole saline proteome extracts from natural matrices to obtain samples containing one family of proteins, only, and/or sets of chemically homogeneous proteins.¹⁸

From the allergonomics perspective, Ole e 1 shows some interesting challenges due to its multiple isoforms and variants present at different degree of distribution. In this report, we present the identification and characterization of multiple forms of Ole e 1 from whole saline proteome extracts of natural matrices using a chemical fractionation procedure followed by MALDI MS and MS/MS of tryptic digests.

* To whom correspondence should be addressed. Prof. Anna Napoli, Via P. Bucci, Cubo 12/C, Dipartimento di Chimica, Università della Calabria, I-87030 Arcavacata di Rende, Italy. Tel., + 39 0984 492852; fax, +39 0984 493307; e-mail, amc.napoli@unical.it.

Experimental Section

Materials. Olive tree pollen of cultivar Villacidro is part of the collection of the Istituto Sperimentale per l'Olivicoltura, C.da Li Rocchi, Rende (CS), Italy.

Solvents (CH_3OH , CH_3CN , CHCl_3 , and H_2O , HPLC grade), ammonium bicarbonate (NH_4HCO_3 , pure $\geq 99.5\%$), trifluoroacetic acid (HPLC grade, pure $\geq 99.0\%$), dithiothreitol (DTT, pure $\geq 99.0\%$), enzymes (trypsin and PNGase F), α -cyano-4-hydroxy-*trans*-cinnamic acid (α -CHCA, pure $\geq 99.0\%$) and 3,5-dimethoxy-4-hydroxycinnamic acid (pure $\geq 99.0\%$) were purchased from Sigma Aldrich Fluka (Milano, Italy).

Allergen Extracts and Digestion. Pure sample solutions of *Ole e 1* were obtained from 50 mg portions of pollen grains from Villacidro cultivar, extracted and treated as reported.¹⁸

Fifty microliters of the pure *Ole e 1* solution was fully reduced at pH 8 by treatment with 10 μL of DTT (50 mM) under magnetic stirring, for 1 h at 57 °C. The sample was then incubated with 4 μL of PNGase F (0.5 unit/ μL) for 3 days at 37 °C. After 78 h, the resulting mixture was digested overnight with 1 μL (4 pmol/ μL) of trypsin.

In a different run, the *Ole e 1* sample was digested overnight with 1 μL (4 pmol/ μL) of trypsin without preventive deglycosylation.

Peptides Mixture Derivatization. Twenty microliters of tryptic digested *Ole e 1* sample (pH 8) was mixed with 16 μL of a Dansyl-Cl solution (0.22 mol/L, in CH_3COCH_3), 10 μL of NaHCO_3 (0.1 M) and 5 μL of saturated Na_2CO_3 . The solution was stirred for 3 h at room temperature, the corresponding dansylated peptides were obtained, and the resulting solution was directly analyzed by MALDI MS and MS/MS.

Protein Identification by MALDI MS and MS/MS Analysis. A 1 μL portion of a premixed solution of each sample and α -CHCA (0.3% in TFA) was spotted on the matrix target, dried at room temperature and directly analyzed by MALDI mass spectrometry.

MALDI-TOF analyses were performed using a 4700 Proteomics Analyzer mass spectrometer, from Applied Biosystems (Foster City, CA), equipped with a 200-Hz Nd:YAG laser at 355-nm wavelength. The MS spectra were acquired in reflectron mode (20-keV accelerating voltage), with 400-ns delayed extraction, averaging 2000 laser shots with a mass accuracy of 20 ppm.

MS/MS experiments were performed at a collision energy of 1–2 kV, defined by the potential difference between the source acceleration voltage (8 kV) and the floating collision cell (7–6 kV); 3000 laser shots were averaged, while the pressure inside the collision cell was 8×10^{-7} Torr.

Linear MALDI MS spectra were acquired averaging 2500 laser shots with a mass accuracy of 500 ppm in default calibration mode that was performed using the following set of standards: insulin (bovine, $[\text{M} + \text{H}]^+_{\text{avg}} = 5734.59$ Da), apomyoglobin (horse, $[\text{M} + \text{H}]^{2+}_{\text{avg}} = 8476.78$ Da, $[\text{M} + \text{H}]^+_{\text{avg}} = 16952.56$ Da), and thioredoxin (*Escherichia coli*, $[\text{M} + \text{H}]^+_{\text{avg}} = 11674.48$ Da).

Proteins were identified by searching a comprehensive protein database using Mascot programs (www.matrixscience.com). Peak harvesting was done automatically using Data Explorer software (Applied Biosystems). Peak resolution was calculated using the Data Explorer software. The query was made for “Other Green Plants” taxonomy allowing maximum mass tolerance of 50 ppm and at most two missed cleavages for tryptic peptides. MS/MS interrogation were carried out, with the same parameters as previously described for the PMF

research. Peptide sequences were automatically identified by database searching of the MS/MS spectra against the NCBI database using the MASCOT search program (http://www.matrixscience.com). However, all spectra were manually checked to verify the validity of the MASCOT results. The mass tolerance of the fragments for MS/MS data search was set at ± 0.25 Da. All MS/MS spectra of dansylated and underivatized glycopeptides were manually processed.

The program GlycoMod (www.expasy.ch/tools/glycomod) was used to predict all possible composition of glycan structures from *Ole e 1* Swiss-Prot/TrEMBL ID imputing glycopeptide determined monoisotopic masses.

Results and Discussion

The high solubility in aqueous media is an important prerequisite for allergen candidates. Their allergenicity is linked either to their concentration or to a rapid release when airborne.²⁰ *Ole e 1* is the major allergen of *O. europaea* and it is characterized by a very high solubility in saline solution.²¹ According to the known hydrophilic and acidic nature of these pollen antigens, a fraction of *Ole e 1* family was obtained by chemical fractionation of the whole Villacidro's proteome extract.¹⁸

The electrophoretic profile resulting from this sample shows only two protein bands, within 14–20 kDa (Figure 1a, lane 1) corresponding to the LMW and HMW forms of *Ole e 1*, respectively. The protein bands were identified by their molecular weights, in agreement with the standard protein markers (Figure 1a, lane M).

The protein content of the sample was monitored by linear MALDI MS. Accordingly, the *Ole e 1* enriched sample was determined by appearance of the ion species of m/z 8946 and 17 811 corresponding to doubly and monocharged forms of the expected antigen, respectively, whereas those of m/z 8237 and 16448 represent probably the double and monocharged low molecular weight form of the same allergen (Figure 1b). The protein profiling displayed by linear MALDI experiments match that obtained by SDS-PAGE chromatography (Figure 1a). The obvious difference between the two methods is represented by the specificity of the mass spectrometric method which allows a preliminary identification of the protein pool through the determination of the average molecular weight of each compound.

The goal of this study was not only the determination *Ole e 1* microheterogeneity from natural matrices, but also the evaluation of the glycosylated isoallergens as well as the characterization of the glycans. Therefore, a preliminary detection of the corresponding glycopeptides followed by the analysis of the carbohydrate moiety was necessary. The glycopeptides can be analyzed by mass spectrometry without cleanup procedures. Peptide backbones generally have higher ionization efficiency than carbohydrates, producing good mass spectral signals and favoring accurate mass measurements. In addition, proteolytic fragments on either side of the amino acids carrying the carbohydrate can be used to reveal the specific site from which the carbohydrate originated.²² The conventional approach to establish the isoform-specific and specific glycosylation pattern requires the separation of the isoforms by HPLC or 2D gel and glycopeptide analysis on the digests of the isolate isoforms.^{23,24}

The chemical fractionation procedure was exploited because of the hydrophilic and acidic nature of the expected antigen which could limit the separation capability of 2D gels. The

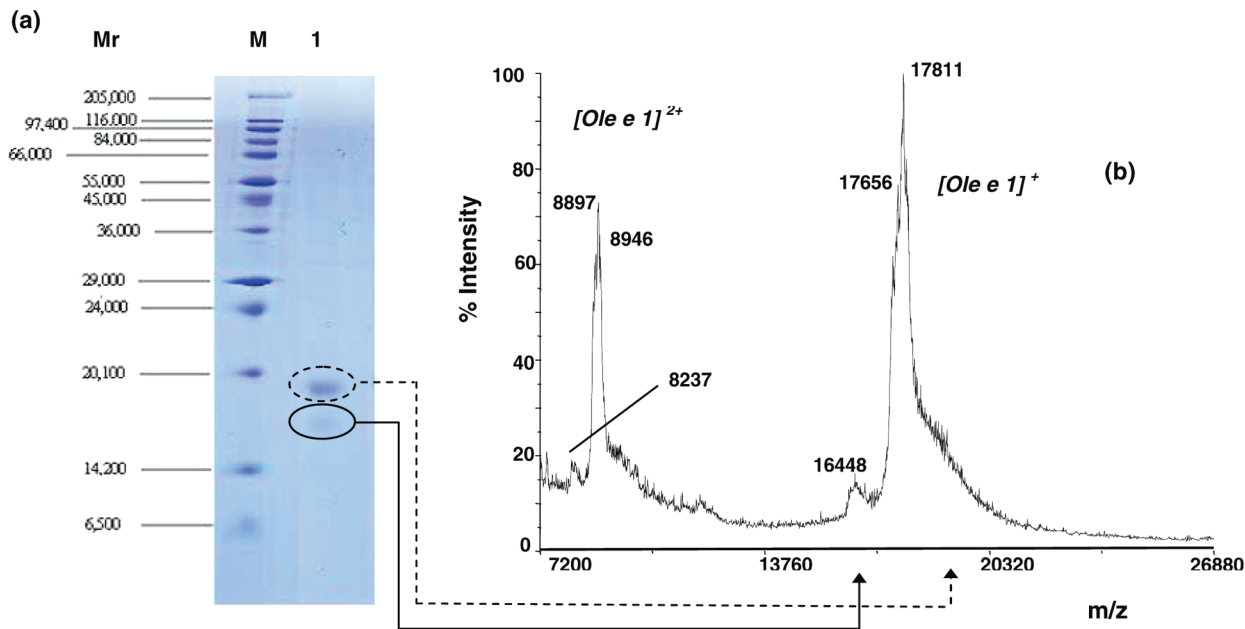


Figure 1. (a) SDS-PAGE (lane M, marker; lane 1, Ole e 1 pure fraction); (b) linear MALDI spectrum of Ole e 1 chemical fraction.

Table 1

	NCBI entry	mass (Da)	score	queries matched	protein name	matched ions (Da)
1	gil13195753	14575	94	7	Main olive allergen [<i>O. europaea</i>]	1793.92; 1600.80; 1472.72; 1267.64; 1787.91; 1235.67; 1022.55.
2	gil33329744	14715	75	6	Major pollen allergen Ole e 1 [<i>O. europaea</i>]	1793.92; 1600.80; 1267.64; 1787.91; 1235.67; 1022.55.
3	gil33329730	14831	75	6	Major pollen allergen Ole e 1 [<i>O. europaea</i>]	1793.92; 1600.80; 1267.64; 1787.91; 1235.67; 1022.55.
4	gil33329734	14904	75	6	Major pollen allergen Ole e 1 [<i>O. europaea</i>]	1793.92; 1600.80; 1267.64; 1787.91; 1235.67; 1022.55.
5	gil37548749	14801	75	6	Major pollen allergen Ole e 1 [<i>O. europaea</i>]	1793.92; 1600.80; 1267.64; 1787.91; 1235.67; 1022.55.
6	gil145313990	15793	75	6	Major pollen allergen Ole e 1 [<i>Olea europaea</i>]	1793.92; 1600.80; 1472.71; 1267.64; 1787.91; 1235.67.
7	gil145313984	15851	75	6	Major pollen allergen Ole e 1 [<i>O. europaea</i>]	1793.92; 1600.80; 1472.71; 1267.64; 1787.91; 1235.67.
8	gil33323443	14873	73	6	Major pollen allergen Ole e 1 [<i>O. europaea</i>]	1793.92; 1600.80; 1267.64; 1787.91; 1235.67; 1022.55.
9	gil33325111	14847	73	6	Major pollen allergen Ole e 1 [<i>O. europaea</i>]	1793.92; 1600.80; 1267.64; 1787.91; 1235.67; 1022.55.

chemical fractionation procedure can be considered similar to a 1D liquid phase separation procedure, which resolves the proteins by their hydrophilic/hydrophobic properties and leads to fractions of chemically homogeneous species.¹⁸ MS/MS experiments offer the second dimension, by analogy with 2D gel, but with the difference of providing the high specificity of the mass spectrometric method; therefore, the initially obtained Ole e 1 sample was digested with trypsin and subjected to MALDI MS and MS/MS. The mass data were processed to assign candidate peptides in the NCBI database using the MASCOT search program. Peptide Mass Fingerprinting (PMF) interrogation unambiguously identified the protein *main olive allergen* [*Olea europaea*] (NCBI entry: gil13195753) using seven *m/z* values corresponding to nine possible sequences of significant score (score >66, Table 1).

The potential site of post-translational modification, located at the C-terminal portion of the protein and of the corresponding tryptic peptides, should be detected, in reflectron ion mode, using sinapinic acid as the matrix. The simple MALDI MS spectrum of the tryptic mixture showed, in fact, the specific

m/z spacing patterns (n203 and n162 + 203u) of glycoforms within the 2–3 kDa mass range (Figure 2).

The absence of larger tryptic glycopeptides and the digestion efficiency was checked in linear mode during 18 h.

A closer inspection of the characteristic patterns displayed in Figure 2 provided some structural information. The presence of three sets of glycoforms and several extra peaks arising from the loss of one or two hexose residues were observed. The fragmentation of high mass glycopeptides affects the mass spectra and makes more difficult the evaluation of the natural glycanic microheterogeneity. The glycopeptide of *m/z* 3182.64, could correspond to a diantennary N-glycan because of the appearance in the spectrum of the peaks at *m/z* 2655.10 and at *m/z* 2128.00 likely arising from consecutive Hex₂-HexNAc losses (Figure 2). The ion peak of *m/z* 2776.18 can be formed by formal loss of two HexNAc units from the glycopeptide of *m/z* 3182.64 or it might correspond to an intact tryptic glycopeptide.

Glycoprotein N-glycans obtained from plants can either be of the high-mannose- and/or the complex-type. High-man-

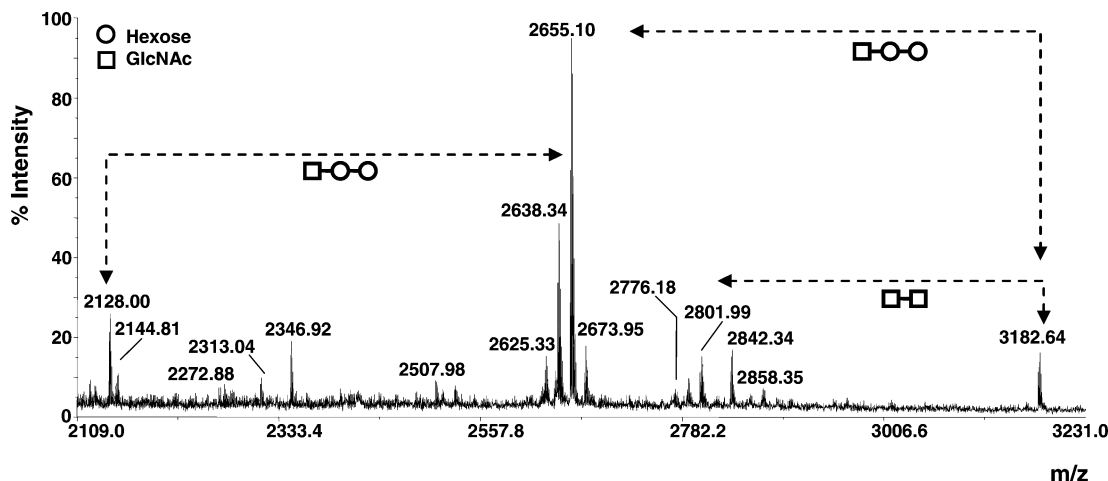


Figure 2. MALDI MS spectrum of Ole e 1 tryptic digest.

nose-N-glycans are usually referred to with the general structure $\text{Man}_{5-9}\text{GlcNAc}_2$. Carbohydrates having terminal GlcNAc residues or larger antennae $\beta(1, 2)$ linked to the $\alpha(1,6)$ - or $\alpha(1,3)$ -mannose constitutive of $(\text{Man}_3\text{-2GlcNAc}_2)$ core have been referred to as complex-type N-glycans. The conservative core could take an extra $\alpha(1, 3)$ -fucose and/or $\beta(1, 2)$ -xylose residue.²⁵

Structural information might be obtained by inputting data derived from preliminary evaluation of the most likely constituents present in the examined glycopeptides and from literature data²⁵ into the freely available software program GlycoMod (www.expasy.ch/tools/glycomod). The experimentally determined mass of glycopeptide is used to find all possible compositions of the glycan moiety when the glycosylation site and protein structure are known. The program works on the basis of the congruence of their molecular masses, only.

Hybrid complex and high mannose structure type were assigned to the observed glycopeptides on the base of the tryptic cleavage site K^{106} (P19963, gil14424429) that give rise to LNTVNGTTR (missed cleavage 0) instead of FKLNTVNGTTR (missed cleavage 1) peptides. The composition of the occurring oligosaccharides was calculated by inputting the presence of at least three Hexoses (Hex) and 2 *N*-acetylhexoses (HexNAc), representing the conservative core. Accordingly, the oligosaccharide mixture should correspond to the list reported in Table 2.

The glycan moiety of the peptide of m/z 3182.64 should therefore be represented by the hybrid complex $(\text{Hex})_3$ $(\text{HexNAc})_2$ $(\text{Deoxyhexose})_2$ $(\text{Pent})_1 + (\text{Man})_3(\text{GlcNAc})_2$ (calculated MW: 2208.79), thus, providing a theoretical molecular weight for the given glycopeptide differing by -524 ppm from that experimentally observed (Table 2, entry 1).

It was therefore necessary to exploit the MS/MS approach to validate both the carbohydrate composition suggested by the Glycomod program previously used and the amino acid sequence of glycopeptide. Although the number of fragments obtained from the precursor ion of m/z 3182.64 is low, direct cross ring fragmentation leads to the release of $[(\text{HexNAc})_2(\text{Hex})_2+43]$ and the consecutive loss of two hexose residues (data not shown). This behavior clearly suggests that the hybrid complex structure provided by Glycomod does not match the observed fragmentation pathway. Nevertheless, a biantennary glycan structure is supported by both theoretical and experimental data. No information was

obtained about the peptide itself $[\text{P}+\text{H}]^+$ and the related ionic species $[\text{P} + \text{H} - 17]^+ / [\text{P} + \text{GlcNAc} + \text{H}]^+ / [\text{P} + \text{H} + 83]^+$.

This uncertainty could be solved by N-terminal derivatization of the tryptic mixture by dansyl chloride. The latter is known for its capability to improve either the ionization efficiency or the amount of structural information.²⁶ Moreover, the number of N-terminal driven fragments should be enhanced on going from dansylated to underivatized peptides. An increasing interest in the use of this peptide-sequence directing group has been recently observed as a consequence of its ability to improve the score as well as the sequence coverage, when a PMF database search is performed.²⁷ Dansylation increases the m/z value of each individual peptide and enlarges the mass-scale of the experiment. Therefore, the increased ionization efficiency and the better signal-to-noise ratio thus achieved, allows the identification, by MALDI, of a number of peptides greater than when the same underivatized tryptic mixture is sampled.

A few microliters of the tryptic mixture were treated with commercially available Dns-Chloride (see Experimental Section) and directly analyzed by the MALDI TOF/TOF workstation in single MS mode. Examination by PMF of the spectrum thus obtained from the dansylated mixture (Figure 3) allows the identification of Ole e 1 using nine masses corresponding to 13 possible sequences (Table 3).

The mass spectrum (Figure 3) shows more peaks than that of underivatized peptides. Peaks arise from peptides with one or two dansyl modifications, whereas unmodified peptides were not observed showing that each tryptic peptide has taken up at least one dansyl group. N-terminal dansylation increases desorption and ionization efficiency in α -CHCA, as previously observed.²⁷

Dansylated peptide peaks present in MS spectrum (Figure 3) showing a glycosidic pattern were selected as the ion precursor. One KeV CID spectra of derivatized peptides afforded, in all cases, three main ion peaks of m/z 1469.14, 1553.73 and 1672.72. The first corresponds to $[\text{M}_{\text{pep}} + \text{H}]^+$, while 1553.73 and 1672.72 arise from cleavage at the innermost *N*-acetylglucosamine residues $[\text{M}_{\text{pep}} + 83 + \text{H}]^+$ and from Y-type cleavage of the chitobiose core $[\text{M}_{\text{pep}} + 203 + \text{H}]^+$ with the retention of the peptide moiety. The latter suggested that all of the glycopeptides investigated do not contain a fucosylated core.

Table 2. Composition of Oligosaccharides Calculated by Glycomod^a

	glycan mass (Da)	Δ mass ^a (ppm)	structure	type	peptide mass [M] (calc)	peptide sequence	glycopeptide mass (calc)
1	2208.79	−524.81	(Hex) ₃ (HexNAc) ₂ (Deoxyhexose) ₂ (Pent) ₁ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	974.51	LNTVNGTTR	3184.31
2	1866.69	45.05	(Hex) ₁ (HexNAc) ₄ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	974.51	LNTVNGTTR	2842.21
3	1663.61	−300.25	(Hex) ₁ (HexNAc) ₃ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	974.51	LNTVNGTTR	2639.13
4	1930.69	398.06	(Hex) ₃ (HexNAc) ₂ (Deoxyhexose) ₁ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	1249.68	FKLNTVNGTTR (Miss 1)	3181.38
5	1606.59	376.99	(Hex) ₁ (HexNAc) ₂ (Deoxyhexose) ₁ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	1249.68	FKLNTVNGTTR (Miss 1)	2857.27
6	1590.59	373.17	(HexNAc) ₂ (Deoxyhexose) ₂ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	1249.68	FKLNTVNGTTR (Miss 1)	2841.28
7	1592.57	−323.14	(Hex) ₁ (HexNAc) ₂ (Pent) ₁ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	1249.68	FKLNTVNGTTR (Miss 1)	2843.26
8	1524.53	347.63	(Hex) ₃ (Deoxyhexose) ₁ + (Man) ₃ (GlcNAc) ₂	high_man	1249.68	FKLNTVNGTTR (Miss 1)	2775.22
9	1403.51	341.45	(Hex) ₁ (HexNAc) ₁ (Deoxyhexose) ₁ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	1249.68	FKLNTVNGTTR (Miss 1)	2654.19
10	1387.51	53.24	(HexNAc) ₁ (Deoxyhexose) ₂ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	1249.68	FKLNTVNGTTR (Miss 1)	2638.20
11	1373.50	436.28	(HexNAc) ₁ (Deoxyhexose) ₁ (Pent) ₁ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	1249.68	FKLNTVNGTTR (Miss 1)	2624.18

^a Calculated vs measured glycopeptide mass.

The assignment of the ions indicates that the molecular mass of the deglycosylated peptide is 1235.47 Da ($M_{\text{pep}} + H - \text{Dns}$), which is higher than that of the expected tryptic peptide L₁₀₇–R₁₁₅ (Ole e 1; P19963 or gil14424429; missed cleavage 0; 975.62 Da). The observed mass increment could correspond to the formation of the superior homologue F₁₀₅–R₁₁₅ (1250.68 Da) taking into account the partial cleavage of the substrate. On the other hand, the peptide F₁₀₅–R₁₁₅ is 15.21 Da higher than 1235.47 Da ($M_{\text{pep}} + H - \text{Dns}$), therefore the presence of a point mutation can be hypothesized at K₁₀₆. It was consequently necessary to exploit the 2 KeV CID of all dansylated-glycopeptides to validated the amino acid composition. The high energy CID spectrum of one dansylated-glycopeptide is shown in Figure 4 as a representative example for all the investigated glycopeptides.

The fragmentation of the peptide chain and the absence of the glycan fragment ions are typical spectral features, common to the other species reacting under theses experimental conditions. On the basis of this fragment ion analysis, the deduced sequence for the N-derivatized peptide should be Dns-FILNTVNGTTR (Theoretical Mass 1468.72 Da). The latter is likely the structure of the peak of m/z 1468.69 detected within −20 ppm error in the MALDI stable ion spectrum and arising from the release of the neutral carbohydrate moiety (Figure 3). Fragments of the proposed sequence match the ion peaks visible in the MS/MS spectrum (Figure 5).

Direct submission of MS/MS data (Figure 5) to the NCBI database for protein identification against other green plants allowed the identification of the isoallergens Ole e 1.0102 (gil2465127), Major pollen allergen Ole e 1 (gil33329740) and main olive allergen [*O. europaea*] (gil13195753). The latter are

the three known isoallergens of Ole e 1 containing a point mutation K₁₀₆ → I.

Hybrid complex and high mannose structure type were assigned to the observed glycopeptides on the base of the extensively tryptic cleavage site K₁₀₄ giving rise to the peptide FILNTVNGTTR (missed cleavage 0) (Table 4). The structure **1** (Table 4), now easily confirms the consecutive neutral loss of 527 (Hex₂-HexNAc) and 1054 (2 Hex₂-HexNAc) (Figure 2).

Whereas **2**, **3**, **5** and **6** (Table 4) probably arise from **1** by consecutive losses of Hex and HexNAc residues during the ionization stage, they could represent wild-type “truncated structures”.²⁸

The high energy CID spectrum of one glycopeptide is shown in Figure 6 as a representative example of all the investigated glycopeptides. It showed fragments originating on either side of the glycosylation site and fragments belonging mainly to the cross ring cleavages of the glycan moieties. The latter correspond to a rearrangement processes typical of metastable ion dissociations not easily observed in MS/MS experiments.²⁹ On the basis of this fragment ion, analysis the N-linked biantennary glycan structure **5** (Table 4) is confirmed as the tryptic segment of m/z 2655.10 (Figure 6). A very similar structural feature was observed for glycopeptides. The structure of the glycoforms **4** and **8** (Table 4) is in agreement with that published by Van Ree et al.,⁶ while the validated structure **7** (Table 4) may be considered as a higher analogue of the well-known glycoform **8** (Table 4).

The amino acid sequences of some Ole e 1 variants showing potential site of post-translational modification are displayed in Figure 7. The sequences 38–51 and 83–99 are responsible for the high microheterogeneity of the nine deduced cDNA

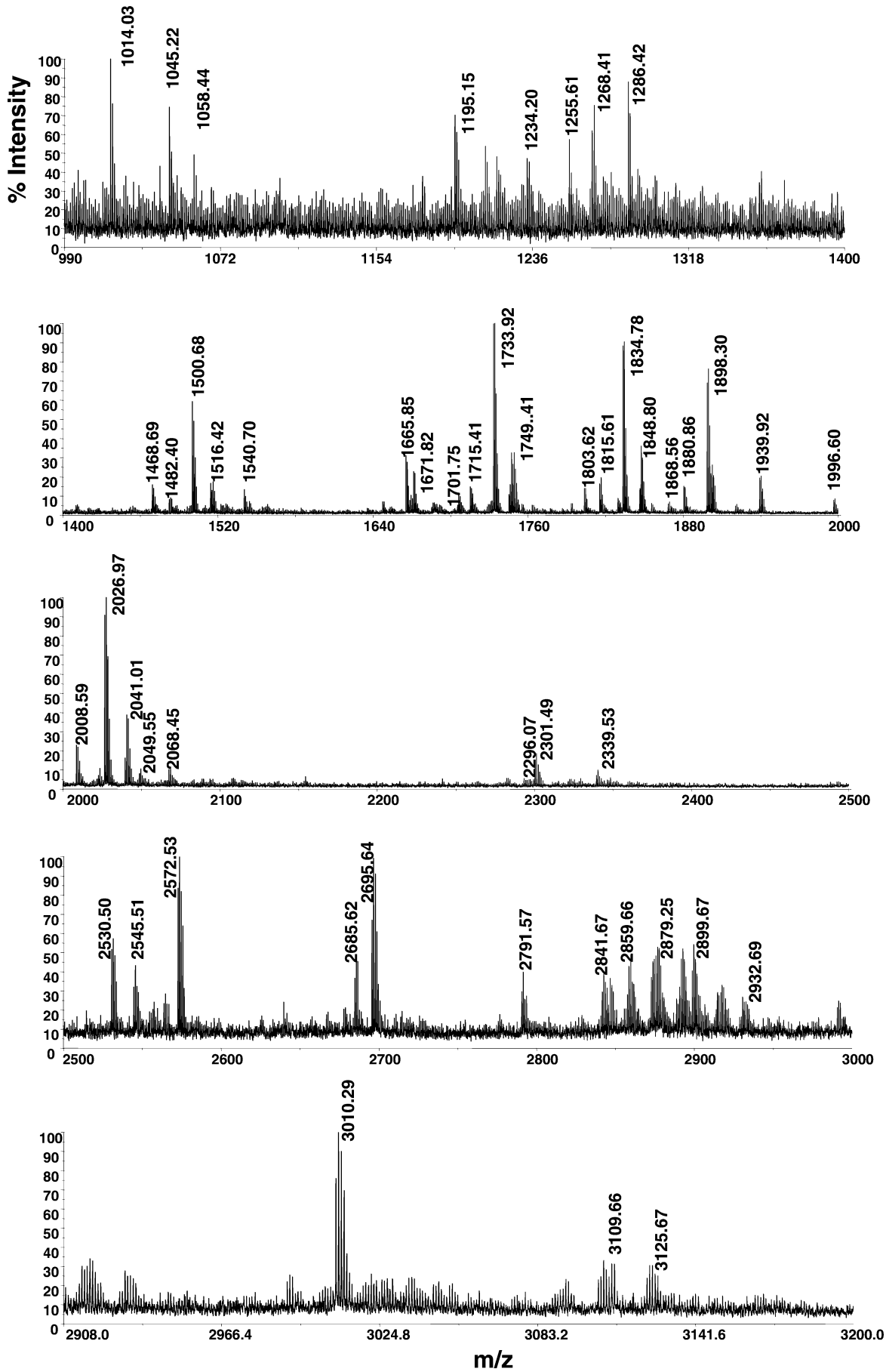


Figure 3. MALDI MS spectrum of Ole e 1 dansylated peptides mixture.

Table 3. PMF Identification of Ole e 1 Using Dansylated Tryptic Mixture

	accession number	score	queries matched	protein name	matched ion (Da)
1	gil2465131	148	9	Ole e 1 protein [<i>O. europaea</i>]	3010.29; 2879.25; 1898.30; 1848.80; 1500.68; 1701.75; 1468.69; 1255.61; 1058.44.
2	gil2465127	129	8	Ole e 1.0102 protein [<i>O. europea</i>]	3010.29; 2879.25; 2026.97; 1834.78; 1500.68; 1701.75; 1468.69; 1058.44.
3	gil2465129	111	7	Ole e 1.0103 protein [<i>O. europea</i>]	3010.29; 2879.25; 2026.97; 1834.78; 1500.68; 1701.75; 1058.44.
4	gil14424429	102	7	Major pollen allergen (Allergen Ole e1) (Ole e I)	2893.27; 2296.07; 2041.01; 1255.61; 1500.68; 1701.75; 1058.44.
5	gil33329750, gil33329744, gil33329730, gil37548749, gil33329734	101	7	Major pollen allergen Ole e 1 [<i>O. europaea</i>]	2026.97; 1834.78; 1500.68; 1701.75; 1468.69; 1255.61; 1058.44
6	gil33329742, gil145313982, gil33323443, gil33325111	100	7	Major pollen allergen Ole e 1 [<i>O. europaea</i>]	2026.97; 1834.78; 1500.68; 1701.75; 1468.69; 1255.61; 1058.44

forms of Ole e 1.³⁰ These regions are the most hydrophilic and antigenic areas of Ole e 1.³¹

The identification of the post-translational modification of Ole e 1 and of the specific amino acid residues was performed by the evaluation of underivatized and dansylated large peptides, while the microheterogeneity can be obtained from the analysis of tryptic peptides after the removal of N-glycans.

In an attempt to simplify the spectral data and obtain some separation of potential sites of point mutation, the Ole e 1 sample was submitted to trypsin after N-deglycosylation by PNGase F. The tryptic digest was examined in MS reflectron mode, whereas MS/MS experiments were used for peptide sequence discrimination. The partial MS spectrum, shown in Figure 8, displays ion peaks corresponding to protonated tryptic peptides of Ole e 1 easily detected with mass errors of 20 ppm. The interrogation of the PMF (MASCOT search program, NCBI

database) unambiguously identified Ole e 1 using 9 masses corresponding to five possible sequences characterized by significant protein scores (score >80) (Table 5).

Protein identification by PMF is considered a rapid and useful method to identify proteins previously separated by 2D gels. The adopted criteria for protein discrimination of variants is represented by the number of peptides matching the conservative domain, *Major pollen allergen (Allergen Ole e 1) (Ole e I)* (Table 5, entry3; gi[14424429], P19963) was in fact identified, matching 6 peptides, with a score of 92 and 50% sequence coverage (Table 6).

A difference of 3.97 mass units between the peaks at m/z 1665.85 and 1669.82 associated to a more complex isotopic pattern of peaks suggests that the Ole e 1 sample contains a mixture of at least two components (Figure 8). Among the published Ole e 1 sequences, only two variants, differing by

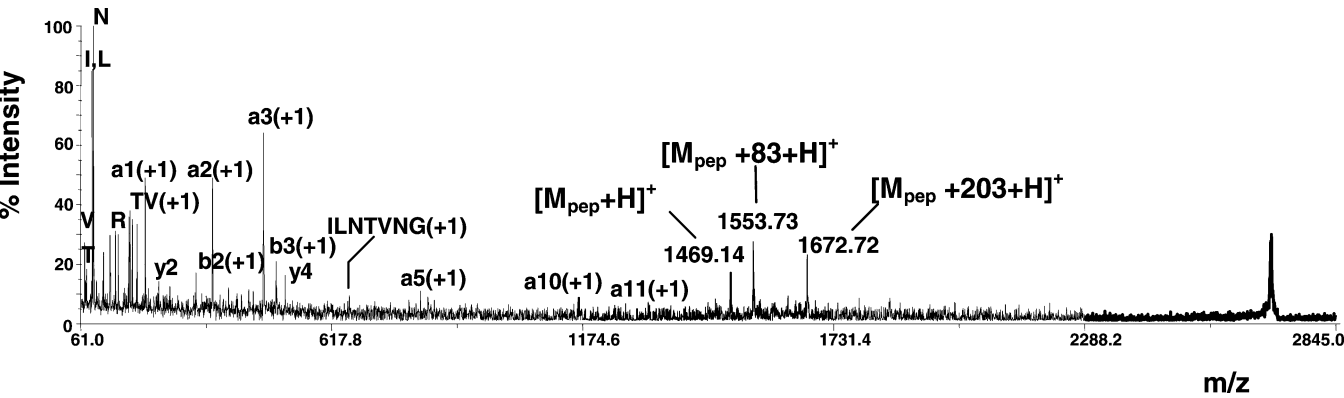
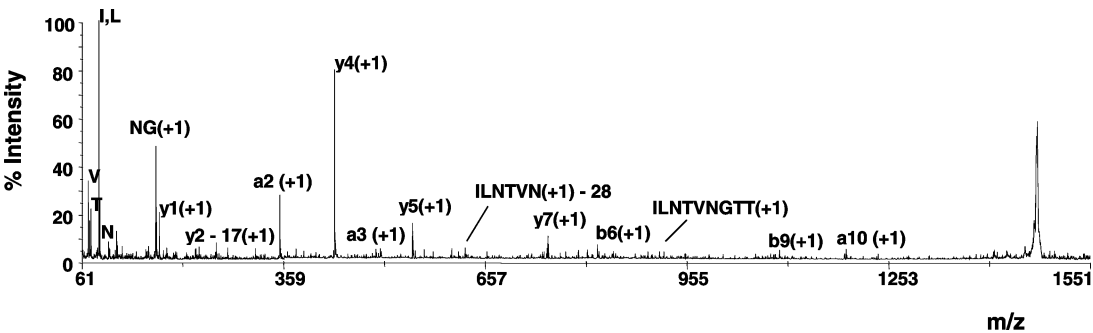

Figure 4. The 2KeV CID spectrum of glycopeptides at m/z 2695.64.

Figure 5. CID spectrum of the sequence DnsFILNTVNGTTR (m/z 1468.69).

Table 4. Composition of Oligosaccharides Determined by MS/MS Measurements

	glycoform mass (calc, Da)	structure	type	FILNTVNGTTR, MW (calc, Da)	MW (calc, Da)	observed (Da)
1	1947.78	(Hex) ₄ (HexNAc) ₂ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	1234.67	3182.44	3182.64
2	1623.49	(Hex) ₂ (HexNAc) ₂ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	1234.67	2858.16	2858.35
3	1607.49	(Hex) ₁ (HexNAc) ₂ (Deoxyhexose) ₁ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	1234.67	2842.16	2842.34
4	1541.39	(Hex) ₄ + (Man) ₃ (GlcNAc) ₂	high_man	1234.67	2776.06	2776.18
5	1420.30	(Hex) ₂ (HexNAc) ₁ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	1234.67	2654.96	2655.10
6	1403.51	(Hex) ₁ (HexNAc) ₁ (Deoxyhexose) ₁ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	1234.67	2638.17	2638.34
7	1389.49	(Hex) ₁ (HexNAc) ₁ (Pent) ₁ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	1234.67	2625.17	2625.33
8	1227.44	(HexNAc) ₁ (Pent) ₁ + (Man) ₃ (GlcNAc) ₂	hybrid/complex	1234.67	2463.11	2463.27
9	892.32	(Hex) ₃ (HexNAc) ₂	-	1234.67	2127.99	2128.00

the exchange of two amino acids at the positions 31 (L → F) and 39 (S → G), respectively, lead to a mass difference of 3.97 u. The theoretical calculated isotopic distribution and the sum of the isotopic distributions of both peptides, containing the previously discussed point mutations, suggest that Ole e 1 (gil14424429, Figure 7) and Ole e 1c (gil1362131, Figure 7) variants exist in the examined mixture. The MS/MS experiments of these peptides validated the sequences ²⁷FITELSEFIPGASVR⁴¹ and ²⁶FITEFSEFIPGAGVR⁴⁰ belonging to Ole e 1 (gil14424429, Figure 7) and Ole e 1c (gil1362131, Figure 7), respectively (Supporting Information Figure 1 and 2).

The peptide mass of 1793.91 Da (Figure 8), identified by database search, and validated by MS/MS experiment as the peptide ²⁵AGFITELSEFIPGASVR⁴¹, matched two different variant of Ole e 1 (Supporting Information Figure 3). This sequence is contained in Ole e 1.0102 (gil 2465127, Figure 7) and Ole e 1.0103 (gil2465129, Figure 7).

A mass difference of 14.01 units between the adjacent peaks at *m/z* 1793.91 (²⁵AGFITELSEFIPGASVR⁴¹, gil 2465127) and 1807.92 (Figure 8) can be justified by the exchange of one amino acid at position 40 (L → V), indicating the presence of a new variant. The MS/MS data confirm the AGFITELSEFIP-

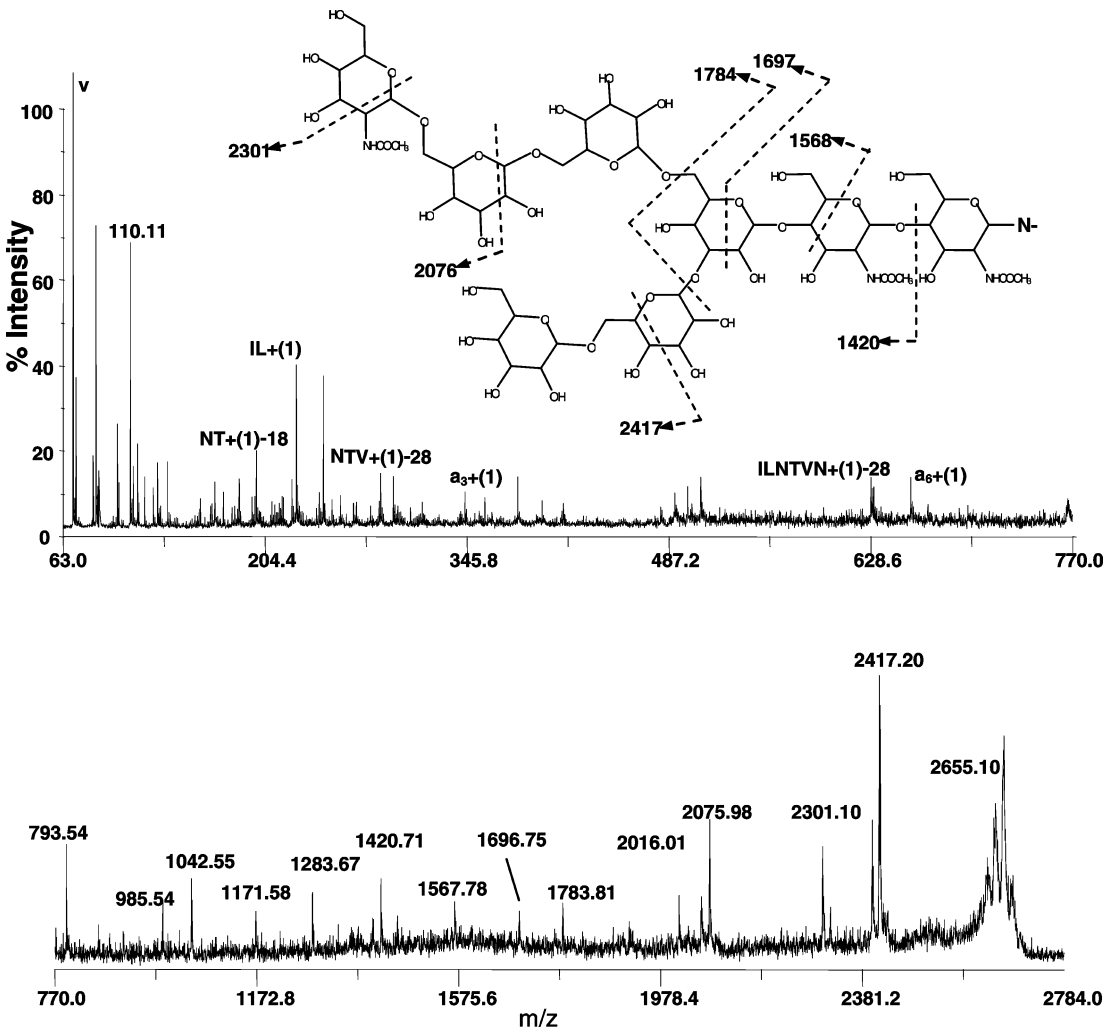


Figure 6. CID spectrum of ion species at *m/z* 2655.10.

gi 14424429	2	DIPQPPVSQFHIQQGVYCDTCRAGFITELSEFIPGASLRLQCKDKENGDTVTFTEVGYTRA	61
gi 2465129	3	.V.....V.....I.....	62
gi 2465127	3	.V.....V.....I.....	62
gi 13195753	2	E.K...I.....	61
gi 33329740	2	GV.....NI.....	61
gi 1362131	2	.V...I...YV.....TR...F.....GV.....G...KI.....	61
gi 14424429	62	EGLYSMLVERDHKNEFCBITLISSGRKDCNEIPTEGWAKPSLKFKLNTVNGTTRTVNPLG	121
gi 2465129	63	I.S.....	122
gi 2465127	63	I.....R	122
gi 13195753	62	S.....I.....	121
gi 33329740	62	D.....V.....I.....I.....	121
gi 1362131	62	I.....L.S...D.....V...V...I.....I.....	121
gi 14424429	122	FFKKEALPKCAQVYNKLGMYPPNM	145
gi 2465129	123	.Y.....	146
gi 2465127	123	.Y.....	146
gi 13195753	122		130
gi 33329740	122		129
gi 1362131	122	P..F.....	145
gi 2465129 Identities = 138/144 (95%), Positives = 143/144 (99%), Gaps = 0/144 (0%)			
gi 2465127 Identities = 138/144 (95%), Positives = 142/144 (98%), Gaps = 0/144 (0%)			
gi 13195753 Identities = 123/129 (95%), Positives = 127/129 (98%), Gaps = 0/129 (0%)			
gi 33329740 Identities = 119/128 (92%), Positives = 124/128 (96%), Gaps = 0/128 (0%)			
gi 1362131 Identities = 122/144 (84%), Positives = 134/144 (93%), Gaps = 0/144 (0%)			

Figure 7. The amino acid sequence of some variants of Ole e 1. The potential site of post-translational modification residue, N₁₁₁, is reported in bold character.

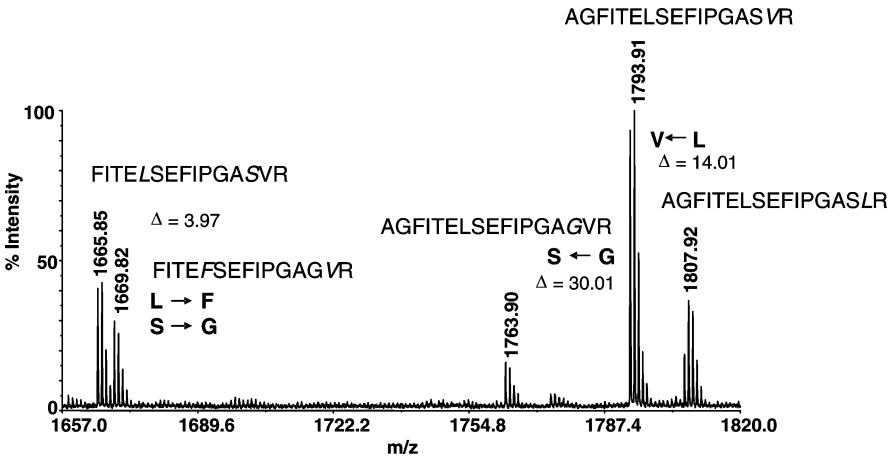


Figure 8. Portion of the MALDI spectrum of tryptic mixture after deglycosylation by PNGase F.

Table 5

	accession number	score	queries matched	protein name	matched ions (Da)
1	gil13195753	163	9	main olive allergen [<i>O. europea</i>]	3234.45; 1793.94; 2523.20; 1600.80; 1267.63; 3156.46; 1787.91; 1235.67; 1022.55.
2	gil145313986	92	6	Major pollen allergen Ole e 1 [<i>O. europaea</i>]	1807.95; 2523.20; 1830.87; 1267.63; 1235.67; 1022.55
3	gil14424429	92	6	Major pollen allergen (Allergen Ole e 1) (Ole e I)	1807.95; 2523.20; 1830.87; 1267.63; 1787.91; 1022.55.
4	gil37724593	90	6	Major pollen allergen Ole e 1 [<i>O. europaea</i>]	1807.95; 2523.20; 1830.87; 1267.63; 1235.67; 1022.55.
5	gil37724597	88	6	Major pollen allergen Ole e 1 [<i>O. europaea</i>]	1807.95; 2523.20; 1830.87; 1267.63; 1235.67; 1022.55.

GASLR proposed sequence identifying unambiguously Ole e 1 (gil14424429, Figure 7) (Supporting Information Figure 4).

The exchange of one amino acid at the position 39 (S → G) leading to a mass difference of 30.01 mass units allows to

Table 6. Peptides Identified by PMF (Major pollen allergen (Allergen Ole e 1) (Ole e I); gi|14424429])

start– end	observed (Da)	MW (expt, Da)	MW (calc, Da)	sequence
24–40	1807.92	1806.92	1806.95	R.AGFITELSEFIPGASLR.L
24–46	2523.28	2522.27	2522.32	R.AGFITELSEFIPGASLR.LQCKDK.E
45–60	1830.81	1829.81	1829.84	K.DKENGDTVTFTEVGYTR.A
61–71	1267.61	1266.60	1266.63	R.AEGLYSMLVER.D
89–104	1787.83	1786.82	1786.86	K.DCNEIPTEGWAKPSLK.F
116–124	1022.55	1021.54	1021.56	R.TVNPLGFFK.K

attribute the sequence AGFITELSEFIPGAGVR to the experimental mass 1763.90 Da with 20 ppm error. This sequence is supported by the fragmentation pattern of the corresponding [M + H]⁺ species matching all peaks visible in the spectrum. The peptide AGFITELSEFIPGAGVR does not match the protein sequences identified by PMF. Therefore, the *Major pollen allergen Ole e 1* (gi|33329740, Figure 7) was assigned as the protein source of the peptide sequence AGFITELSEFIPGAGVR on the basis of the sequence congruency with the NCBI database. Accordingly, MS/MS data confirm the amino acid sequence AGFITELSEFIPGAGVR.

These results clearly show that the natural matrix contains *Major pollen allergen (Allergen Ole e1) (Ole e I)* (gi|14424429), *Ole e 1.0102 protein [O. europea]* (gi|2465127), *Ole e 1.0103 protein [O. europea]* (gi|2465129), *Major pollen allergen Ole e 1* (gi|33329740), *main olive allergen [O. europaea]* (gi|13195753) and *Ole e 1c* (gi|1362131) variants at different degree of distribution. Among these, only *Major pollen allergen (Allergen Ole e1) (Ole e I)* (gi|14424429) and *Ole e 1.0103 protein [O. europea]* (gi|2465129) are nonglycosylated isoallergens.

Conclusion

Structural function relationship is an important issue when dealing with post-translationally modified proteins. The situation is more complicated in the full characterization of glycoproteins.

A precise evaluation of the detailed structure of the isoforms of Olive pollen allergens could provide means for targeted therapies when needed. The results discussed above show that the combined use of chemical fractionation and MALDI MS/MS measurements provide a perfect two-dimensional approach for an unambiguous determination of the sequence, site of modification and glycan structure of the examined glycoproteins. An important role was also played by the N-terminal derivatization of the tryptic mixtures with dansyl chloride. Special attention was paid to the proper use of the available databases and programs. The widely used Glycomod program failed, in fact, to provide a correct answer when provided with input data available from acknowledged databases, whereas it contributed to a great extent in the final evaluation of the structure of the examined allergen isoforms when MS and MS/MS data, experimentally obtained as described within the manuscript, were used.

Acknowledgment. Paper was supported by PRIN 2006 and FIRB no. RBIP06SXM-R-004.

Supporting Information Available: MS/MS spectra and sequence identification by database search of the ion

species at *m/z* 1665.85, 1669.82, 1793.91 and 1807.92. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

(1) De Dios Alche, J.; Castro, A. J.; Olmedilla, A.; Fernandez, M., C.; Rodriguez, R.; Villalba, M.; Rodriguez-Garcia, M. I. *J. Cell Sci.* **1999**, *112*, 2501–2509.

(2) Villalba, M.; Lopez-Otin, C.; Martin-Orozco, E.; Monsalve, R. I.; Palomino, P.; Lahoz, C.; Rodriguez, R. *Biochem. Biophys. Res. Commun.* **1990**, *172*, 523–528.

(3) Villalba, M.; Batanero, E.; Monsalve, R. I.; Gonzalez De La Pena, M. A.; Lahoz, C.; Rodriguez, R. *J. Biol. Chem.* **1994**, *269*, 15217–15222.

(4) Lauzurica, P.; Giurbino, C.; Maruri, N. *Mol. Immunol.* **1988**, *25*, 329–335.

(5) Boluda, L.; Sastre, J.; Casanovas, M.; Fernandez-Caldas, E. *J. Immunol. Method* **1999**, *223*, 17–26.

(6) Van Ree, R.; Cabanes-Macheteau, M.; Akkerdaas, J.; Milazzo, J. P.; Loutelier-Bourhis, C.; Rayon, C.; Villalba, M.; Koppelman, S.; Aalberse, R.; Rodriguez, R.; Faye, L.; Lerouge, P. *J. Biol. Chem.* **2000**, *275*, 11451–11458.

(7) Wilson, I. B. H. D.; Altman, F. *Glycoconjugate J.* **1998**, *15*, 1055–1070.

(8) Villalba, M.; Batanero, E.; Lopez-Otin, C.; Sanchez, L. M.; Monsalve, R. I.; Gonzalez De La Pena, M. A.; Lahoz, C.; Rodriguez, R. *Eur. J. Biochem.* **1993**, *216*, 863–869.

(9) Villalba, M.; Batanero, E.; Lopez-Otin, C.; Sanchez, L. M.; Monsalve, R. I.; Gonzalez De La Pena, M. A.; Lahoz, C.; Rodriguez, R. *Eur. J. Biochem.* **1993**, *216*, 863–869.

(10) Lomardero, M.; Barbas, J. A.; Moscoso del Prado, J.; Carreira, J. *Clin. Exp. Allergy* **1994**, *24*, 765–770.

(11) Villalba, M.; Batanero, E.; Monsalve, R. I.; Gonzalez De La Pena, M. A.; Lahoz, C.; Rodriguez, R. *J. Biol. Chem.* **1994**, *269*, 15217–15222.

(12) Beretta, L. *Nat. Methods* **2007**, *4*, 785.

(13) Jensen, O. N. *Curr. Opin. Chem. Biol.* **2004**, *8*, 33–41.

(14) Gorg, A.; Weiss, W.; Deeth, H. C.; Alewood, P. F. *Proteomics* **2004**, *4*, 743–752.

(15) Lauber, W. M.; Carroll, J. A.; Dufield, D. R.; Kiesel, J. R.; Radabaugh, M. R.; Malone, J. P. *Electrophoresis* **2001**, *22*, 906–918.

(16) Gygi, S. P.; Corthals, G. L.; Zhang, Y.; Aebersold, R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9390–9395.

(17) Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R. *Nat. Biotechnol.* **1999**, *17*, 676–682.

(18) Napoli, A.; Aiello, D.; Di Donna, L.; Sajjad, A.; Perri, E.; Sindona, G. *Anal. Chem.* **2006**, *78*, 3434–3443.

(19) Napoli, A.; Aiello, D.; Di Donna, L.; Prendushi, H.; Sindona, G. *Anal. Chem.* **2007**, *79*, 5941–5948.

(20) Vrtala, S.; Grote, M.; Duchene, M. *Int. Arch. Allergy Immunol.* **1995**, *107*, 160–169.

(21) Carri, es, J.; Fernandez- Caldas, E.; Boluda, L.; Casanovas, M.; Sastre, J.; Lluch Bernal, M.; Blanca, M. *Allergy* **2002**, *57*, 798–804.

(22) Dell, A.; Morris, H. R. *Science* **2001**, *291*, 2351–2356.

(23) Holland, J. W.; Deeth, H. C.; Alewood, P. F. *Proteomics* **2006**, *6*, 3087–3095.

(24) Kolarich, D.; Weber, A.; Turecek, P.; Schwarz, H. P.; Altmann, F. *Proteomics* **2006**, *6*, 3369–3380.

(25) Rayon, C.; Lerouge, P.; Faye, L. *J. Exp. Bot.* **1998**, *49*, 1463–1473.

(26) Roth, K. D. W.; Huang, Z. H.; Sadagopan, N.; Watson, J. T. *Mass Spectrom. Rev.* **1998**, *17*, 255–274.

(27) Park, S. J.; Song, J. S.; Kim, H. J. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 3089–3096.

(28) Altmann, F. *Glycoconjugate J.* **1998**, *15*, 79–82.

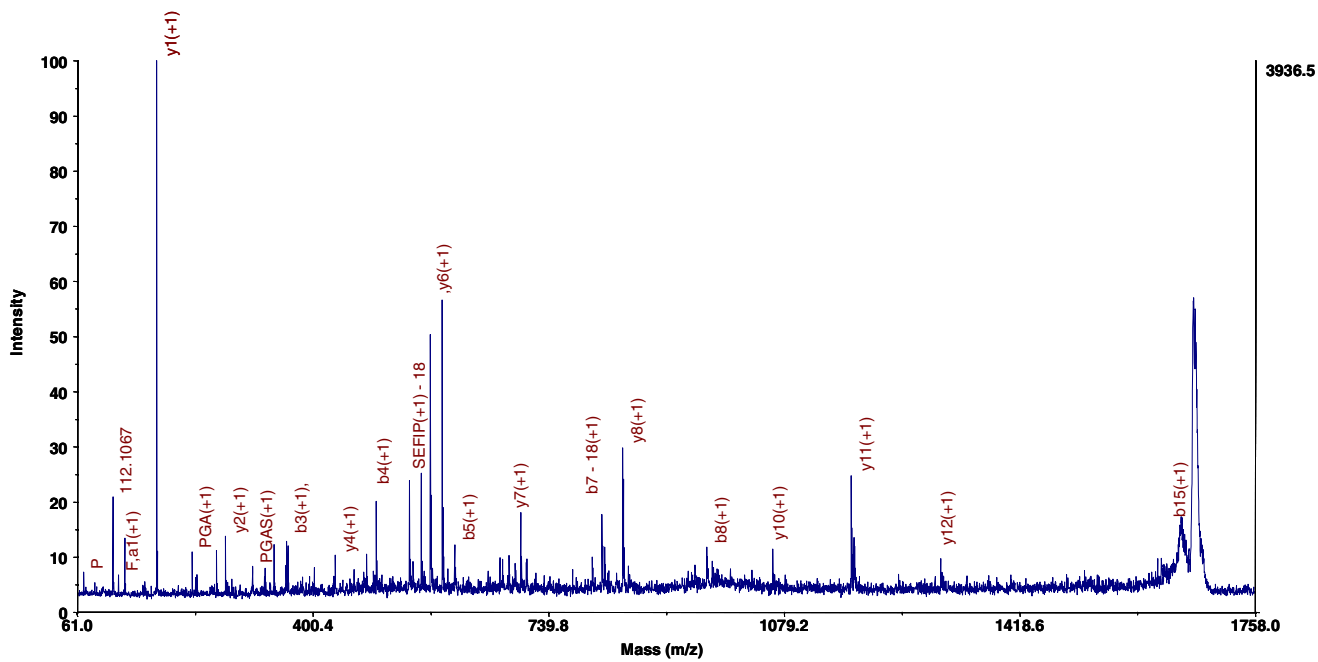
(29) Mechref, Y.; Novotny, M. V.; Krishnan, C. *Anal. Chem.* **2003**, *75* (18), 4895–4903.

(30) Asturias, J. A.; Arilla, M. C.; Gomez-Bayon, N.; Martinez, J.; Martinez, A.; Palacios, R. *J. Allergy Clin. Immunol.* **1997**, *100*, 365–372.

(31) Villalba, M.; Batanero, E.; Lopez-Otin, C.; Sanchez, L. M.; Monsalve, R. I.; Gonzalez De La Pena, M. A.; Lahoz, C.; Rodriguez, R. *Eur. J. Biochem.* **1993**, *216*, 863–869.

PR700855U

Supporting information Figure 1: (a) MS/MS spectrum of the ion specie at m/z 1665.85; (b) sequence identification by data base search.

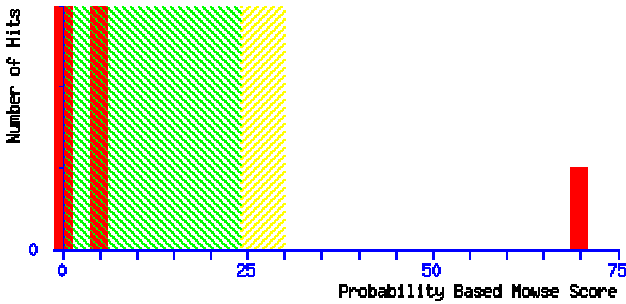


MASCOT SEARCH RESULTS

Database : NCBIInr (1411415 sequences; 454141287 residues)
Taxonomy : Other green plants (58941 sequences)
Significant hits: [gil14424429](#) Major pollen allergen (Allergen Ole e 1) (Ole e I)

Probability Based Mowse Score

Score is $-10 \times \log(P)$, where P is the probability that the observed match is a random event.
Individual ions scores > 24 indicate peptides with significant homology ($p < 0.05$).
Individual ions scores > 30 indicate identity or extensive homology ($p < 0.05$).



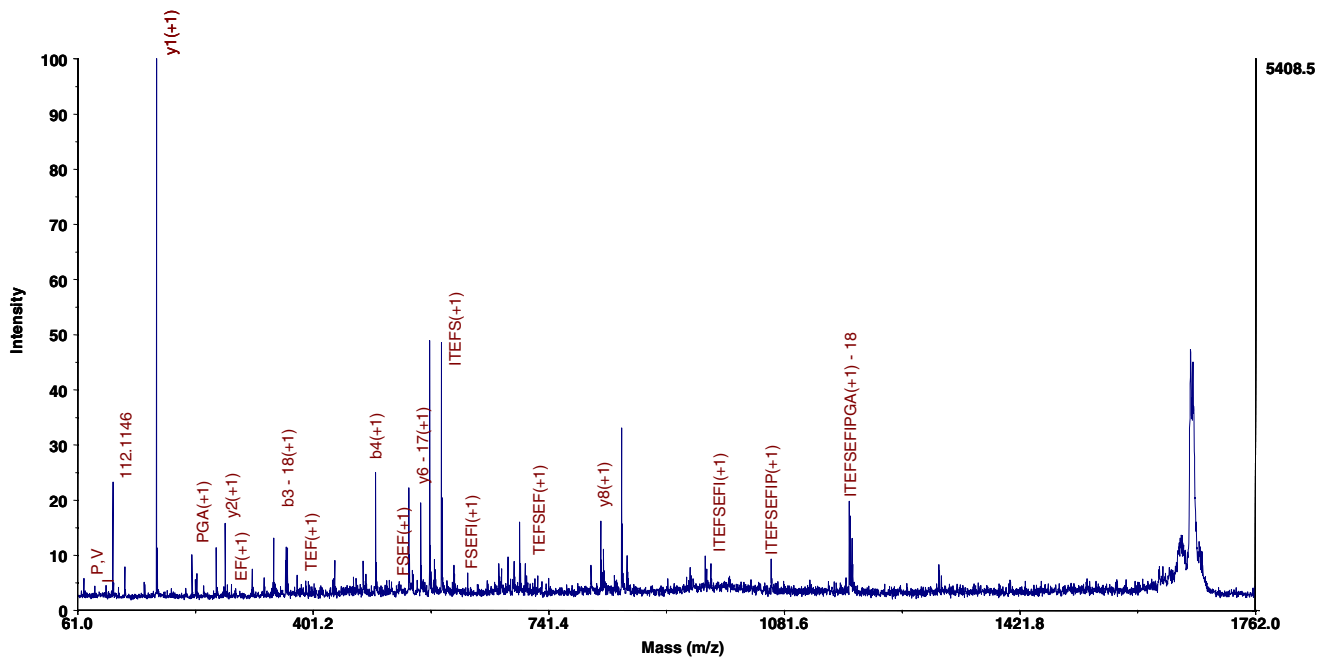
Peptide Summary Report

[1gil14424429](#) Mass: 16319 Total score: 70 Peptides matched: 1 Major pollen allergen (Allergen Ole e 1) (Ole e I)
Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
1	1665.85	1664.84	1664.88	-0.03	0	70	1	FITELSEFIPGASVR

Proteins matching the same set of peptides:
[gil1362136](#) Mass: 16404 Total score: 70 Peptides matched: 1 major allergen OLE5c - common olive
[gil2465131](#) Mass: 16605 Total score: 70 Peptides matched: 1 Ole e 1 protein [Olea europaea]

Supporting information Figure 2: (a) MS/MS spectrum of the ion specie at m/z 1669.82; (b) sequence identification by data base search.



MASCOT SEARCH RESULTS

Database : NCBIInr (1411415 sequences; 454141287 residues)

Taxonomy : Other green plants (58941 sequences)

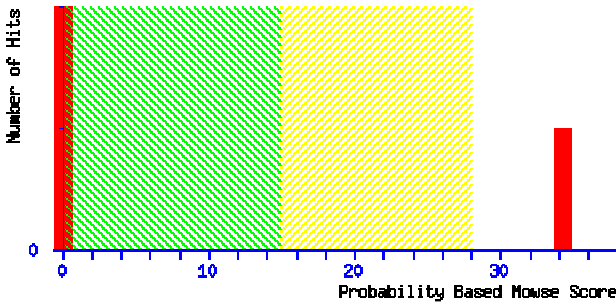
Significant hits: [gil1362131](#) major allergen OLE1c - common olive (fragment)

Probability Based Mowse Score

Score is $-10 \times \log(P)$, where P is the probability that the observed match is a random event.

Individual ions scores > 15 indicate peptides with significant homology ($p < 0.05$).

Individual ions scores > 28 indicate identity or extensive homology ($p < 0.05$).

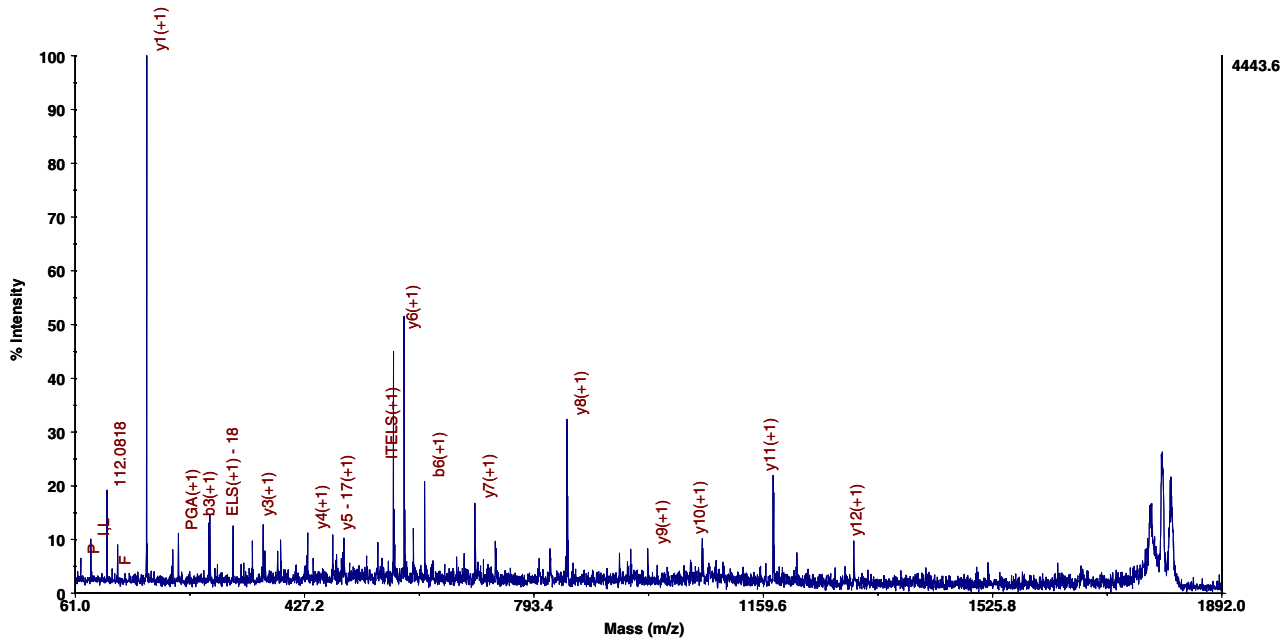


Peptide Summary Report

1. [gil1362131](#) Mass: 16474 Total score: 34 Peptides matched: 1 major allergen OLE1c - common olive (fragment)
Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
1	1669.82	1668.81	1668.85	-0.04	0	34	1	FITEFSEFIPGAGVR

Supporting information Figure 3: (a) MS/MS spectrum of the ion specie at m/z 1793.91; (b) sequence identification by data base search.



MASCOT SEARCH RESULTS

Database : NCBIInr (1411415 sequences; 454141287 residues)

Taxonomy : Other green plants (58941 sequences)

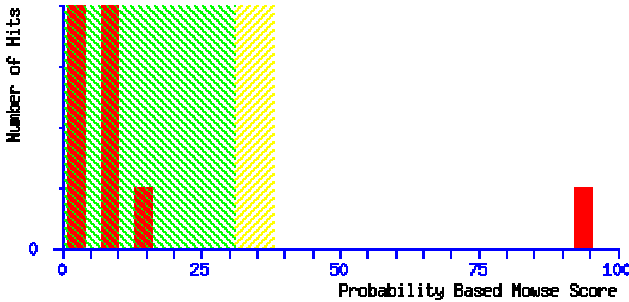
Significant hits: 1. [gi2465127](#) Ole e 1.0102 protein [Olea europaea]

Probability Based Mowse Score

Score is $-10 \cdot \log(P)$, where P is the probability that the observed match is a random event.

Individual ions scores > 31 indicate peptides with significant homology ($p < 0.05$).

Individual ions scores > 38 indicate identity or extensive homology ($p < 0.05$).



Peptide Summary Report

1. [gi2465127](#) Mass: 16536 Total score: 94 Peptides matched: 1 Ole e 1.0102 protein [Olea europaea]

Check to include this hit in error tolerant search or archive report

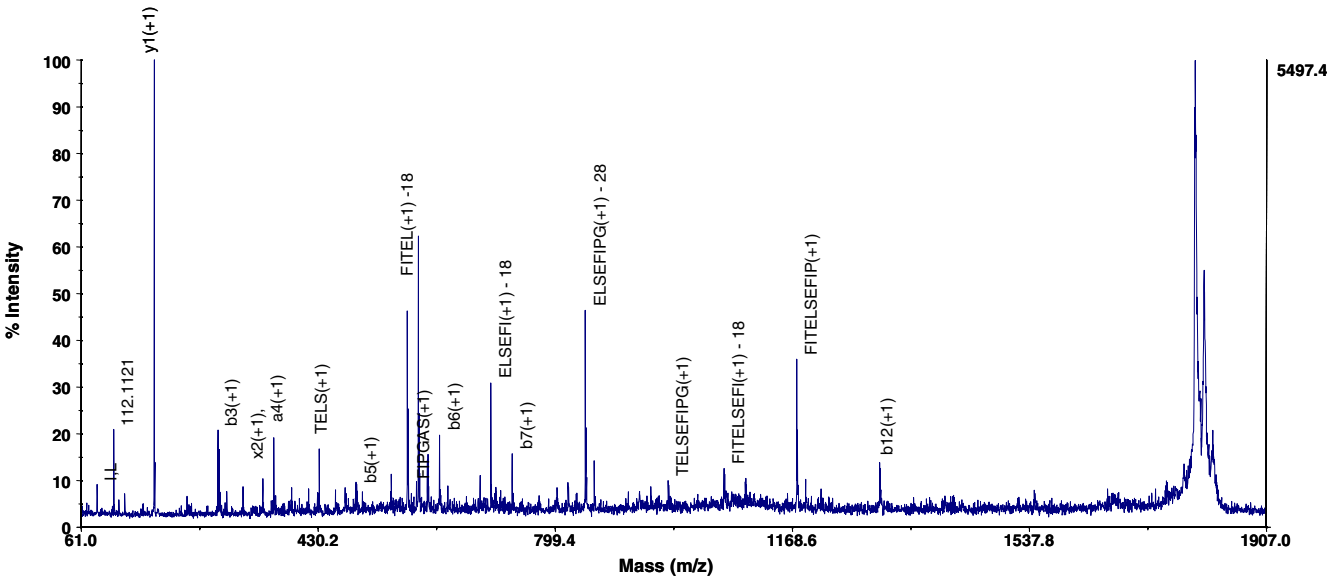
Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
1	1793.91	1792.90	1792.94	-0.04	0	94	1	AGFITELSEFIPGASVR

Proteins matching the same set of peptides:

[gi2465129](#) Mass: 16410 Total score: 94 Peptides matched: 1 Ole e 1.0103 protein [Olea europaea]

[gi1362133](#) Mass: 15288 Total score: 94 Peptides matched: 1 major allergen OLE26 - common olive (fragment)

Supporting information Figure 4: (a) MS/MS spectrum of the ion specie at m/z 1807.92; (b) sequence identification by data base search.

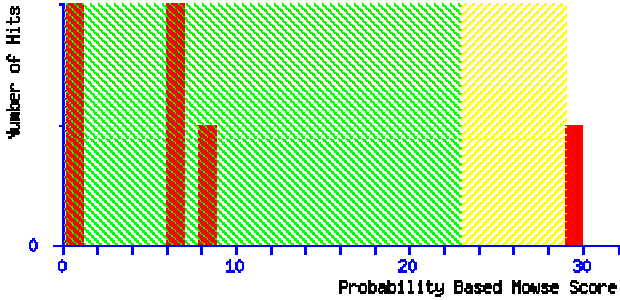


MASCOT SEARCH RESULTS

Database : NCBIInr (1411415 sequences; 454141287 residues)
 Taxonomy : Other green plants (58941 sequences)
 Significant hits: [gil14424429](#) Major pollen allergen (Allergen Ole e 1) (Ole e I)

Probability Based Mowse Score

Score is $-10 \times \log(P)$, where P is the probability that the observed match is a random event.
 Individual ions scores > 23 indicate peptides with significant homology ($p < 0.05$).
 Individual ions scores > 29 indicate identity or extensive homology ($p < 0.05$).



Peptide Summary Report

1. [gil14424429](#) Mass: 16319 Total score: 29 Peptides matched: 1 Major pollen allergen (Allergen Ole e 1) (Ole e I)

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
1	1807.92	1806.91	1806.95	-0.04	0	29	1	AGFITELSEFIPGASLR

Proteins matching the same set of peptides:

[gil1362135](#) Mass: 16318 Total score: 29 Peptides matched: 1 major allergen OLE3c - common olive
[gil1362134](#) Mass: 15384 Total score: 29 Peptides matched: 1 major allergen OLE33/OLE37 - common olive (fragment)