Journal of Proteomics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Proteomics

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

A recombinant isoform of the Ole e 7 olive pollen allergen assembled by *de novo* mass spectrometry retains the allergenic ability of the natural allergen

Carmen Oeo-Santos[a], Salvador Mas[a], Sara Benedé[a], María López-Lucendo[b], Joaquín Quiralte[c], Miguel Blanca[d,e], Cristobalina Mayorga[e], Mayte Villalba[a,*], Rodrigo Barderas[a,f,*]

[a] *Dpto. Bioquímica y Biología Molecular I, Universidad Complutense de Madrid, Madrid, Spain*
[b] *Centro de Investigaciones Biológicas, CSIC, Madrid, Spain*
[c] *Dpto. Alergología, Hospital Universitario Virgen del Rocío, Sevilla, Spain*
[d] *Servicio de Alergia, Hospital Infanta Leonor, Madrid, Spain*
[e] *Unidad de Alergología, Hospital Universitario Regional de Málaga-IBIMA, UMA, Málaga, Spain*
[f] *Unidad de Proteómica Funcional, UFIEC, CROSADIS, ISCIII, Majadahonda, Madrid, Spain*

ARTICLE INFO

Keywords:
Recombinant Ole e 7 allergen
Proteomics
de novo mass spectrometry
nsLTP
Olive pollen allergy

ABSTRACT

The allergenic non-specific lipid transfer protein Ole e 7 from olive pollen is a major allergen associated with severe symptoms in areas with high olive pollen levels. Despite its clinical importance, its cloning and recombinant production has been unable by classical approaches. This study aimed at determining by mass-spectrometry based proteomics its complete amino acid sequence for its subsequent expression and characterization. To this end, the natural protein was in-2D-gel tryptic digested, and CID and HCD fragmentation spectra obtained by nLC-MS/MS analyzed using PEAKS software. Thirteen out of the 457 *de novo* sequenced peptides obtained allowed assembling its full-length amino acid sequence. Then, Ole e 7-encoding cDNA was synthesized and cloned in pPICZαA vector for its expression in *Pichia pastoris* yeast. The analyses by Circular Dichroism, and WB, ELISA and cell-based tests using sera and blood from olive pollen-sensitized patients showed that rOle e 7 mostly retained the structural, allergenic and antigenic properties of the natural allergen. In summary, rOle e 7 allergen assembled by *de novo* peptide sequencing by MS behaved immunologically similar to the natural allergen scarcely isolated from pollen.

Significance: Olive pollen is an important cause of allergy. The non-specific lipid binding protein Ole e 7 is a major allergen with a high incidence and a phenotype associated to severe clinical symptoms. Despite its relevance, its cloning and recombinant expression has been unable by classical techniques. Here, we have inferred the primary amino acid sequence of Ole e 7 by mass-spectrometry. We separated Ole e 7 isolated from pollen by 2DE. After in-gel digestion with trypsin and a direct analysis by nLC-MS/MS in an LTQ-Orbitrap Velos, we got the complete *de novo* sequenced peptides repertoire that allowed the assembling of the primary sequence of Ole e 7. After its protein expression, purification to homogeneity, and structural and immunological characterization using sera from olive pollen allergic patients and cell-based assays, we observed that the recombinant allergen retained the antigenic and allergenic properties of the natural allergen. Collectively, we show that the recombinant protein assembled by proteomics would be suitable for a better *in vitro* diagnosis of olive pollen allergic patients.

1. Introduction

Olive pollen is responsible for Type I respiratory responses in the Mediterranean countries and large areas of America, South Africa, and Australia [1]. Thirteen allergens from *Olea europaea* -Ole e 1 to Ole e 13- have been identified and characterized to date [2, 3]. Among them, Ole e 7, a non-specific lipid binding protein (nsLTP), exhibits a molecular mass of 10 kDa and a basic character with multiple isoforms reported with pIs ranging from 8 to 10 [4]. nsLTPs are a family of proteins widely distributed throughout the plant kingdom [5]. These proteins are integrated into the protein superfamily of prolamins, which includes more than 3000 proteins from more than 250 species mainly found in plants [6]. nsLTPs contain an eight-cysteine residue conserved motif with four disulfide bonds, and an internal hydrophobic cavity where the lipid-binding site is located. Many of these nsLTPs have been described as relevant allergens from different biological sources such as

* Corresponding authors.
E-mail addresses: mvillalb@ucm.es (M. Villalba), r.barderasm@isciii.es (R. Barderas).

https://doi.org/10.1016/j.jprot.2018.06.001
Received 12 February 2018; Received in revised form 29 May 2018; Accepted 3 June 2018

pollen, food or latex [7–15]. Among them, Ole e 7 is a major allergen associated to severe clinical symptoms in those regions with high olive pollen counts, where olive tree is extensively cultivated such as Andalusia in Spain, or Portugal and Italy [3, 16]. Due mainly to its high degree of polymorphism or its mRNA instability, its cloning was unable, and Ole e 7 remained to be produced as a recombinant allergen from its identification in 1999 to be used in clinics [4], and mainly for diagnosis.

The primary structure of the majority of reported allergens has been commonly inferred from their corresponding nucleotide sequence obtained by cloning of their specific cDNA. However, as the genome sequence of many biological sources remains to be elucidated, and therefore if the protein of study is highly polymorphic or its mRNA is highly unstable, it might become virtually impossible to obtain its nucleotide sequence - as occurred with Ole e 7 - [4]. An alternative approach for the completion of amino acid sequences consists of *de novo* peptide sequencing, where a peptide sequence is directly obtained from a MS/MS spectrum and, therefore, a protein database is not required. Thanks to the extensive development of mass spectrometers, the primary structure of proteins from unsequenced biological sources genomes might now be reliably achieved. The combination of different fragmentation techniques and software, enabling the combination of complementary fragmentations (*i.e.* collision-induced dissociation –CID- and higher-energy collisional dissociation –HCD-), can be used for automated *de novo* peptide sequencing with high accuracy at amino acid level [17]. Indeed, these methods can now be applied not only for identifying and/or quantifying peptides and proteins, but also for determining the primary structure of allergens from unsequenced organisms [18–22].

The aim of the present study consists of the assembling of the whole Ole e 7 amino acid sequence from different peptide fragments obtained by multiple *de novo* peptide sequencing by mass spectrometry. Then, we produced Ole e 7 as a recombinant protein in *Pichia pastoris*, determining that its molecular and immunological behavior was similar to the natural allergen. Collectively, the here obtained data assessed the effectiveness of the strategy to determine the primary structure of highly polymorphic allergens from unsequenced genomes, and the usefulness of the recombinant Ole e 7 allergen assembled by proteomics for clinical patient diagnosis.

2. Materials and methods

See the Materials and methods section in this article's Supplementary data for full details about materials, expression and purification of the recombinant allergen, two-dimensional gel electrophoresis, analytical procedures, circular dichroism spectroscopic analyses, IgG and IgE immunoassays, inhibition experiments, and cell-based allergic mediator release assays.

2.1. Sera

Sera from 62 well-characterized patients sensitized to olive pollen were used to carry out immunological assays. Well-characterized sera from sensitized patients to olive pollen were obtained from two regions of Spain with high olive pollen counts: Jaén (n = 39) and Málaga (n = 23) [23–25]. Sera from non-atopic individuals were used as controls. Written informed consent was obtained from all patients. The work was performed accomplishing the Ethic Guidelines of Complutense University of Madrid.

2.2. Ole e 7 digestion, mass spectrometry analysis and de novo peptide sequencing

Ole e 7 (5 µg) from olive pollen was separated by two-dimensional gel electrophoresis -2DE- [23]. The main spot of the protein was in-gel digested with trypsin (Fig. 1) and peptides extracted using 100% acetonitrile (ACN) and 0.5% trifluoroacetic acid.

Extracted peptides were dried and purified using a Zip Tip with 0.6 µL C18 resin (Millipore). Finally, samples were reconstituted in 5 µL 0.1% formic acid/2% ACN before the analysis by nLC-MS/MS in the Proteomics Facility of the Centro de Investigaciones Biológicas (Madrid, Spain). Peptides were trapped in a C18-A1 ASY-Column 2-cm precolumn (Thermo Scientific) and eluted to a reverse phase column (15 cm × 75 µm inner diameter, Nanoseparations) packed with 3 µm particle C_{18} resin, where they were separated using a 150 min gradient from 0 to 35% buffer B (Buffer A: 0.1% Formic acid/2% ACN; Buffer B: 0.1% Formic acid in ACN) at a flow rate of 250 nL/min. The LC system, a NanoEasy HPLC (Proxeon), was coupled to an LTQ Orbitrap Velos (Thermo Scientific) *via* an electrospray ion source (Proxeon). The LTQ Orbitrap Velos was operated in an ion positive mode and in a data dependent manner as indicated [26, 27].

CID and HCD fragmentation spectra were analyzed with PEAKS Studio software (Bioinformatics Solutions Inc.) for *de novo* peptide sequencing. Amino acid sequences were aligned with DIALIGN and visualized with GeneDoc software.

2.3. rOle e 7 expression and purification

The synthetic cDNA-encoding sequence of Ole e 7 was purchased from IDT-DNA Technologies (Supplementary Fig. 1), directly subcloned in pPICZαA and transformed into KM71H *P. pastoris* strain yeast cells to express the protein in the extracellular media. Then, the extracellular media containing the recombinant protein was lyophilized and subsequently purified by chromatography and RP-HPLC.

See the methods section in this article's Supporting information for the full details about the expression and purification of rOle e 7.

3. Results

In previous reports, we identified, purified the allergen Ole e 7 from olive pollen, and analyzed its contribution to olive pollen allergy [3, 4, 28–30]. Although partial amino acid sequences of Ole e 7 were obtained by means of Edman degradation and peptide fingerprinting and deposited in protein databases, the high number of isoforms of Ole e 7 precluded its cloning and the completion of its amino acid sequence by classical techniques [4].

In this context, the main goal of this work consisted of determining the primary amino acid sequence of Ole e 7 by proteomics for its further expression and characterization in comparison to the natural protein isolated from pollen (Fig. 1).

3.1. Assembling of Ole e 7 amino acid sequence by nLC-MS/MS analysis

We performed a proteomic analysis of Ole e 7 isolated from pollen to obtain by *de novo* peptide sequencing as many peptides as possible to try to assemble the complete Ole e 7 primary sequence.

First, we separated Ole e 7 isolated from pollen by 2DE. Second, the main spot of Ole e 7 was in-gel digested with trypsin and directly analyzed by nLC-MS/MS in an LTQ-Orbitrap Velos (Fig. 1). Then, CID and HCD Fragmentation spectra were analyzed using PEAKS software. A total number of 457 *de novo* sequenced peptides were obtained (Table 1 and Supporting Table S1). *De novo* sequenced peptides with individual amino acid ALC confidence factors higher than 60% were used for the manually assembling of Ole e 7, using 3 previously reported amino acid peptides of Ole e 7 obtained by Edman degradation and peptide mass fingerprinting [4], and nsLTPs amino acid sequences deposited in databases. In total, we used for the manual assembling of the complete amino acid sequence of Ole e 7 13 out of 457 *de novo* peptides sequences (Table 1 and Fig. 2A).

Then, the assembled primary sequence of Ole e 7 was compared with allergenic LTPs amino acid sequences deposited in the WHO-IUIS database (Fig. 2B). Identity and similarity values with previously