BBA - Proteins and Proteomics 1865 (2017) 1067–1076

Contents lists available at ScienceDirect

BBA - Proteins and Proteomics

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

A relevant IgE-reactive 28 kDa protein identified from *Salsola kali* pollen extract by proteomics is a natural degradation product of an integral 47 kDa polygalaturonase

Salvador Mas[a,1], Carmen Oeo-Santos[a,1], Javier Cuesta-Herranz[b], Araceli Díaz-Perales[c], Carlos Colás[d], Javier Fernández[e], Domingo Barber[f], Rosalía Rodríguez[a], Vivian de los Ríos[g], Rodrigo Barderas[a,*], Mayte Villalba[a,*]

[a] Dpto. Bioquímica y Biología Molecular I, Universidad Complutense de Madrid, Madrid, Spain
[b] Hospital Fundación Jiménez Díaz, Madrid, Spain
[c] Centro de Biotecnología y Genómica de Plantas (UPM-INIA), Campus de Montegancedo, Pozuelo de Alarcón, Madrid, Spain
[d] Hospital Clínico Universitario 'Lozano Blesa', Zaragoza, Spain
[e] Dpto. Medicina Clínica, Universidad Miguel Hernández, Alicante, Spain
[f] Institute of Applied Molecular Medicine (IMMA), Universidad CEU San Pablo
[g] Centro de Investigaciones Biológicas, CSIC, Madrid, Spain

ARTICLE INFO

Keywords:
Sal k 6
Recombinant allergen
Cross-reactivity
Polygalacturonase
Amaranthaceae pollinosis

ABSTRACT

A highly prevalent IgE-binding protein band of 28 kDa is observed when *Salsola kali* pollen extract is incubated with individual sera from Amaranthaceae pollen sensitized patients. By an immunoproteomic analysis of *S. kali* pollen extract, we identified this protein band as an allergenic polygalacturonase enzyme. The allergen, named Sal k 6, exhibits a pI of 7.14 and a molecular mass of 39,554.2 Da. It presents similarities to Platanaceae, Poaceae, and Cupressaceae allergenic polygalacturonases. cDNA-encoding sequence was subcloned into the pET41b vector and produced in bacteria as a His-tag fusion recombinant protein. The far-UV CD spectrum determined that rSal k 6 was folded. Immunostaining of the *S. kali* pollen protein extract with a rSal k 6-specific pAb and LC-MS/MS proteomic analyses confirmed the co-existence of the 28 kDa band together with an allergenic band of about 47 kDa in the pollen extract. Therefore, the 28 kDa was assigned as a natural degradation product of the 47 kDa integral polygalacturonase. The IgE-binding inhibition to *S. kali* pollen extract using rSal k 6 as inhibitor showed that signals directed to both protein bands of 28 and 47 kDa were completely abrogated. The average prevalence of rSal k 6 among the three populations analyzed was 30%, with values correlating well with the levels of grains/m^3 of Amaranthaceae pollen. Sal k 6 shares IgE epitopes with Oleaceae members (*Fraxinus excelsior*, *Olea europaea* and *Syringa vulgaris*), with IgE-inhibition values ranging from 20% to 60%, respectively. No IgE-inhibition was observed with plant-derived food extracts.

1. Introduction

Russian thistle (*Salsola kali*) is an important source of anemophilous allergenic pollen, which is easily spread over kilometers from its source of production. This weed is becoming an important allergy-inducer in large temperate areas from Middle Eastern countries, USA and Southern Europe [1,2], and, in general, in those regions where the desertification is an important environmental problem. Five allergens have been characterized from this weed pollen so far (for a review see [3]): the specific allergenic marker Sal k 1 (pectin methylesterase) [4], the protein kinase homologue Sal k 2, the cobalamin independent methionine synthase Sal k 3 [5], the panallergen profilin Sal k 4 [6,7], and the Ole e 1-like allergen Sal k 5 [8].

Polygalacturonases (PGs) are glycosylated enzymes belonging to the pectinase group that catalyze the hydrolysis of α-1,4 glycosidic bonds located in the polygalacturonan backbone of α-1,4 linked D-galacturonic acid residues, a component of the plant cell wall. These enzymes are active in processes that involve degradation of the cell wall such as ripening, abscission of fruit or pollen maturation [9]. An important role of these enzymes also resides in the fact that some pathogens like fungi

* Corresponding authors at: Dpto. Bioquímica y Biología Molecular, Facultad de Ciencias Químicas, Universidad Complutense, 28040 Madrid, Spain.
E-mail addresses: rbarderas@quim.ucm.es (R. Barderas), mvillalb@ucm.es (M. Villalba).
[1] Both authors contributed equally.

http://dx.doi.org/10.1016/j.bbapap.2017.05.007
Received 18 September 2016; Received in revised form 7 April 2017; Accepted 8 May 2017
Available online 10 May 2017
1570-9639/ © 2017 Elsevier B.V. All rights reserved.

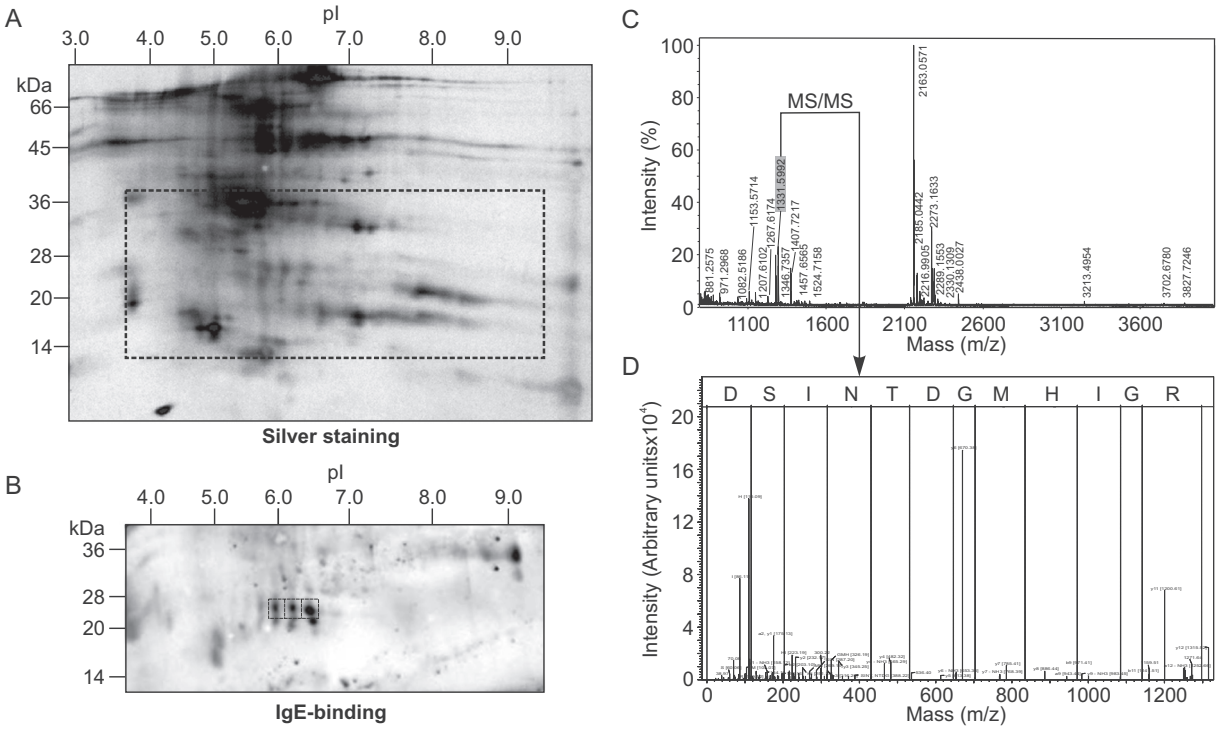


Fig. 1. Identification of polygalacturonase as a new allergen of *S. kali* pollen. (A) Silver staining of 75 μg of *S. kali* pollen protein extract after two-dimensional electrophoretic separation. (B) Identification of IgE-reactive spots in the Box highlighted in A by 2DE using an equivolumetric pool of sera ($n = 3$) from *S. kali* sensitized patients. (C) Boxed spots in B were excised and subsequently used for in-gel trypsin digestion and mass fingerprinting analysis and *de novo* sequencing. (D) The highlighted peptide in (C) was obtained for all three excised spots and used for *de novo* MS/MS amino acid sequencing. The amino acid sequence obtained is showed at the top of the graph. Molecular mass markers and pI are shown in (A) and (B).

use their enzymatic activities to degrade plant cell wall, and then penetrate inside of the host cell [10]. Thus, these proteins have also a significant value for industrial purposes [11,12].

Three biological families of plants with allergenic relevant species have been found to contain allergenic PGs: *Chamaecyparis obtusa* (Cha o 2), *Cryptomeria japonica* (Cry j 2) and *Juniperus ashei* (Jun a 2) from Cupressaceae family [13–15]; *Platanus orientalis* (Pla or 2) and *Platanus acerifolia* (Pla a 2) from Platanaceae family [16]; *Paspalum notatum* (Pas n 13) [17], *Zea mays* (Zea m 13) and *Phleum pratense* (Phl p 13) from Poaceae family [18]. Therefore, the widespread distribution of PGs in the vegetal kingdom pointed out to a potential role of these proteins in allergenic cross-reactivity.

The analysis of the IgE-binding profile of *S. kali* sensitized patients by SDS-PAGE shows several protein bands with important IgE-binding ability, which corresponds to unknown allergens [6]. Interestingly, about 30% of *S. kali* sensitized patients show reactivity against a particular band of a molecular mass of about 28 kDa.

Based on this information, we have here performed the proteomic analysis of the 28 kDa IgE-binding protein (Fig. 1). We separated the *S. kali* pollen protein extract by two-dimensional gel electrophoresis (2DE), and subsequently identified by mass spectrometry that this relevant IgE-binding protein belonged to the PG family. It was named as Sal k 6, following the rules of the allergen nomenclature. To get further insights into the molecular properties of this new *S. kali* allergen, we cloned, sequenced and expressed Sal k 6 in *Escherichia coli* as a recombinant His-tag fusion protein. We performed a structural and immunological characterization of the recombinant protein, and demonstrated by proteomics and immunological analyses that the 28 kDa protein band was actually the proteolytic degradation product of a 47 kDa allergen. This new allergen from *S. kali* pollen was unnoticed because it shares molecular mass and pI with IgE-binding protein bands, including the most prevalent allergen of *S. kali* pollen such as Sal k 1 [4]. We measured the Sal k 6 sIgE levels using sera of 88 *S. kali* allergic patients from three geographically different populations from Spain with very different numbers of Amaranthaceae pollen counts in the living area of the allergic patients; and thus, we obtained a Sal k 6 prevalence ranging between 17 and 39%, which correlated with the pollen counts.

2. Materials and methods

2.1. Two-dimensional gel electrophoresis (2DE)

Isolectric focusing was achieved under reducing conditions in the presence of 3 mM tributylphosphine in a PROTEAN IEF Cell (Bio-Rad) using 7 cm length, pH 3–10 linear ReadyStrip IPG gels (Bio-Rad). After isoelectrofocusing, proteins were separated by SDS-PAGE in 15% polyacrylamide gels under reducing conditions in the presence of 50 mM dithiothreitol and 3.7% iodoacetamide. 2DE gels were alternatively stained with silver, Coomassie Blue G-250 reagents, or transferred to nitrocellulose membranes for immunostaining.

2.2. Salsola kali pollen extraction and de novo mass spectrometry sequencing

S. kali pollen was provided by Biopol Laboratory Inc. (Washington, USA). *S. kali* pollen protein extract (75 μg), obtained as previously described [8], was subjected to 2DE. Then, the gel was incubated during 3 h in 50% methanol containing 2% phosphoric acid, washed 3 times with ultrapure water, followed by an incubation during 1 h in 33% methanol containing 17% ammonium sulphate and 3% phosphoric acid. Finally, the gel was stained with 0.06% Coomassie blue G-250 (Sigma-Aldrich, Química, Spain) during 16 h, and destained with milliQ water. Then, indicated spots were excised to perform peptide mass fingerprinting (PMF) and *de novo* sequencing in the Proteomics Facility of the Complutense University (Madrid, Spain). MS/MS data of indicated excised spots from 2DE gels were acquired on a Bruker-Reflex IV MALDI-TOF (Bruker-Franzer Analytic, Bremen, Germany). As no

```
     atgaagacatttaatctcccattgttggtagccctgttctacctgtttgtatctgtggcaagatcc   66
1     M  K  T  F  N  L  P  L  L  V  A  L  F  Y  L  F  V  S  V  A  R  S
     cagggtcccattgatatcacaaaatttggtgctaagcctaatgcagatgcaacatcggctcttttg  132
23    Q  G  P  I  D  I  T  K  F  G  A  K  P  N  A  D  A  T  S  A  L  L
     gctgcctggaaggaagcgtgtgcagcagcagcgccagcgaaaatcgtagtgccagcaggagagttc  198
45    A  A  W  K  E  A  C  A  A  A  A  P  A  K  I  V  V  P  A  G  E  F
     ttgttgaatgccgtgaagcttcaaggtccatgcaaggctcctcttactattgaaattgctggaaac  264
67    L  L  N  A  V  K  L  Q  G  P  C  K  A  P  L  T  I  E  I  A  G  N
     ttcaaagcaccagcagatgttgcacagatgaaaggtgaagacacatgggttaagatcgagaatgtc  330
89    F  K  A  P  A  D  V  A  Q  M  K  G  E  D  T  W  V  K  I  E  N  V
     caaggcctcaccatcacttgtctcccaactggaggcactttcgatggccaagggcaagccgcatgg  396
111   Q  G  L  T  I  T  C  L  P  T  G  G  T  F  D  G  Q  G  Q  A  A  W
     aagcagaacaagtgcgctcaaagcggcatgtgcaacagtcttccttacaattttaggttcaacacc  462
133   K  Q  N  K  C  A  Q  S  G  M  C  N  S  L  P  Y  N  F  R  F  N  T
     ctgaccaatgctcagatcagcgggatcaaatctttgaacagcaagctgtaccacatgggagttatg  528
155   L  T  N  A  Q  I  S  G  I  K  S  L  N  S  K  L  Y  H  M  G  V  M
     gggtgcaaaaacattacactaacaggattgaccattgatgcaccaaaggacagtcttaacacagat  594
177   G  C  K  N  I  T  L  T  G  L  T  I  D  A  P  K  D  S  L  N  T  D
     ggtatgcacattggacgatcaaatggggtgcatgccaccaactcaaagatcggtacaggagatgac  660
199   G  M  H  I  G  R  S  N  G  V  H  A  T  N  S  K  I  G  T  G  D  D
     tgtatctctatgggagatggggctgttgatgttcacgtcgagggtatcacctgtggacctggccat  726
221   C  I  S  M  G  D  G  A  V  D  V  H  V  E  G  I  T  C  G  P  G  H
     ggtatcagtatcggaagcatgggtaagttcgctaatgaggctcccaacactggtatctttgtcaag  792
243   G  I  S  I  G  S  M  G  K  F  A  N  E  A  P  N  T  G  I  F  V  K
     aactgcagcttcacagacactgacaacggtgttaggatcaagtcttggatgaactcttttgaggct  858
265   N  C  S  F  T  D  T  D  N  G  V  R  I  K  S  W  M  N  S  F  E  A
     agcgcttctgatctccacttcgaggacatcaccgttaccaacgttttgaaccctgttatcatcgat  924
287   S  A  S  D  L  H  F  E  D  I  T  V  T  N  V  L  N  P  V  I  I  D
     caggagtactgcccctacaaccactgcaaggagaagactccatcaaaggttaagctcagcaagatc  990
309   Q  E  Y  C  P  Y  N  H  C  K  E  K  T  P  S  K  V  K  L  S  K  I
     agtttcaagaacgttcatggagctgcaaaatcagctgaagtcgtcaagctattgtgcagcagcgca 1056
331   S  F  K  N  V  H  G  A  A  K  S  A  E  V  V  K  L  L  C  S  S  A
     gttccctgtgatggtgttgagcttgctgacatcgacctcactttccctggtggtgccgctgtctca 1122
353   V  P  C  D  G  V  E  L  A  D  I  D  L  T  F  P  G  G  A  A  V  S
     cagtgcaagaatgttaagccgattgtcactggcaagcagaaccctgtcgcttgtggtgcacctgct 1188
375   Q  C  K  N  V  K  P  I  V  T  G  K  Q  N  P  V  A  C  G  A  P  A
     acccctgctgccccataa                                                 1212
397   T  P  A  A  P  -
```

Fig. 2. Nucleotide and amino acid sequence of PG from *S. kali* pollen. The signal peptide identified using SignalP 4.0 is highlighted in gray. The peptide obtained by *de novo* sequencing by mass spectrometry is boxed with a continuous line. The nucleotide sequence derived from the amino acid sequence identified by mass spectrometry and used to design the sense and antisense oligonucleotides for Sal k 6 is highlighted in italics. Aspartic residues described to be involved in the enzymatic activity of PGs are boxed with a dashed line. The two potential N-glycosylation sites at amino acid 180 and 265 in mature Sal k 6 are also highlighted with a circle. Numbers in the left of the figure, number of the initial amino acid of the line. Numbers in the right of the figure, number of the last nucleotide in the line. The sequence is deposited in the GenBank database with the entry KY883988 for the complete cDNA sequence, and KC920919 for the entry of the mature Sal k 6 without the signal peptide.

protein assignment was obtained by PMF since *S. kali* genome has not been reported, *de novo* sequencing was carried out and the same peptide for all analyzed protein spots was obtained -DSINTDGMHIGR-.

2.3. Cloning of the allergenic polygalacturonase from S. kali pollen

A sense 5′-AAYACNGAYGGNATGCAYATH-3′ and antisense 5′-DATRTGCATNCCRTCNGTRTT-3′ degenerate oligonucleotides were designed from the partial amino acid sequence NTDGMHI with sequence identity with PGs obtained by *de novo* sequencing by mass spectrometry (Fig. 2). We synthesized the cDNA from total RNA *S. kali* pollen using the SMART RACE cDNA amplification kit (BD Biosciences, Clontech, Madrid, Spain). Then, we amplified by PCR the cDNA-encoding the 5′ and 3′ sequences of *S. kali* PG, using the previously described sense and antisense oligonucleotides and the universal primer contained in the same kit according to the manufacturer instructions. The PCR fragments obtained corresponding to partial 5′ and 3′ cDNA sequences were cloned into the pCR2.1 (Invitrogen, Groningen, The Netherlands) and sequenced. Then, a sense specific oligonucleotide was designed from the N-terminus of *S. kali* PG (QGPIDI) 5′-aatcatATGCAGGGTCCCATTGATATC-3′, containing a *Nde*I restriction site (underlined) without the signal peptide of *S. kali* PG identified with SignalP software [19] (Fig. 2). The antisense specific oligonucleotide was designed from the C-terminus (PATPAAP) without the stop codon and an *Xho*I restriction site (underlined) 5′-taCTCGAGTGGGGCAGCAGGGGTAGCAGG-3. Finally, the complete cDNA-encoding the whole PG nucleotide sequence was subcloned, using the restriction enzymes *Nde*I and *Xho*I, into the pET41b plasmid (Novagen, Billerica, MA, USA) as a fusion protein linked to an 8xHis-tag at the C-terminus (His-tag).

2.4. Expression and purification of S. kali recombinant polygalacturonase

The pET41b/Salk6 construct was used to transform BL21(DE3) *E. coli* cells. LB supplemented with 100 μg/mL kanamycin containing these cells was grown overnight, and then, diluted ten times and maintained at 37 °C until an optical density at 600 nm reached a value of 1.0. Then, cell cultures were induced with 1 mM isopropyl thio-β-D-galactoside at indicated times and temperatures to determine the best conditions for the production of the recombinant protein. Then, the culture was centrifuged at 6000*g* during 20 min at 4 °C. Inclusion bodies were solubilized during 1 h with 20 mM Tris-HCl pH 8.0, 0.5 M NaCl, 5 mM imidazole and 1 mM 2-mercaptoethanol containing 6 M guanidine hydrochloride, and subsequently clarified by centrifugation at 12000*g* during 30 min at 4 °C and the supernatant filtered through 0.22 μm prior to be applied at 0.5 mL/min into an His Trap FF crude (GE healthcare, Madrid, Spain). The column was washed with 20 mM

Tris-HCl pH 8.0, 0.5 M NaCl, 5 mM imidazole and 1 mM 2-mercaptoethanol containing 6 M urea. Then, a 60 min gradient to 20 mM Tris-HCl pH 8.0, 0.5 M NaCl, 5 mM imidazole and 1 mM 2-mercaptoethanol was performed for on-column refolding. Finally, the recombinant protein was eluted by an isocratic gradient with the same buffer containing 0.5 M imidazole. Fractions containing the purified protein were visualized by SDS-PAGE, pooled, dialyzed against 20 mM ammonium bicarbonate, lyophilized and stored at − 80 °C until use.

2.5. Sera and antibodies

Sera from 88 patients from different regions of Spain, suffering from pollinosis to Amaranthaceae pollen were collected from Alicante (n = 32), Zaragoza (n = 28) and Murcia (n = 28). Written informed consent was obtained from all patients [20].

Mouse polyclonal antisera against recombinant PG from *S. kali* pollen was obtained by weekly injections of the protein (1 μg) in complete Freud's adjuvant [21], accomplishing the Ethic Guidelines of Complutense University of Madrid. Horseradish peroxidase-labeled goat polyclonal antibody against rabbit IgG was obtained from Bio-Rad (Richmond, Calif). Horseradish peroxidase-labeled goat polyclonal antibody against mouse IgG was purchased from Pierce Chemical Co. (Rockford, Ill). Horseradish peroxidase-labeled mouse monoclonal antibody against His-tag was from Sigma-Aldrich.

2.6. Analytical procedures

SDS-PAGE was performed under reducing conditions in 15% polyacrylamide gels. Proteins were stained with Coomassie Blue R-250 (Sigma-Aldrich). Molecular mass determinations were done with unstained protein molecular weight markers SM0431 (Fermentas).

The concentration of the purified recombinant PG was calculated by measuring the absorbance at 280 nm in a DU-7 spectrometer (Beckman) using the theoretical extinction coefficient ($E^{0.1\%}$) of 0.7 calculated with ProtParam. For pollen and plant-derived food extracts, the bicinchoninic acid protein assay reagent (Thermo Scientific, Rockford, IL, USA) was used to determine the total concentration of proteins.

2.7. CD spectroscopic analysis

The CD spectrum was recorded in the far UV on a JASCO J-715 spectropolarimeter (Japan Spectroscopic Co., Tokyo, Japan) at a 200 μg/mL protein concentration in sodium phosphate 50 mM, pH 7.5 at room temperature. CDNN CD spectra deconvolution software (Applied Photophysics) was used to deconvolute the CD spectra recorded. Theoretical percentages of secondary structure were obtained with GOR [22].

2.8. Immunological characterization

Proteins were blotted onto nitrocellulose membranes (Amersham Biosciences, Barcelona, Spain) after SDS-PAGE or 2DE. Membranes were incubated with individual human sera (diluted 1:10), anti-human IgE monoclonal antibody (diluted 1:5000) and anti-mouse IgG horseradish peroxidase-labeled antibody. Alternatively, a mouse polyclonal antisera raised against rSal k 6 (diluted 1:50,000) followed by an anti-mouse IgG horseradish peroxidase-labeled polyclonal antibody (1:3000). Alternatively, to detect the recombinant fusion protein, a horseradish peroxidase-labeled mouse monoclonal antibody against His-tag (1:2000) was used. For immunoblotting inhibition assays, mouse polyclonal antisera raised against rSal k 6 (diluted 1:5000), individual sera or an equivolumetric pool of sera (n = 5) (diluted 1:10) was preincubated with PBS, different amounts of recombinant PG or pollen and plant-derived food extracts. Chemiluminiscent signal was developed by ECL-Western blotting reagent (Amersham Bioscience).

Indirect ELISA was performed in 96 well plates coated with 500 ng of extract or 100 ng of purified recombinant protein per well. Prevalence of Sal k 6 was calculated with sera from *S. kali* sensitized patients (diluted 1:10), taking as positive reactions those absorbance values at 492 nm greater than three-fold the standard deviations above the mean of controls from non-atopic donors. For inhibition assays, the specific polyclonal antisera raised against rSal k 6 (diluted 1:5000) or an equivolumetric pool of human sera (n = 5) (diluted 1:10) were previously adsorbed to different concentrations of pollen and plant-derived food extracts as inhibitors. Binding of a PG-specific polyclonal antisera was detected by goat anti-mouse IgG horseradish peroxidase-labeled antibody (1:2500) and the binding of human IgE was detected by an anti-human IgE monoclonal antibody (diluted 1:5000) followed by rabbit anti-mouse IgG horseradish peroxidase-labeled antibody (diluted 1:2500). Signal was measured at 492 nm in an iMark Microplate Absorbance Reader (Bio-Rad). In ELISA experiments with individual serum, absorbance values above 0.1 were taken as positives. The percentage of inhibition was calculated according to the following formula: inhibition (%) = [1 − ($OD_{492\ nm}$ with inhibitor / OD_{492nm} without inhibitor)] × 100.

2.9. LC-MS/MS analysis

S. kali pollen protein extract was subjected to size-exclusion chromatography onto a Sephadex G-75 column chromatography as previously done [8]. Fractions of 2.5 mL were analyzed by 15% SDS-PAGE and immunostaining using a pAb against Sal k 6. Indicated enriched fractions in Sal k 6 were concentrated on an SDS-PAGE, and stained with Coomassie Brilliant Blue G-250. Subsequently, in-gel trypsin digestion was performed and peptide extraction performed according to previously published protocols at the Proteomics Facility of the Centro de Investigaciones Biológicas (Madrid, Spain) [23,24]. Extracted peptides were then directly subjected to MS/MS analysis.

Briefly, all peptide separations were carried out on an Easy-nLC 1000 nano system (Thermo Scientific). For each analysis, the sample was loaded into a precolumn Acclaim PepMap 100 (Thermo Scientific) and eluted in a RSLC PepMap C18, 15 cm long, 75 μm inner diameter and 2 μm particle size (Thermo Scientific). The mobile phase flow rate was 300 nL/min using 0.1% formic acid in water (solvent A) and 0.1% formic acid and 100% acetonitrile (solvent B). The gradient profile was set as follows: 0–35% solvent B for 90 min, 35%–100% solvent B for 4 min, 100% solvent B for 8 min. Four microliters of each sample were injected.

MS analysis was performed using a Q Exactive mass spectrometer (Thermo Scientific). For ionization, 1800 V of liquid junction voltage and 270 °C capillary temperature was used. The full scan method employed a m/z 400–1500 mass selection, an Orbitrap resolution of 70,000 (at m/z 200), a target automatic gain control (AGC) value of 3×10^6, and maximum injection times of 100 ms. After the survey scan, the 15 most intense precursor ions were selected for MS/MS fragmentation. Fragmentation was performed with a normalized collision energy of 27 eV and MS/MS scans were acquired with a starting mass of m/z 100, AGC target was 2×10^5, resolution of 17,500 (at m/z 200), intensity threshold of 8×10^3, isolation window of 2 m/z units and maximum IT was 100 ms. Charge state screening was enabled to reject unassigned, singly charged, and equal or more than seven protonated ions. A dynamic exclusion time of 20s was used to discriminate against previously selected ions.

2.10. MS data analysis

MS data were analyzed with Proteome Discoverer (version1.4.1.14) (Thermo) using standardized workflows. Mass spectra *.raw files were searched against the Uniprot Viridiplantae database containing the Sal k 6 here described protein sequences (3,336,992 sequences protein entries) using SEQUEST search engine. Precursor and fragment mass tolerance were set to 10 ppm and 0.02 Da, respectively, allowing 2

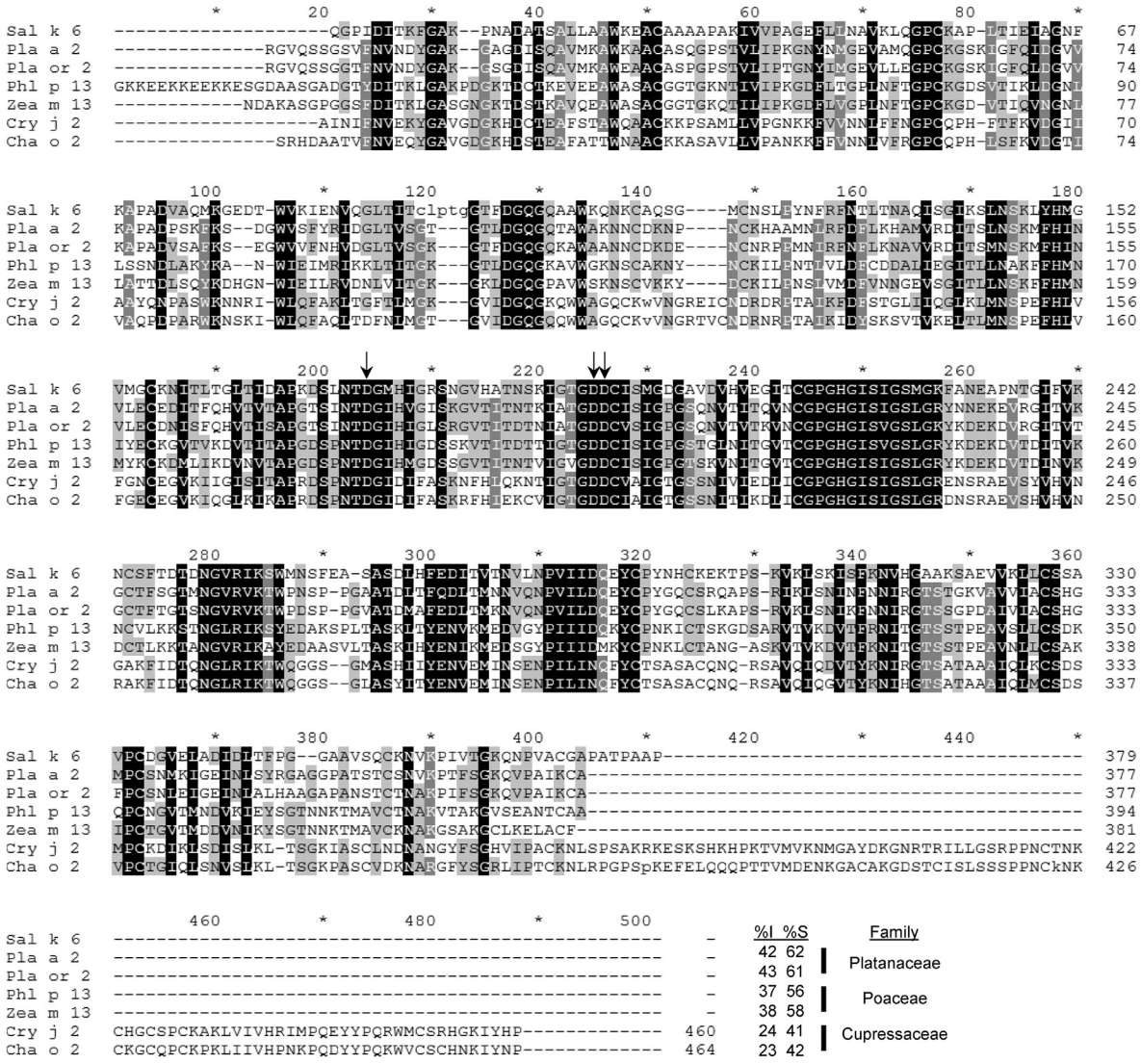


Fig. 3. Alignment of Sal k 6 amino acid sequence with allergenic pollen PGs from indicated taxonomic plant families. Identity and similarity percentages were calculated using as reference the mature amino acid sequence of Sal k 6. Arrows indicate the aspartic residues described to be involved in the enzymatic activity of PG.

missed cleavages, carbamidomethylation of cysteines as a fixed modification and methionine oxidation as a variable modification. Identified peptides were filtered using Percolator algorithm with a q-value threshold of 0.01 [25].

2.11. Basophil activation test (BAT)

This test was performed as previously described [26]. After blood-cell separation, 50 μL of each patient's cell suspension were incubated with 50 μL of the indicated protein concentration. To evaluate background basal values without stimulation (negative control), we added 50 μL of stimulation buffer (cRPMI), containing IL-3 (2 ng/mL) in the cell suspension. As a positive control, a monoclonal anti-IgE anti-receptor antibody was used at a final concentration of 1 μg/mL. A positive response was concluded for values ≥ 15%, and stimulation index (antigen-specific response/basal level) ≥ 2.

3. Results

3.1. Identification of an IgE-reactive protein by de novo sequencing

The analysis of the IgE-binding profile of *S. kali* pollen extract with individual sera revealed the presence of a protein band of 28 kDa recognized by the IgE of approximately 30% of the Amaranthaceae pollen sensitized patients (Supplementary material, Fig. S1) [4,8]. Here, we aimed to identify and characterize this allergenic protein band.

To this end, we firstly set up a proteomic analysis for the identification of this allergenic protein. We performed a 2DE analysis of *S. kali* pollen protein extract by silver staining and immunostaining with sera from sensitized patients reactive to the 28 kDa protein band (Fig. 1A and B). After immunostaining, three spots with IgE-binding ability, corresponding to the 28 kDa protein and pI ranging from 5.9 to 6.5, were excised and analyzed by *de novo* sequencing by mass spectrometry (Fig. 1B and D). In all three analyzed spots, we obtained the same amino acid sequence–DSINTDGMHIGR-(Fig. 1C and D). BlastP analyses revealed amino acid sequence identity with the PG protein family. No other amino acid peptide sequences were obtained from the analysis with identity to any other protein deposited in the database. Finally, in addition to the three analyzed spots, we also observed three others with lower-molecular masses and IgE binding ability, which should correspond to the non-glycosylated forms of the identified PG.

3.2. Cloning, sequencing and analysis of the cDNA-encoding the allergen Sal k 6

The peptide obtained by mass spectrometry was used to design

sense and antisense degenerate oligonucleotides to amplify by PCR the cDNA codifying for the N-terminal and C-terminal end of the *S. kali* PG. In a second round of PCR, we amplified the whole cDNA-encoding sequence by using specific oligonucleotides synthesized from the 5′- and 3′-end nucleotide sequence of the PG previously sequenced. This nucleotide sequence was deposited in the GenBank® database with the accession number KY883988 for the complete sequence and KC920919 for the mature protein without the signal peptide and named Sal k 6 according to the allergen nomenclature (Fig. 2).

The analysis of the deduced amino acid sequence of Sal k 6 with SignalP 4.0 and ProtParam showed the presence of two potential N-glycosylation sites at amino acids 180 and 265, highlighting the glycosylation character of these enzymes. In addition, a leader sequence in the N-terminus amino acid sequence was also identified. The mature protein without the signal peptide consisting of 379 amino acids in length exhibited a theoretical pI of 7.14 and a molecular mass of 39,554.2 Da (Fig. 2) [19,27]. This signal peptide, consisting of 22 amino acids (66 nucleotides), was removed from the cDNA-encoding sequence of Sal k 6 to express the recombinant mature protein.

The comparison of the amino acid sequence of mature Sal k 6 with reported sequences of allergenic PGs deposited in the IUIS database (www.allergen.org) rendered highly variable percentages of identity and similarity depending on the plant families they belong (Fig. 3). Therefore, Pla a 2 and Pla or 2, both PGs from the Platanaceae family, exhibited 42 and 43% identity and 62 and 61% similarity with rSal k 6, respectively; whereas Phl p13 and Zea m 13 from the Poaceae family showed 37 and 38% identity and 56 and 58% similarity, respectively-; Cupressaceae Cry j 2 and Cha o 2 PGs showed the lowest percentages of identity - 24 and 23% - and similarity - 41 and 42% -, respectively (Fig. 3).

Despite the low identity percentages observed, the seven compared PGs shared some large amino acid segments highly conserved (Fig. 3) [28].

3.3. Expression, purification and structural characterization of recombinant Sal k 6

The cDNA-encoding sequence of mature Sal k 6 was subcloned into the pET41b vector and directly used to transform BL21(DE3) *E. coli* cells to produce the recombinant protein fused to a His-tag in the C-terminal end.

The recombinant protein –rSal k 6- remained insoluble at all conditions tested, with the highest amount of protein obtained at 16 °C after 48 h of induction with IPTG (Fig. 4A). After solubilization of inclusion bodies with chaotropic agents, we performed on-column refolding prior to elute the protein with an isocratic gradient of imidazole (Fig. 4B). Fractions containing the solubilized recombinant protein were pooled, dialyzed and lyophilized, obtaining a final yield of ~10 mg of purified soluble protein per litre of cell culture.

The protein was then subjected to far-UV CD experiments to determine if rSal k 6 was adequately refolded (Fig. 4C). A content on α-helix of 28.7%, β-sheet of 52.4% and a random coil of 18.7% was determined, which is in concordance with the high content in β-sheet of the PG protein family [29]. Accordingly, we confirmed that rSal k 6 was folded at secondary structure level (Fig. 4C).

3.4. Assignment of the 28 kDa band as a fragment derived from a high molecular mass allergenic polygalacturonase

PGs are enzymes comprising heterogeneous molecular masses ranging from 39 to 60 kDa. These data are in agreement with the electrophoretic mobility of the purified recombinant protein of about 47 kDa calculated by SDS-PAGE and the theoretical molecular mass of 39,554.2 Da calculated with the amino acid sequence derived from the mature Sal k 6-encoding cDNA. Then, we hypothesized that the 28 kDa protein spots used to obtain the PG sequence might be a shorter PG isozyme or a degradation product of a higher molecular mass PG present in the *S. kali* pollen extract.

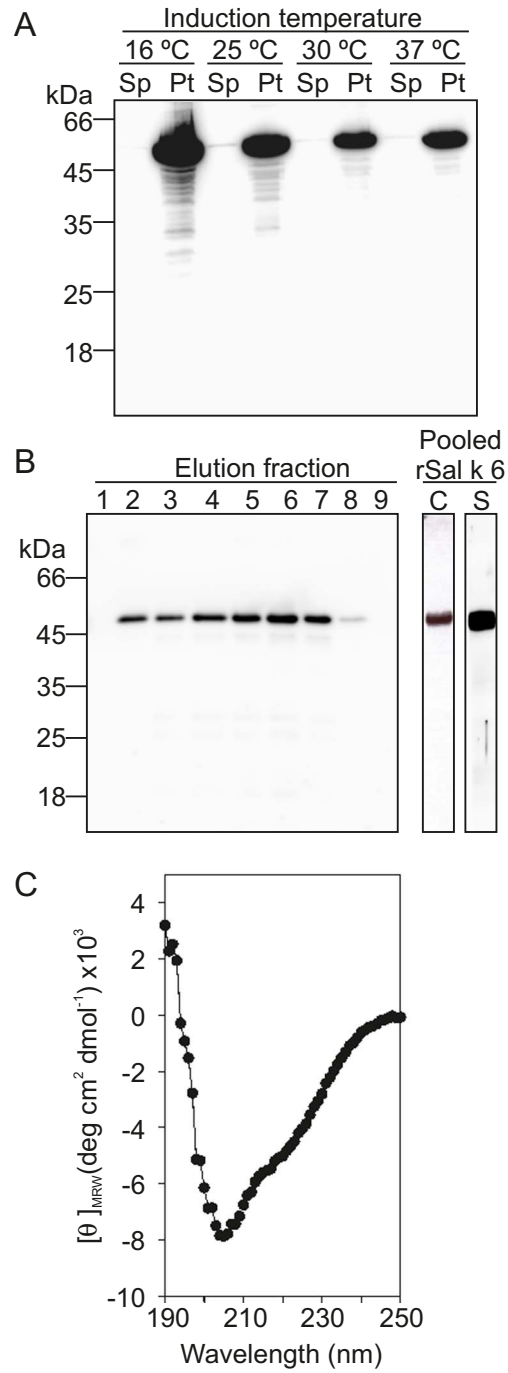


Fig. 4. Expression, purification and structural characterization of PG from *S. kali* pollen. (A) Coomassie Blue staining analysis of the recombinant expression of Sal k 6 in *E. coli* by SDS-PAGE at different temperatures and times of induction with 1 mM IPTG. Sp, supernatant. Pt, pellet. (B) Analysis of the purification by Ni-NTA affinity chromatography of rSal k 6 by SDS-PAGE and immunostaining. C, Coomassie Blue staining. S, pool of five sera. Molecular mass markers are shown in (A) and (B). (C) Circular dichroism spectra of rSal k 6.

To address this question, we deeply analyzed the *S. kali* pollen protein extract by i) immunostaining using the anti-Sal k 6 mouse pAb and inhibition assays using sera from sensitized patients and the recombinant protein, and ii) LC-MS/MS proteomics analysis of the 47 kDa and the 28 kDa forms from pollen. First, we performed an immunostaining of the *S. kali* pollen protein extract separated by 1DE gel with the mouse pAb raised against rSal k 6. Remarkably, rSal k 6

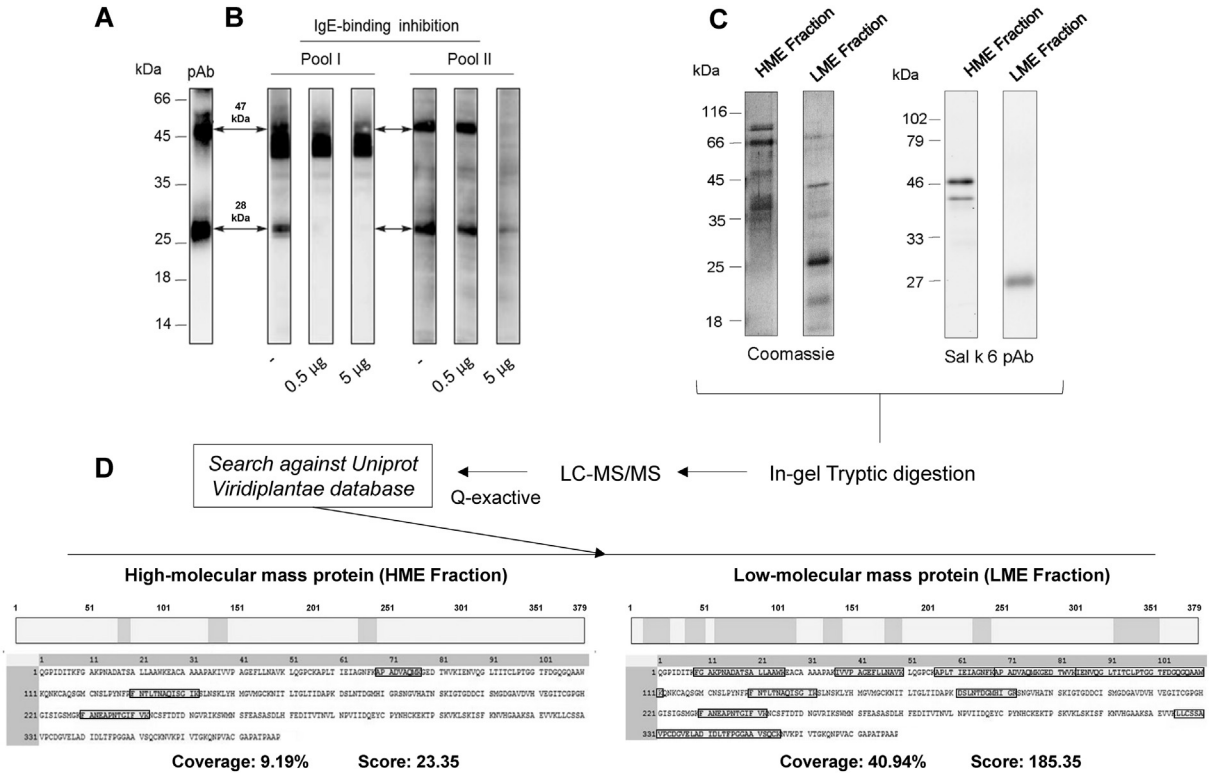


Fig. 5. Assignment of the 28 kDa band as a fragment of the integral PG of 47 kDa by immunostaining and proteomics analyses. (A) Recognition of the 47 and 28 kDa protein bands after staining with a pAb raised against the rSal k 6 allergen. (B) Inhibition of the IgE-binding capability of two Sal k 6-positive pool of sera (pool I: Sal k 1-positive and pool II: Sal k 1-negative sensitized pool of patients' sera) to natural PG from *S. kali* pollen using rSal k 6 as inhibitor (0.5 and 5 µg). (C) Coomassie Blue staining and immunostaining of the fractions enriched in high-molecular mass proteins (HME Fraction) and low-molecular-mass proteins (LME Fraction) containing the 47 kDa and the 28 kDa forms of Sal k 6, respectively. Fractions were obtained from the separation by size-exclusion chromatography of *S. kali* pollen protein extract (data not shown). (D) Scheme of the LC-MS/MS work flow together with the identified peptides in gray located at the primary amino acid sequence of Sal k 6 for both fractions enriched in high-molecular and low-molecular mass proteins. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

pAb was able to simultaneously recognize the protein band of 47 kDa and the low molecular band of 28 kDa (Fig. 5A). These results supported the hypothesis that the 28 kDa band is the main degradation product with IgE-binding ability of the integral *S. kali* pollen PG. To corroborate this observation, we performed IgE-inhibition assays of *S. kali* pollen protein extract using rSal k 6 as inhibitor and IgE from two pools of sera from: i) Sal k 6-positive patients and Sal k 1-negative patients, and ii) Sal k 6- and Sal k 1-positive patients to distinguish IgE reactivity to Sal k 6 and Sal k 1, which possess similar pI and molecular mass (Fig. 5B). Interestingly, the IgE-inhibition experiments showed that the specific IgEs from Sal k 6-positive sera directed to the bands of 28 and 47 kDa were completely abrogated when rSal k 6 was used as inhibitor. However, the inhibition of the specific IgEs from Sal k 6- and Sal k 1-positive patients with rSal k 6 could only completely inhibit the protein band of 28 kDa, whereas the band of 47 kDa was only partially inhibited due to the electrophoretic overlapping of these two allergens.

To further investigate these observations, we performed a proteomics analysis of the 47 kDa and the 28 kDa protein forms of Sal k 6 from *S. kali* pollen extract to confirm that both protein bands corresponded to the allergenic polygalacturonase. We first separated the pollen protein extract by size-exclusion chromatography and two fractions enriched in the 47 and the 28 kDa forms were in-gel digested with trypsin and subsequently analyzed by LC-MS/MS on a Q-exactive (Fig. 5C). Remarkably, we observed among the first three hits of both fractions the presence of Sal k 6 with 3 peptides for the 47 kDa protein band and 10 peptides for the 28 kDa protein band. In addition, we obtained for both fractions a coverage for Sal k 6 of 9.19% and 40.94%, and a score of 23.35 and 185.35, respectively (Fig. 5D, and Supplementary material, Table S1). Moreover, since 9 out of the 10 identified peptides for the 28 kDa protein band of Sal k 6 were located at the N-terminal of the protein, we might speculate that the natural degradation product of the 28 kDa form of Sal k 6 lacks the C-terminal end.

Collectively, these data confirmed that the protein band of 28 kDa with IgE-binding ability corresponds to a natural degradation product of the higher molecular mass allergen Sal k 6, and thus, validating the identification of a new *S. kali* allergen as a PG.

3.5. Immunological properties of rSal k 6

Eighty-eight sera from *S. kali* sensitized patients from three different regions of Spain (Zaragoza, Murcia and Alicante) were tested by ELISA for IgE reactivity against rSal k 6 (Fig. 6A-C). The average prevalence of rSal k 6 was calculated to be 30%, with prevalence values of 18, 39 and 39% for Alicante, Murcia and Zaragoza, respectively. Interestingly, these data correlate with the levels of grains/m^3 of Amaranthaceae pollen observed in the three populations (Supplementary material, Fig. S2). The highest values of prevalence and potency of Sal k 6, determined as the mean value of the positive sera, were observed in the population of Zaragoza, which possessed the highest amounts of grains/m^3 of *S. kali* pollen during its flowering period (Supplementary material, Fig. S2).

To complete the immunological characterization of rSal k 6, we tested the ability of the recombinant protein to activate basophils of *S. kali* sensitized patients (Fig. 6D). At 20 µg/mL of rSal k 6, we observed a positive activation in both allergic patients, and especially in one patient, in comparison to Alt a 1 allergen that was used as negative control in these non-sensitized patients to alternaria. The non-atopic control was non-responder at all concentrations and allergen tested.

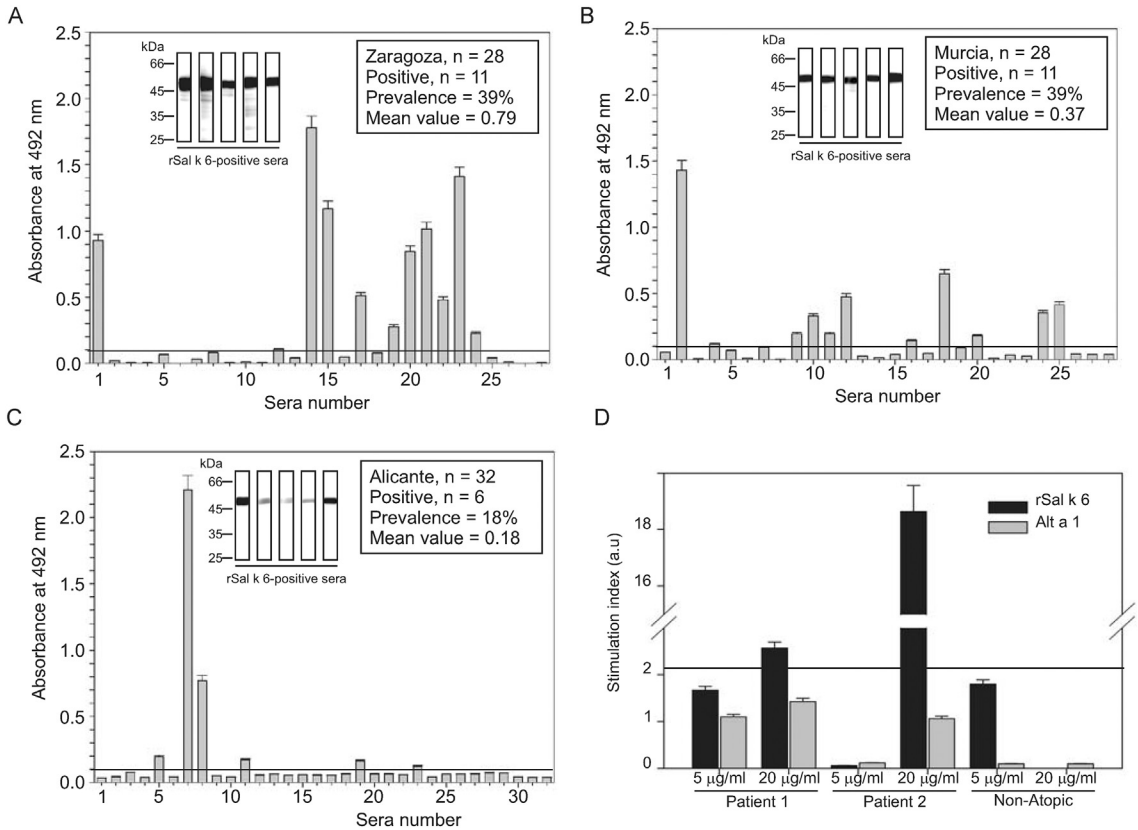


Fig. 6. Immunological characterization of Sal k 6. (A–C) Analysis of the sIgE levels of three populations of patients from Spain by ELISA and five randomly selected Sal k 6-positive sera from each population by immunoblotting (inset). Prevalence of the allergen is indicated. Potency of the allergen was calculated as the mean value of the positive sera. (D) Basophil activation assay. Stimulation index of patients allergic to *S. kali* pollen and the non-atopic control patient at different concentrations of the allergens are shown.

3.6. Cross-reactivity of Sal k 6 with PGs from pollen and plant-derived food

Due to the presence in all allergenic PGs of highly conserved regions, we hypothesized a role of Sal k 6 in cross-reactivity processes. To address this question, we tested the presence of PGs similar to Sal k 6 in different pollen and plant-derived food extracts by immunoblotting with the anti-rSal k 6 pAb (Supplementary material, Fig. S3). Remarkably, although we observed the presence of PGs sharing IgG epitopes with Sal k 6 in most pollen protein extracts at different extents, we observed an absence of PGs or the presence of PGs not sharing IgG epitopes with Sal k 6 in the plant-derived food extracts tested (Supplementary material, Fig. S3). Next, we performed IgG-inhibition assays using the anti-rSal k 6 pAb by immunoblotting or ELISA (Fig. 7A and B). All pollen protein extracts were able to inhibit the IgG-binding to rSal k 6 at different extents from 16% to 52% by immunoblotting and from 23% to 40% by ELISA, while *S. kali* pollen protein extract reached the highest IgG inhibition levels of 78% by immunoblotting and 100% by ELISA.

Then, we determined the potential role of PGs in IgE cross-reactivity with pollen and pollen-food cross-reactivity. To that end, we performed IgE-inhibition assays either by immunoblotting or ELISA (Fig. 7C and D). Only six out of the nine pollen protein extracts containing IgG-reactive PGs were able to inhibit the IgE-binding to rSal k 6 (Fig. 6C and D). The maximum levels of IgE-inhibition were observed for pollens from *Betula verrucosa* (79%), the Oleaceae family (*Fraxinus excelsior*, *Olea europaea* and *Syringa vulgaris*), with IgE-inhibition values of 37, 42, and 68%, and the Amaranthaceae *Chenopodium album* (34%). Interestingly, although PG enzymes have been described in *Cupressus arizonica*, *Platanus acerifolia* and *Lolium perenne,* no inhibition to IgE-binding to rSal k 6 was observed with these pollen protein extracts. In addition, no IgE-inhibition was observed with plant-derived food extracts (Supplementary material, Fig. S3), as expected from the absence of IgG reactive PGs in their extracts.

Collectively, these results support a role of rSal k 6 in pollen-pollen IgE cross-reactivity with related and non-related allergenic pollen sources.

4. Discussion and conclusions

S. kali is an important allergy-inducing agent in temperate countries such as Spain, Middle East or USA. Several factors such as the desertification process and the use of *S. kali* in greening programs and the pollution derived from industries have increased the presence of this weed worldwide. In Spain, this weed can be found in almost the whole territory and the number of sensitized patients to this pollen has largely increased in the recent years [3]. In some extensive regions of Spain - the South-East coast (*i.e.* Alicante and Murcia) and the Center-East (*i.e.* Zaragoza) - the presence of high number of pollen counts during its flowering period is frequent (Supplementary material, Fig. S2).

The knowledge of the allergenic components present in a particular biological source facilitates the diagnosis and the clinical treatment of patients by means of a more suitable standardization of the clinically used extracts. Although in *S. kali* pollen five allergens have been described so far [3], several relevant IgE-binding protein bands, which do not correspond to any of these allergens, have been observed by immunostaining of the *S. kali* protein extract using sera from *S. kali* pollen sensitized patients.

In this work, we have been able to identify by proteomics three IgE-reactive spots corresponding to a 28 kDa band as a fragment of a PG enzyme (EC 3.2.1.15) that belongs to the family 28 of glycosyl hydrolases. This group of enzymes displays heterogeneous molecular masses ranging from 39 to 60 kDa, and some of them have been described as allergens –PGs from grasses, plane tree and Japanese

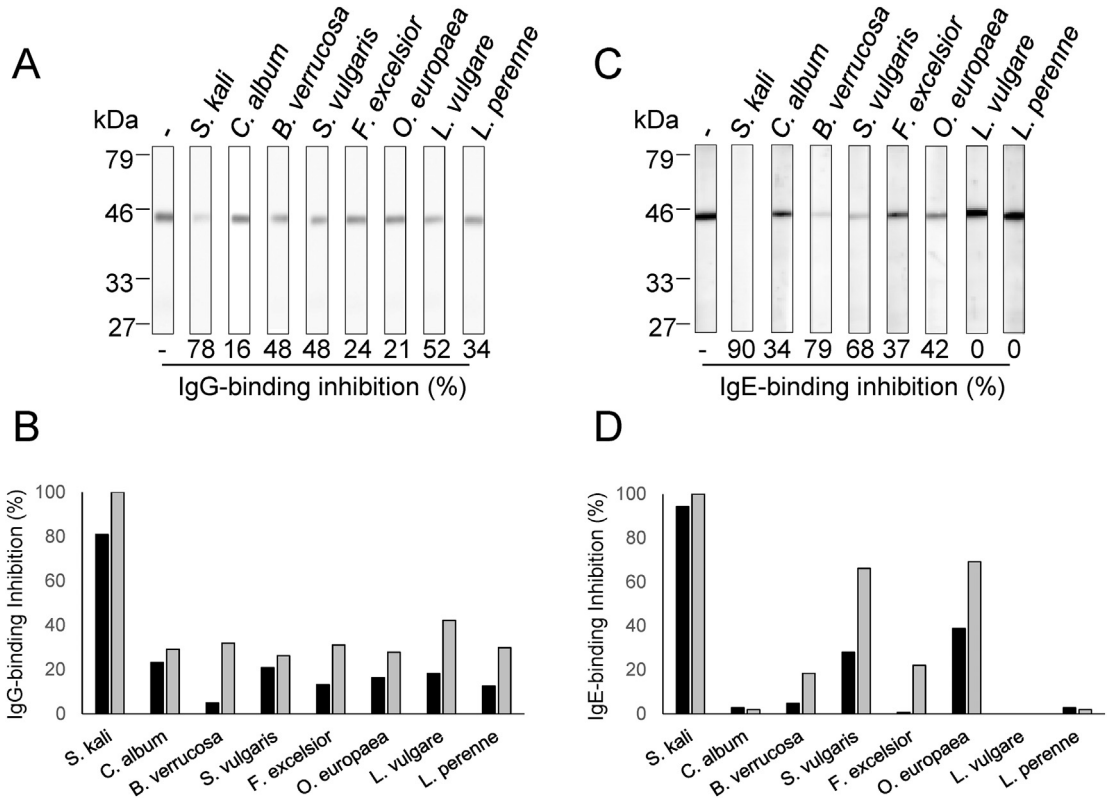


Fig. 7. Identification of IgE cross-reactive polygalacturonases from different pollen sources with Sal k 6. (A–B) Inhibition of the IgG-binding ability or (C–D) IgE-binding ability to 500 μg (A, C) or 20 μg and 200 μg (B, D) of rSal k 6 using different pollen protein extracts as inhibitors and an anti-rSal k 6 pAb (A, C) or a pool of five Sal k 6-positive sera (B, D) by immunoblotting (A, C) or ELISA (B, D). Values of inhibition in percentage were calculated as the signal lost in the samples inhibited regarding to the non-inhibited control (−) by immunoblotting or according to: [1 − ($OD_{492\ nm}$ with inhibitor / $OD_{492\ nm}$ without inhibitor)] × 100 by ELISA. Molecular mass markers are shown.

cypress- [14,17,18]. Although the molecular mass of the identified fragment did not correspond to these molecular masses, we demonstrated by proteomics and IgE inhibition experiments using rSal k 6 allergen produced in *E. coli* that this natural protein band unequivocally corresponded to a spontaneous degradation product from a PG with an electrophoretic mobility of 47 kDa. Similar degradation issues occurring by i) the presence of enzymes in extracts (*i.e.* in house dust mite extracts), ii) the instability of allergens or pollen protein extracts (*i.e.* cypress pollen extract) or iii) during the isolation procedure (*i.e.* food allergens) have been previously reported [30–32]. Other examples of the existence of lower molecular masses protein bands than the integral allergen occurred in allergenic protein bands with homology to domains of larger enzymes (Ole e 9 and Ole e 10 [33]) or allergenic proteolytic fragments derived from allergens of higher molecular masses (Ole e 4 and Ole e 9 [34]). Remarkably, the reported instability of other allergenic PGs - Phl p 13 and Zea m 13 [35,36] - is claimed by the authors as the reason for the later identification of the major allergen Phl p 13 in comparison to other major *Phleum pretense* allergens [35]. In *S. kali*, the intact Sal k 6 allergen has probably been remained hidden because the marker of sensitization of *S. kali* pollen –the PME Sal k 1- shares similar physicochemical properties: electrophoretic mobility, pI and a molecular mass of about 43 kDa [4]. Therefore, Sal k 6-positive patients could have been misidentified by SDS-PAGE as Sal k 1-positive patients, especially because about 10% of these patients were non-reactive to Sal k 1 by ELISA (personal observations).

Although Sal k 6 was observed to be a minor allergen in the three populations of sera tested with a prevalence ranging 19% and 39%, we confirmed the capability of rSal k 6 to induce basophils activation from blood of *S. kali* pollen-sensitized patients. Remarkably, we observed a clear correlation between the counts of grains of Amaranthaceae pollen in the atmosphere appearing in the three regions tested, and the prevalence and the sIgE levels to Sal k 6 in the sensitized patients (Fig. 6; Supplementary material, Fig. S2). In this sense, although rSal k 6 prevalence was similar in the sensitized patients from Zaragoza and Murcia, the potency of the allergen in Zaragoza was higher, which correlated with the highest counts of pollen observed in that region. On the other hand, sera from sensitized patients from Alicante showed the lowest prevalence and potency of the allergen, which again correlated with the lowest counts of grains of Amaranthaceae pollens. Collectively, these data seemed to indicate the high importance of the count levels of Amaranthaceae pollen with the potency and prevalence of Sal k 6.

Since PGs shared highly conserved regions, and Platanaceae, Poaceae and Cupressaceae plant families have been described to contain allergenic PGs [13–18], we investigated the potential association of Sal k 6 in pollen-pollen and pollen-plant food polysensitization. We have here identified that Oleaceae, Betulaceae and Amaranthaceae plant families contain allergenic pollen PGs, probably a subfamily of allergens apart from the previously identified allergenic PGs from Platanaceae, Poaceae and Cupressaceae plant families. Remarkably, *P. acerifolia, L. perenne* and *C. arizonica* were the only pollen protein extracts unable to cross-react at IgE level with rSal k 6. PGs from pollen protein extracts from the Oleaceae family and *B. verrucosa* showed the highest cross-reactive ability with rSal k 6. On the other hand, we observed an absence of potential PGs sharing IgG or IgE epitopes with Sal k 6 in plant-derived food extracts.

In conclusion, the production of the recombinant PG of *S. kali* has demonstrated to be a good approach to study the allergenic properties of this protein. Therefore, rSal k 6 becomes a good candidate to gain further insights in the allergenic properties of this family of allergens. The here presented work shows that Sal k 6 is a relevant allergen of *S. kali* pollen and is implicated in pollen-pollen cross-reactivity. Although neither IgG nor IgE cross-reactive PGs were found in plant-derived foods, we cannot discard the presence of allergenic PGs in other plant-

derived food extracts.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbapap.2017.05.007.

Funding

This work was supported by grants SAF2011-26716 and SAF2014-53209-R from the Ministerio de Economía y Competitividad and RIRAAF Network RD12/0013/0015 from the ISCIII. R.B. was a fellow of the Ramón y Cajal program of the Ministerio de Economía y Competitividad (Spain). C·O-S. is supported by a contract of the Programa Operativo de Empleo Juvenil y la Iniciativa de Empleo Juvenil (YEI) with the participation of the Consejería de Educación, Juventud y Deporte de la Comunidad de Madrid y del Fondo Social Europeo. The Proteomics facilities of the UCM and CIB-CSIC acknowledge the support of ProteoRed-ISCIII network.

Conflict of interest

The authors declare no conflict of interest.

Transparency document

The http://dx.doi.org/10.1016/j.bbapap.2017.05.007 associated with this article can be found, in online version.

Acknowledgements

We thank the excellent technical support of Sara Abián.

References

[1] D. Barber, F. de la Torre, F. Feo, F. Florido, P. Guardia, C. Moreno, J. Quiralte, M. Lombardero, M. Villalba, G. Salcedo, R. Rodríguez, Understanding patient sensitization profiles in complex pollen areas: A molecular epidemiological study, Allergy 63 (11) (2008) 1550–1558.

[2] C. Colás, S. Monzón, M. Venturini, A. Lezaun, Double-blind, placebo-controlled study with a modified therapeutic vaccine of *Salsola kali* (Russian thistle) administered through use of a cluster schedule, J. Allergy Clin. Immunol. 117 (4) (2006) 810–816.

[3] M. Villalba, R. Barderas, S. Mas, C. Colás, E. Batanero, R. Rodríguez, Amaranthaceae pollens: review of an emerging allergy in the Mediterranean area, J. Investig. Allergol. Clin. Immunol. 24 (6) (2014) 371–381.

[4] R. Barderas, J. Garcia-Selles, G. Salamanca, C. Colas, D. Barber, R. Rodriguez, M. Villalba, A pectin methylesterase as an allergenic marker for the sensitization to Russian thistle (*Salsola kali*) pollen, Clin. Exp. Allergy 37 (7) (2007) 1111–1119.

[5] M.A. Assarehzadegan, M. Sankian, F. Jabbari, M. Tehrani, R. Falak, A. Varasteh, Identification of methionine synthase (Sal k 3), as a novel allergen of *Salsola kali* pollen, Mol. Biol. Rep. 38 (1) (2011) 65–73.

[6] S. Mas, R. Barderas, C. Colas, J. Quiralte, R. Rodriguez, M. Villalba, The natural profilin from Russian thistle (*Salsola kali*) contains a low IgE-binding ability isoform–molecular and immunological characterization, FEBS J. 279 (23) (2012) 4338–4349.

[7] M.A. Assarehzadegan, A. Amini, M. Sankian, M. Tehrani, F. Jabbari, A. Varasteh, Sal k 4, a new allergen of *Salsola kali*, is profilin: A predictive value of conserved conformational regions in cross-reactivity with other plant-derived profilins, Biosci. Biotechnol. Biochem. 74 (7) (2010) 1441–1446.

[8] L. Castro, S. Mas, R. Barderas, C. Colás, J. García-Sellés, D. Barber, J. Rodríguez, M. Villalba, Sal k 5, a memeber of the widespread Ole e 1-like protein family, is a new allergen of Russian thistle (*Salsola kali*) pollen, Int. Arch. Allergy Immunol. 163 (2014) 142–153.

[9] K.A. Hadfield, A.B. Bennett, Polygalacturonases: many genes in search of a function, Plant Physiol. 117 (2) (1998) 337–343.

[10] A. Bezier, B. Lambert, F. Baillieul, Cloning of a grapevine botrytis-responsive gene that has homology to the tobacco hypersensitivity-related hsr203J, J. Exp. Bot. 53 (378) (2002) 2279–2280.

[11] I.G. Sandri, R.C. Fontana, D.M. Barfknecht, M.M. da Silveira, Clarification of fruit juices by fungal pectinases, LWT Food Sci. Technol. 44 (10) (2011) 2217–2222.

[12] D.R. Kashyap, P.K. Vohra, S. Chopra, R. Tewari, Applications of pectinases in the commercial sector: a review, Bioresour. Technol. 77 (3) (2001) 215–227.

[13] X. Zhou, Y. Zhang, K.J. Niklas, Sensitivity of growth and biomass allocation patterns to increasing nitrogen: a comparison between ephemerals and annuals in the Gurbantunggut Desert, north-western China, Ann. Bot. (2013).

[14] M. Namba, M. Kurose, K. Torigoe, K. Hino, Y. Taniguchi, S. Fukuda, M. Usui, M. Kurimoto, Molecular cloning of the second major allergen, Cry j II, from Japanese cedar pollen, FEBS Lett. 353 (2) (1994) 124–128.

[15] M. Yokoyama, M. Miyahara, K. Shimizu, K. Kino, H. Tsunoo, Purification, identification, and cDNA cloning of Jun a 2, the second major allergen of mountain cedar pollen, Biochem. Biophys. Res. Commun. 275 (1) (2000) 195–202.

[16] I. Ibarrola, M.C. Arilla, A. Martinez, J.A. Asturias, Identification of a polygalacturonase as a major allergen (Pla a 2) from *Platanus acerifolia* pollen, J. Allergy Clin. Immunol. 113 (6) (2004) 1185–1191.

[17] J.M. Davies, A. Voskamp, T.D. Dang, B. Pettit, D. Loo, A. Petersen, M.M. Hill, J.W. Upham, J.M. Rolland, R.E. O'Hehir, The dominant 55 kDa allergen of the subtropical Bahia grass (*Paspalum notatum*) pollen is a group 13 pollen allergen, Pas n 13, Mol. Immunol. 48 (6–7) (2011) 931–940.

[18] R. Suck, A. Petersen, S. Hagen, O. Cromwell, W.M. Becker, H. Fiebig, Complementary DNA cloning and expression of a newly recognized high molecular mass allergen phl p 13 from timothy grass pollen (*Phleum pratense*), Clin. Exp. Allergy 30 (3) (2000) 324–332.

[19] T.N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: discriminating signal peptides from transmembrane regions, Nat. Methods 8 (10) (2011) 785–786.

[20] S. Mas, P. Boissy, R.I. Monsalve, J. Cuesta-Herranz, A. Diaz-Perales, J. Fernandez, C. Colas, R. Rodriguez, R. Barderas, M. Villalba, A recombinant Sal k 1 isoform as an alternative to the polymorphic allergen from *Salsola kali* pollen for allergy diagnosis, Int. Arch. Allergy Immunol. 167 (2) (2015) 83–93.

[21] P. Barral, M. Villalba, R. Rodriguez, E. Batanero, The role of major olive pollen allergens Ole e 1, Ole e 9, and Ole e 10 on mice sensitization, Ann. Allergy Asthma Immunol. 96 (3) (2006) 466–471.

[22] A. Kloczkowski, K.L. Ting, R.L. Jernigan, J. Garnier, Combining the GOR V algorithm with evolutionary information for protein secondary structure prediction from amino acid sequence, Proteins 49 (2) (2002) 154–166.

[23] A. Pelaez-Garcia, R. Barderas, R. Batlle, R. Vinas-Castells, R.A. Bartolome, S. Torres, M. Mendes, M. Lopez-Lucendo, R. Mazzolini, F. Bonilla, A. Garcia de Herreros, J.I. Casal, A proteomic analysis reveals that Snail regulates the expression of the nuclear orphan receptor nuclear receptor subfamily 2 group F member 6 (Nr2f6) and interleukin 17 (IL-17) to inhibit adipocyte differentiation, Mol. Cell. Proteomics 14 (2) (2015) 303–315.

[24] A. Pelaez-Garcia, R. Barderas, M. Mendes, M. Lopez-Lucendo, J.C. Sanchez, A. Garcia de Herreros, J.I. Casal, Data from proteomic characterization of the role of Snail1 in murine mesenchymal stem cells and 3T3-L1 fibroblasts differentiation, Data Brief 4 (2015) 606–613.

[25] L. Kall, J.D. Canterbury, J. Weston, W.S. Noble, M.J. MacCoss, Semi-supervised learning for peptide identification from shotgun proteomics datasets, Nat. Methods 4 (11) (2007) 923–925.

[26] A. Diaz-Perales, M.L. Sanz, G. Garcia-Casado, R. Sanchez-Monge, F.J. Garcia-Selles, M. Lombardero, F. Polo, P.M. Gamboa, D. Barber, G. Salcedo, Recombinant Pru p 3 and natural Pru p 3, a major peach allergen, show equivalent immunologic reactivity: a new tool for the diagnosis of fruit allergy, J. Allergy Clin. Immunol. 111 (3) (2003) 628–633.

[27] M.R. Wilkins, E. Gasteiger, A. Bairoch, J.C. Sanchez, K.L. Williams, R.D. Appel, D.F. Hochstrasser, Protein identification and analysis tools in the ExPASy server, Methods Mol. Biol. 112 (1999) 531–552.

[28] Y. van Santen, J.A. Benen, K.H. Schroter, K.H. Kalk, S. Armand, J. Visser, B.W. Dijkstra, 1.68-A crystal structure of endopolygalacturonase II from *Aspergillus niger* and identification of active site residues by site-directed mutagenesis, J. Biol. Chem. 274 (43) (1999) 30474–30480.

[29] H.J. Rozeboom, G. Beldman, H.A. Schols, B.W. Dijkstra, Crystal structure of endo-xylogalacturonan hydrolase from *Aspergillus tubingensis*, FEBS J. 280 (23) (2013) 6061–6069.

[30] Y. Shahali, Z. Pourpak, M. Moin, A. Mari, A. Majd, Instability of the structure and allergenic protein content in Arizona cypress pollen, Allergy 64 (12) (2009) 1773–1779.

[31] R. Falak, M. Sankian, R. Noorbakhsh, M. Tehrani, M.A. Assarehzadegan, F.J. Azad, A. Abolhasani, A.R. Varasteh, Identification and characterisation of main allergic proteins in *Vitis vinifera* Vitis, Food Agric. Immunol. 24 (3) (2013) 255–268.

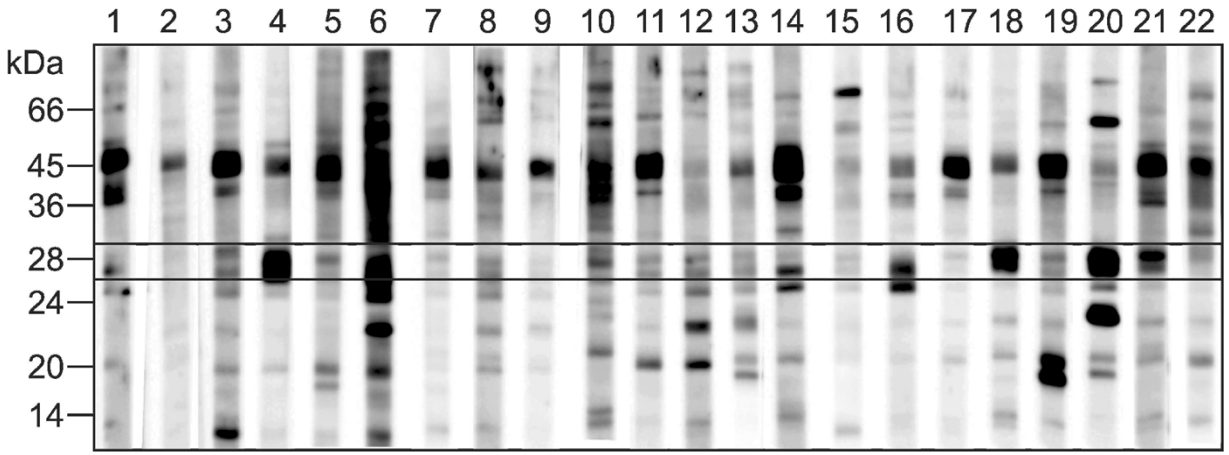
[32] L.N. Soldatova, E.J. Paupore, S.H. Burk, R.W. Pastor, J.E. Slater, The stability of house dust mite allergens in glycerinated extracts, J. Allergy Clin. Immunol. 105 (3) (2000) 482–488.

[33] P. Barral, E. Batanero, O. Palomares, J. Quiralte, M. Villalba, R. Rodriguez, A major allergen from pollen defines a novel family of plant proteins and shows intra- and interspecie cross-reactivity, J. Immunol. 172 (6) (2004) 3644–3651.

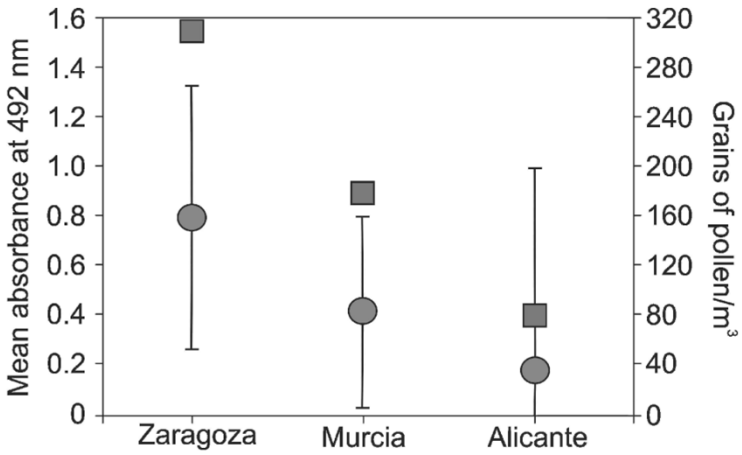
[34] M. Villalba, R. Rodriguez, E. Batanero, The spectrum of olive pollen allergens. From structures to diagnosis and treatment, Methods 66 (1) (2014) 44–54.

[35] A. Petersen, R. Suck, S. Hagen, O. Cromwell, H. Fiebig, W.M. Becker, Group 13 grass allergens: Structural variability between different grass species and analysis of proteolytic stability, J. Allergy Clin. Immunol. 107 (5) (2001) 856–862.

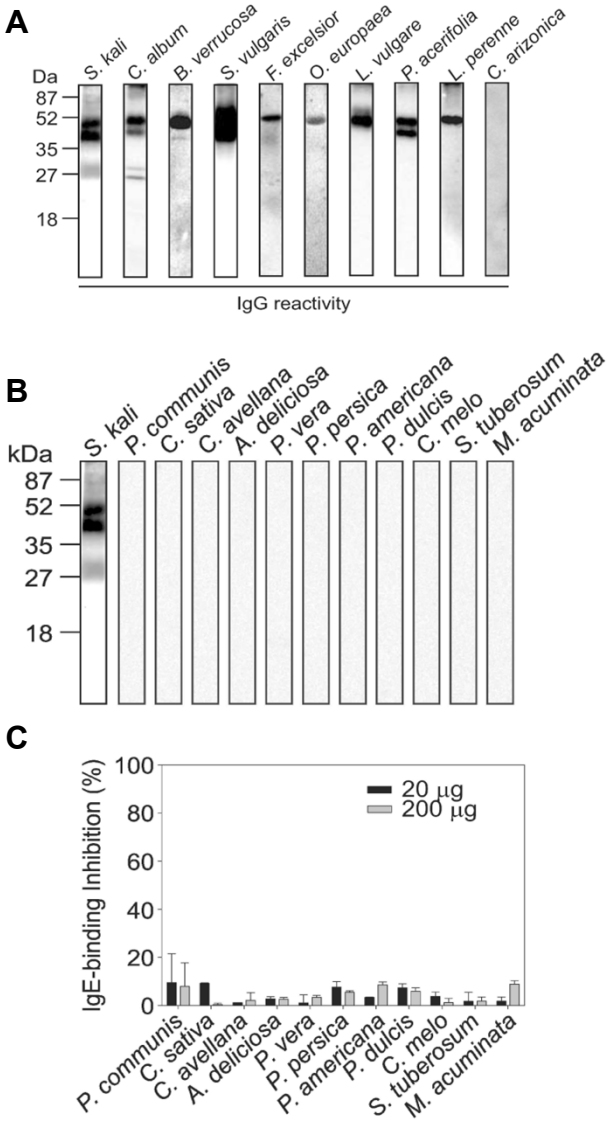
[36] I. Swoboda, M. Grote, P. Verdino, W. Keller, M.B. Singh, N. De Weerd, W.R. Sperr, P. Valent, N. Balic, R. Reichelt, R. Suck, H. Fiebig, R. Valenta, S. Spitzauer, Molecular characterization of polygalacturonases as grass pollen-specific marker allergens: expulsion from pollen via submicronic respirable particles, J. Immunol. 172 (10) (2004) 6490–6500.



Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3