# Identification of the cysteine protease Amb a 11 as a novel major allergen from short ragweed

**Julien Bouley, PhD,* Rachel Groeme, MSc,* Maxime Le Mignon, PhD, Karine Jain, MSc, Henri Chabre, PhD, Véronique Bordas-Le Floch, PhD, Marie-Noëlle Couret, PhD, Laetitia Bussières, MSc, Aurélie Lautrette, PhD, Marie Naveau, PhD, Véronique Baron-Bodo, PhD, Vincent Lombardi, PhD, Laurent Mascarell, PhD, Thierry Batard, PhD, Emmanuel Nony, PhD, and Philippe Moingeon, PhD** *Antony, France*

**Background: Allergy to pollen from short ragweed *(Ambrosia artemisiifolia)* is a serious and expanding health problem in the United States and in Europe.**
**Objective: We sought to investigate the presence of undescribed allergens in ragweed pollen.**
**Methods: Ragweed pollen proteins were submitted to high-resolution gel electrophoresis and tested for IgE reactivity by using sera from 92 American or European donors with ragweed allergy. Pollen transcriptome sequencing, mass spectrometry (MS), and recombinant DNA technologies were applied to characterize new IgE-binding proteins.**
**Results: High-resolution IgE immunoblotting experiments revealed that 50 (54%) of 92 patients with ragweed allergy were sensitized to a 37-kDa allergen distinct from Amb a 1. The full-length cDNA sequence for this molecule was obtained by means of PCR cloning after MS sequencing of the protein combined with ragweed pollen RNA sequencing. The purified allergen, termed Amb a 11, was fully characterized by MS and confirmed to react with IgEs from 66% of patients. This molecule is a 262-amino-acid thiol protease of the papain family expressed as a combination of isoforms and glycoforms after proteolytic removal of N- and C-terminal propeptides from a proform. Three-dimensional modeling revealed a high structural homology with known cysteine proteases, including the mite Der p 1 allergen. The protease activity of Amb a 11, as well as its capacity to activate basophils from patients with ragweed allergy, were confirmed. The production of a nonglycosylated recombinant form of Amb a 11 in *Escherichia coli* established that glycosylation is not required for IgE binding.**
**Conclusion: We identified the cysteine protease Amb a 11 as a new major allergen from ragweed pollen. Given the similar physicochemical properties shared by the 2 major allergens, we hypothesize that part of the allergenic activity previously ascribed to Amb a 1 is rather borne by Amb a 11. (J Allergy Clin Immunol 2015;136:1055-64.)**

***Key words:*** *Allergen, Amb a 11,* Ambrosia artemisiifolia*, cysteine protease, mass spectrometry, short ragweed*

*Abbreviations used*
AIT: Allergen immunotherapy
2D: Two-dimensional
IUIS: International Union of Immunological Societies
MS: Mass spectrometry
MS/MS: Tandem mass spectrometry
MW: Molecular weight
nanoLC: Nanoflow liquid chromatography
pAb: Polyclonal antibody
PDB: Protein Data Bank
RACE: Rapid amplification of cDNA ends

Pollen allergens from short ragweed (*Ambrosia artemisiifolia* var. *elatior*) cause severe type I respiratory allergies.[1] Short ragweed pollen exposure has been consistently and significantly expanding in North America and Europe,[2] possibly as a consequence of climate change.[3] As of today, allergen immunotherapy (AIT) with ragweed pollen extract is the only treatment option available to patients to address the underlying mechanisms of allergic inflammation. Specifically, AIT reduces both rhinoconjunctivitis symptoms and use of symptomatic medications as a consequence of a reorientation of inappropriate ragweed pollen–specific $T_H2$ responses.[4-9] Importantly, to be efficacious, allergen extracts used for AIT should contain all major/relevant allergens causing allergic inflammation in patients.

To date, up to 10 distinct allergens have been identified in *A artemisiifolia* pollen, according to the existing record of the International Union of Immunological Societies (IUIS) allergen nomenclature subcommittee.[10] Among those, the pectate lyase Amb a 1 is currently considered the most allergenic molecule, with greater than 90% of patients with ragweed allergy sensitized to it.[11-13] Ragweed pollen also contains minor allergens, such as Amb a 3 and Amb a 7 (plastocyanins),[14] Amb a 4 (homolog to the major mugwort allergen Art v 1),[15] Amb a 5 (an approximately 5-kDa allergen),[16] Amb a 6 (a nonspecific lipid transfer protein), Amb a 8 (profilin), and Amb a 9 and Amb a 10 (calcium-binding proteins) panallergens.[17] Importantly, however, there is clear evidence that some patients with ragweed allergy exhibit a strong IgE reactivity against various ragweed pollen proteins distinct from known allergens (unpublished results).[11,12]

In this context we investigated the presence of unidentified ragweed allergens, combining high-resolution gel electrophoresis and mass spectrometric (MS) analysis of pollen proteins exhibiting IgE reactivity. We report the identification and characterization of a new major allergen belonging to the cysteine protease family, which is now recorded as Amb a 11.0101 in the IUIS official nomenclature.

From Research & Pharmaceutical Development, Stallergenes.
*These authors contributed equally to this work.
Disclosure of potential conflict of interest: J. Bouley declares no relevant conflicts of interest. The rest of the authors are employed by Stallergenes.
Received for publication March 28, 2014; revised February 16, 2015; accepted for publication March 2, 2015.
Available online April 10, 2015.
Corresponding author: Philippe Moingeon, PhD, Research and Development, Stallergenes, 6 rue Alexis de Tocqueville, 92183 Antony Cedex, France. E-mail: pmoingeon@stallergenes.com.
0091-6749/$36.00

http://dx.doi.org/10.1016/j.jaci.2015.03.001

# METHODS

## Patient population

Sera from both European and American allergic donors were selected for proteomics experiments. European patients (n = 69) included in a phase I clinical study (ClinicalTrials.gov identifier NCT01224834) after obtaining informed consent were selected based on a history of symptoms associated with seasonal ragweed pollen allergy, a positive skin prick test response to ragweed pollen, and specific IgE levels to ragweed pollen of greater than 0.70 kU/L, as measured by means of ImmunoCAP (Thermo Scientific, Uppsala, Sweden). Plasma samples from American patients given a diagnosis of ragweed allergy (n = 23, IgE levels to ragweed pollen >1 kU/L) were purchased from PlasmaLab International (Everett, Wash).

## Identification of proteins reactive to human IgE

Total proteins were extracted from ragweed pollen (Greer Laboratories, Lenoir, NC) by using ammonium bicarbonate, fractionated by means of 1-dimensional (n = 23) or 2-dimensional (2D; n = 69) PAGE and then blotted onto a nitrocellulose membrane before probing with individual sera from donors with ragweed allergy for IgE reactivity or with a polyclonal antibody (pAb), as described in the Methods section in this article's Online Repository at www.jacionline.org. Protein spots were excised from 2D gels by using an EXquest spot cutter (Bio-Rad Laboratories, Marnes-la-Coquette, France), processed by means of tryptic in-gel digestion, and analyzed by using MS, as described in the Methods section in this article's Online Repository. A combination of 2D PAGE, immunoblotting, and Edman sequencing was used to determine the N-terminal amino acid sequence of Amb a 11.[18] The conjugated N-terminal peptide of Amb a 11 (GSAPGSIDTDPNKDF) was used to immunize rabbits and produce anti–Amb a 11 pAb (AGRO-BIO, La-Ferté-Saint-Aubin, France). Basophil activation tests were performed, as described in the Methods section in this article's Online Repository.

## Ragweed pollen RNA sequencing and Amb a 11 cDNA cloning

Deep mRNA sequencing was performed by Beckman Coulter Genomics (Grenoble, France) with a 454 sequencer using titanium chemistry. Up to 43,149 entries were obtained after *de novo* assembly and open reading frame translation with Newbler software (Roche Diagnostics, Meylan, France).

Total RNA was used as a template for 5′ and 3′ rapid amplification of cDNA ends (RACE) PCR experiments using the SMARTer cDNA synthesis kit (Clontech, Saint-Germain-en-Laye, France), as described in the Methods section in this article's Online Repository.

## Purification and analysis of nAmb a 11 and rAmb a 11

The mature form of nAmb a 11 was purified by means of off-gel electrophoresis (Agilent Technologies, Les-Ulis, France), followed by reversed-phase HPLC separation with a high-recovery macroporous reversed-phase $C_{18}$ column (Agilent Technologies, Santa Clara, Calif), as described in the Methods section in this article's Online Repository. Purified nAmb a 11.0101 was fully characterized by means of nanoflow liquid chromatography (nanoLC)-MS(/MS). Full-length Amb a 11.0101 was produced as a recombinant molecule in *Escherichia coli* and tested for enzymatic activity, as described in the Methods section in this article's Online Repository.

## Bioinformatic analysis and Amb a 11 3-dimensional structure modeling

The following sequences were obtained by running BLAST and PSI-BLAST against the nrNCBI and UniProt databases with the Amb a 11 amino acid sequence as a query: the plant-derived food allergens Act d 1 (ACTN_ACTDE), Ana c 2 (BROM1_ANACO), the cysteine protease-homolog from *Carica papaya* (PAPA1_CARPA, Papain), and the mite allergens Der f 1 (PEPT1_DERFA), Der p 1 (PEPT1_DERPT), and Der p 1 homolog from *Aleuroglyphus ovatus* (A7UNT9_9ACAR, grain mite group 1 allergen). Alignments were made with CLUSTALW and ESPript[19] software. Sequences that were not annotated as "allergen" were removed. Geno3D[20] was used as an automatic Web server for protein molecular modeling. The 3-dimensional crystal structure of vignain (Protein Data Bank [PDB] ID 1S4V) determined by using x-ray diffraction[21] was used as a template structure. The PyMOL software (Schrödinger, Mannheim, Germany) was used to display the 3-dimensional structure of Amb a 11.

# RESULTS

## Evidence for a new major ragweed pollen allergen

Proteins from an aqueous extract were separated by means of 2D gel electrophoresis, stained with a fluorescent dye (CyDye), and immunoblotted by using individual sera from 69 donors with ragweed allergy (European, n = 46; American, n = 23) to identify new ragweed pollen allergens (Fig 1, *A-D*). All IgE-reactive protein spots were retrieved from the gels, trypsin digested, and submitted to tandem mass spectrometry (MS/MS). Among 19 spots analyzed, 14 could be assigned to known ragweed allergens (ie, Amb a 1, Amb a 5, and Amb a 8; see Table E1 in this article's Online Repository at www.jacionline.org). The amino acid sequence of the first 15 N-terminal residues of spot number 14 was identified as GSAPGSIDTDPNKDF, without any homology to known proteins. As expected, spots corresponding to the major allergen Amb a 1 (spots 2-10, 12, 13, and 16 in Fig 1, *B*) were confirmed to react to serum IgE from a majority of allergic patients, with 59 (86%) of 69 sera reacting to at least 1 Amb a 1 isoallergen. The 2D immunoblot analysis confirmed that the most prevalent allergens in patient sensitization include Amb a 1.01 and 1.03 isoforms (spots 2, 3, and 6-9; Fig 1, *B* and *C*). Protein spots reactive to serum IgE from approximately 15% of subjects were classified as minor ragweed allergens (ie, spots 1, 11, and 18, as well as spots 17 and 19 corresponding to Amb a 8 and Amb a 5, respectively). It is noteworthy that spots 14 and 15 were reactive to serum IgE from 37 (54%) of 69 donors (Fig 1, *C* and *D*). The same tryptic peptide sequences were obtained by using MS/MS for both spots 14 and 15, as well as for 2 more acidic spots, establishing that they represent a unique unknown protein. On the basis of an extensive characterization illustrated below, the latter was officially recorded as ragweed pollen allergen Amb a 11.0101 by the IUIS allergen nomenclature subcommittee. Collectively, this seroepidemiologic analysis documents 4 prominent patterns of IgE sensitization among patients, involving either Amb a 1 alone (Fig 1, *B*), multiple allergens (data not shown), predominantly Amb a 1 and Amb a 11 (Fig 1, *C*) or mostly Amb a 11 (Fig 1, *D*), representing around 40%, 15%, 30%, and 15% of patients, respectively.

In our 1-dimensional immunoblotting assays with sera from 23 additional European donors with ragweed allergy, Amb a 1 and Amb a 11 resolved as distinct bands, with evidence for a strong reactivity with patient's IgE for both allergens (Fig 1, *E*, lane 2). By using a specific rabbit pAb directed against Amb a 11, the latter was essentially observed as a mature form with an apparent molecular weight (MW) of approximately 37 kDa on SDS-PAGE and a minor band at approximately 52 kDa, likely corresponding to a proform (Fig 1, *E*, lane 1). As determined by using SDS-PAGE, the observed MW of Amb a 11 is larger than expected based on the amino acid sequence shown below, likely because of a glycosylated N-terminal domain, rigid (proline rich) C-terminus, or both.