

Cloning and Expression of Immunologically Active Recombinant *Amb a V* Allergen of Short Ragweed (*Ambrosia artemisiifolia*) Pollen¹

Balaram Ghosh,² Michael P. Perry, Thorunn Rafnar, and David G. Marsh

Division of Clinical Immunology, Johns Hopkins Asthma and Allergy Center, Johns Hopkins University School of Medicine, Baltimore, MD 21224

ABSTRACT. We have cloned and sequenced *Amb a V*, an *Ambrosia artemisiifolia* (short ragweed) pollen allergen that has proved to be particularly useful in the genetic analysis of human immune responsiveness. The amino acid sequence deduced from the cloned cDNA sequence corresponds to the published sequence of the protein, except that the cDNA sequence encodes an extra 10 amino acids at the C-terminus. The expressed 55-residue protein is then presumably cleaved enzymatically at the C-terminal lysine found in the 45-residue protein isolated from the pollen. The cloning and sequencing of *Amb a V* genomic DNA confirmed the cDNA sequence and showed that the *Amb a V* gene has no introns. Recombinant *Amb a V* allergen, expressed in *Escherichia coli*, bound to IgG and IgE antibodies in all *Amb a V*-allergic individuals tested and inhibition studies demonstrated that the recombinant protein contains a subset of the antigenic epitopes found on native *Amb a V*. In addition, recombinant *Amb a V* released histamine efficiently from basophils from *Amb a V*-allergic patients. The recombinant *Amb a V* allergen and mutants of *Amb a V* should, therefore, be useful in studies of allergen epitopes in humans, as well as providing a diagnostic tool. *Journal of Immunology*, 1993, 150: 5391.

A large proportion (at least 15–20%) of the populations in the developed countries suffers from IgE-mediated atopic diseases such as allergic rhinitis and bronchial asthma (1, 2). Among the wind-pollinated weeds, ragweed pollens represent a major source of allergenic proteins in the United States (3). The most extensively studied pollen is short ragweed (*Ambrosia artemisiifolia*), from which several allergens have been purified and characterized (4–10). *Amb a V*³ (Ra5) and its *Amb V* homologues, *Amb t V* (from giant ragweed) and

Amb p V (from western ragweed), are proteins of low relative m.w. (4400–5000), which have been studied as model systems because of the striking associations between immune responsiveness to these proteins and the atopic subject possessing the HLA-DRB1*1501 allele and the expressed HLA-DR2/Dw2 specificities (11–14). *Amb a V*-specific T cells have recently been cloned from ragweed-allergic individuals. In experiments using these T cells, and by DNA sequencing and DNA typing of HLA-D region genes, it has been shown that immune responsiveness to the *Amb a V* allergen is primarily restricted to the HLA-DR (α , β 1 *1501) molecule (formerly DR2b or DR $\alpha\beta$ 1 2.2) (15, 16). These data provide strong evidence for a causal relationship between the presence of the DR (α , β 1 *1501) molecule and responsiveness toward the *Amb V* allergens.

Cloning and sequencing of the cDNA encoding these allergens is important: 1) for the production and purification of large quantities of allergen proteins, and 2) to perform site-directed mutagenesis and express these mutant allergens for the purpose of studying the molecular mechanism of their interactions with MHC class II molecules, TCR and antibodies. We report the molecular cloning and sequence analysis of the *Amb a V* allergen and also dem-

Received for publication August 10, 1992. Accepted for publication March 25, 1993.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¹ This work was supported by Research Grant AI-19727 from the National Institutes of Health and a Johns Hopkins Institutional Grant.

² Address correspondence and reprint requests to Dr. B. Ghosh, Division of Clinical Immunology, Johns Hopkins Asthma and Allergy Center, Johns Hopkins University School of Medicine, 5501 Hopkins Bayview Circle, Baltimore, MD 21224.

³ Abbreviations used in this paper: *Amb a V*, allergen from *Ambrosia artemisiifolia*; *Amb t V*, allergen from *Ambrosia trifida*; PCR, polymerase chain reaction; HLA, human leukocyte Ag; DARIA, double antibody radioimmunoassay; GST, glutathione S-transferase; BBS, borate-buffered saline.

Table I
Oligonucleotides used for reverse-transcription and/or PCR amplification

Name ^a	Sequence ^b	Derived from <i>Amb a V</i> aa Positions ^c
BG15	5'-AGgaattcC/TIGTICCTGc/TGCITGGGC	1-7
BG16	5'-AGgaattcc/TTICCG/ACAT/cTTT/cTTTc/cA/cA	40-35
BG18	5'-T/cTCG/ATAG/ACAACIAC/cTGCCAIGGc/ACA	34-26
BG20	5'-GAgattcGTCGACATCGA(T) ₁₇	
BG21	5'-GAgattcGTCGACATCGAT	
BG22	5'-GCATATTGTTGTAGCGACCCA	16-22
BG24	5'-ACCACTTGCCAAGGACAGTACCG	31-24
BG28	5'-AGgaattcGTTGTGGTAAAAGCGTG	10-16
BG29	5'-AGgaattcTATTTCGCTGGATTCTGTAAC	38-32
BG31	5'-GAgattcG/cGTTCTTAGTGACGTTCA	
BG33	5'-AGgaattcATTACTTTGTTAATTTTA	

^a The degenerate nucleotides are shown in small caps. Inosine (I) was included at positions of fourfold degeneracy. Oligonucleotides were synthesized by the phosphoramidite method using a DNA synthesizer (Milligen/Bioscience) and purified by HPLC at the Johns Hopkins School of Hygiene and Public Health.

^b Sequences of oligonucleotides used in this study. *EcoRI* restriction sites are shown in lower-case letters.

^c The amino acid positions of *Amb a V* to which the oligonucleotides correspond. For the oligonucleotides BG15, BG24, BG28, and BG29, only parts of the codons of the aas at positions 7, 31, 16, and 32, respectively, are used (see Fig. 2).

onstrate that the expressed recombinant *Amb a V* (r*Amb a V*) allergen contains a subset of the epitopes found on native *Amb a V* (n*Amb a V*) and binds to IgE and IgG antibodies from allergic individuals. Furthermore, r*Amb a V* releases histamine from basophils of *Amb a V*-allergic individuals.

Materials and Methods

Reverse transcription and PCR

The overall strategy of PCR amplification is schematically shown in Figure 1 and is similar to that described in detail elsewhere (17). Total RNA was isolated from short ragweed pollen (Greer Labs, Lenoir, NC) using a modified procedure of plant RNA isolation (17). sscDNA was synthesized from the short ragweed total RNA (17). Two degenerate primers BG15 and BG16 were synthesized based on the N- and C-terminal sequence information, respectively (Table I). Because the codon usage of this plant species was not known, we synthesized a mixture of oligonucleotides; to reduce the degeneracy, inosine (I) was introduced at positions of fourfold degeneracy. The sscDNA (1.5 μ l) was amplified through 30 PCR cycles in the presence of 75 pmol each of BG15 and BG16 primers (Table I)/50 mM KCl/10 mM Tris-HCl pH 8.4/1.5 to 4.0 mM MgCl₂/0.01% gelatin/0.7 mM dNTP as described (18). The cDNA was initially denatured for 5 min at 95°C followed by the addition of 2.5 units of *Taq* polymerase (Perkin-Elmer/Cetus, Emeryville, CA). The conditions for PCR amplification were as follows: denaturation at 92°C for 2.0 min, annealing at 47°C for 1 min, and extension at 72°C for 1.5 min. At the end of the last cycle, the extension was carried out for an additional 10 min at 72°C. The amplified DNA products (5–10 μ l) were analyzed by electrophoresis on an agarose gel (NuSieve, GTG agarose, FMC Bioproducts, Rockland, ME) and a Southern blot hybridization was performed with the oligonucleotide, BG18. The expected 151-nucleotide fragment was isolated from the agarose gel and subcloned

into M13 mp10 for sequencing. Based on this sequence, a number of *Amb a V*-specific oligonucleotides (BG22, BG24, BG28, and BG29) were synthesized (Table I).

3'-end amplification by PCR

Total RNA was mixed with 1 μ l (0.5 μ g/ μ l) (dT)₁₇-adapter-primer (BG20), heated at 65°C for 2 min and quenched on ice. The reverse transcription was performed as described (17). The reaction mixture was diluted to 500 μ l with TE (10 mM Tris-HCl pH 7.4, 1mM EDTA buffer and stored at 4°C. Two microliters of the cDNA pool were amplified under the same conditions as described before, except for the following changes: 1) 30 pmol each of BG21(adapter) and BG28 (*Amb a V* specific) primer were used, 2) 2 mM of each dNTP was used, and 3) annealing was performed at 50°C. The amplified product (1–2 μ l) was then reamplified with 30 pmol each of BG22 and BG21 under the above conditions. These PCR-amplified fragments were analyzed by agarose gel electrophoresis and a Southern blot hybridization was performed with an internal oligonucleotide, BG24.

5'-end amplification by anchored-PCR

Total RNA was reverse-transcribed using the *Amb a V* gene-specific primer BG16 in a total volume of 60 μ l as described above, and the resultant cDNA was diluted to 2 ml with water. The excess primer was removed by Centricon 30 spin-filtration (Amicon Corp., Danvers, MA) and a poly (A) "tail" was added as described (17). The resultant 5'-cDNA pool (10 μ l) was amplified under the same conditions as described for the 3'-end amplification except for the following changes: 1) 5 pmol of BG20 (dT₁₇- adapter), 30 pmol of BG21 (adapter-primer), and 30 pmol of BG29 (*Amb a V*-specific primer) were used, 2) 1.5 mM MgCl₂ was present in the PCR reaction, and 3) annealing was

performed at 52°C. The PCR products were analyzed by agarose gel electrophoresis and a Southern blot hybridization was performed with the internal oligonucleotide, BG24.

Direct amplification of genomic *Amb a V* DNA from pollen

Petroleum ether-defatted short ragweed pollen was incubated at 37°C for 30 min in PCR buffer (50 mM KCl/10 mM Tris-HCl, pH 8.4/2.0 mM MgCl₂/0.01% gelatin) containing 0.05 mg/ml of proteinase K, 20 mM dithiothreitol and 1.7 μM SDS. The mixture was heated at 85°C for 10 min to inactivate the proteinase K. Ten μl of the extract/30 pmol each of *Amb a V* specific primers BG33 and BG31 (Table I)/10 μl of PCR buffer/10 μl of dNTP (2 mM each) in a total volume of 100 μl were denatured at 95°C for 10 min. *Taq* polymerase (2.5 U) was added and the DNA was amplified in a thermal cycler as follows: denaturation 93°C for 1.5 min, annealing at 37°C for 45 s, and extension at 72°C for 45 s, with a final extension for 10 min.

Sequencing of PCR products

The PCR products were digested with *Eco*RI and ligated into an *Eco*RI-digested dephosphorylated M13 mp10 vector as described (19). For the isolation of 5'- and 3'-end specific clones, the M13 plaques were screened as described (17). Oligonucleotides BG22 and BG24 were used as probes separately for 5'-end clones to obtain both complementary strands. For the 3'-end clones, the identical procedure was followed except that only BG24 was used. The ssDNA was isolated from M13 clones and sequenced using Sequenase (U.S. Biochemicals, Cleveland, OH), according to the supplier's protocol.

Genomic Southern blot analysis

Short ragweed genomic DNA was digested separately with *Eco*RI, *Bam*HI, and *Rsa*I restriction enzymes. The digested DNA was separated on 0.7% agarose gel. The DNA was transferred onto a nylon membrane (GeneScreen Plus) following standard protocol (19). The blot was prehybridized for 3 h in 6X SSPE/5X Denhardt's solution/100 μg/ml denatured salmon sperm DNA/0.1% SDS and was hybridized for 18 h in fresh hybridization solution in the presence of a ³²P-labeled *Amb a V* cDNA probe. The membrane was washed with increasing stringency to a final wash in 0.5X SSC/0.1% SDS at 55°C for 30 min and exposed to x-ray film for autoradiography.

Expression of recombinant *Amb a V* protein in *Escherichia coli*

The coding sequence for the *Amb a V* protein was amplified using PCR with the following primers: 5' primer: 5'-GA-

AcatatgggatccTTAGTGCCGTGTGCGTGGGCGGGG-3', and 3' primer: 5'-GAACatatggaattcATTATTTGCCGCA-TTTTTTGGTGCATAT-3'. The amplified products were digested with *Eco*RI and *Bam*HI and subcloned into the pGEX-2T vector, inframe with a 26-kDa polypeptide of GST from *Schistosoma japonicum* (20). After transformation of *E. coli* strain JM109 with the recombinant plasmid, the fusion protein was expressed and purified as described (21). The purified fusion protein was then cleaved with thrombin and the resulting *rAmb a V* and GST fragments were separated by reverse phase HPLC (21).

DARIA

Antibody binding to native and recombinant *Amb a V* was measured by DARIA (9). Human sera were obtained from ragweed-allergic individuals who had been found to possess substantial levels of IgE antibody to *Amb a V*, and had not received ragweed immunotherapy for at least 5 yr. Briefly, 1 ng of ¹²⁵I-labeled *nAmb a V* or *rAmb a V* was incubated with diluted human serum (1/2 for IgE and 1/10 for IgG analyses). A total of 50 μl of each of goat anti-human IgE antiserum and human IgE myeloma PS (1/250) for the IgE analyses, or 100 μl of goat anti-human IgG antiserum was added for the IgG antibody analyses. The mixtures were incubated at 4°C for 18 h. The precipitates were pelleted, washed four times in BBS and counted to assess direct Ag binding.

Antigenic cross-reactivity between recombinant and native *Amb a V* was measured by competitive DARIA. One ng of ¹²⁵I-labeled *nAmb a V* protein was incubated for 4 h at 23°C with human allergic serum, diluted as above, in the presence of 10-fold dilutions (100, 10, 1, and 0.01 ng) of unlabelled *rAmb a V* or *nAmb a V*. The second antibody (anti-IgE or anti-IgG) was added as before and the precipitate was washed and counted. The percent inhibition (%I) was calculated using the formula: %I = (cpm of test serum without inhibitor) - (cpm of test serum with inhibitor)/(cpm of test serum without inhibitor) - (cpm of negative control serum) × 100.

Similar inhibition experiments were performed using radiolabeled *rAmb a V*

Histamine release assays

Basophil histamine release sensitivities to recombinant and native *Amb a V* were assayed by a previously described procedure (22). Briefly, the RBC were sedimented from whole blood by dextrose sedimentation. The leukocytes were separated from plasma by centrifugation. The cell pellet was washed with PAG (Pipes, albumin, and glucose) and resuspended in PAG containing Ca²⁺ and Mg²⁺ salts. Aliquots of the final cell suspension were incubated with recombinant or native *Amb a V* at 37°C in a water bath for 45 min. After a brief centrifugation, the supernatants were

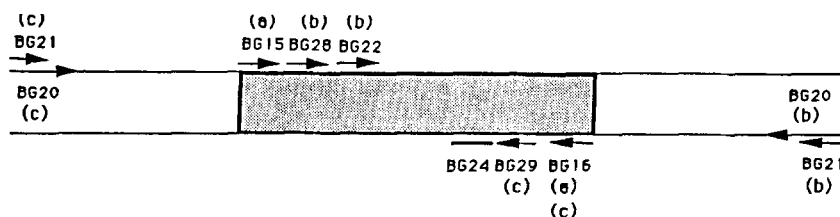


FIGURE 1. Schematic representation of the strategy for *Amb a V* cloning. The stippled area represents the portion of *Amb a V* where the amino acid sequence was known. This portion was amplified using primers BG15 and BG16 (a). For the 3'-end amplification (b), the ssDNA (produced using BG20) was amplified successively using BG28 plus BG21, followed by BG22 plus BG21 to increase the specificity. For the 5'-end amplification (c), the ssDNA (produced using BG16) was subjected to poly (A) tailing. The product was annealed with BG20 and then amplified using BG29 and BG21.

collected into plastic cups and used for measuring histamine release in an autoanalyzer (Technicon, South Plainfield, New Jersey) using *o*-phthaldialdehyde. The results are expressed as the percentage of histamine release induced by the allergen relative to that induced by lysis of the cells with 8% perchloric acid.

Results

Amb a V cDNA sequence

After the strategy of PCR amplification described in Figure 1, we obtained three sets of overlapping cDNA sequences (see *Materials and Methods*). From these overlapping sequences, the total nucleotide sequence of *Amb a V* was compiled (Fig. 2). The resultant full-length *Amb a V* cDNA sequence is compatible with the published amino acid sequence, except that the presumed translation-initiation codon (ATG, shown in bold) is one nucleotide out-of-frame in comparison with the published amino acid sequence (7). To test whether the cDNA sequence may have an artifact resulting from the anchored-PCR, we obtained independent clones from multiple PCR reactions. We have sequenced a total of six full-length independent cDNA clones in both orientations and all have the identical sequence.

Amb a V genomic DNA

To compare the cDNA sequence with the *Amb a V* genomic DNA sequence, we cloned and sequenced the *Amb a V* genomic DNA from short ragweed pollen. Primers BG33 and BG31 were used to amplify *Amb a V* DNA directly from pollen. The amplified DNA was subcloned into M13 and sequenced. The sequences of all the seventeen genomic DNA clones obtained exactly corresponded to the cDNA sequence obtained earlier, including the out-of-frame ATG codon. This finding also showed that the *Amb a V* gene contains no introns.

Comparison of the nucleotide sequences of *Amb a V* and *Amb t V*

Both the amino acid and nucleotide sequences of these two allergens are compared in Figure 2. From the amino acid

sequence deduced from the cDNA sequence, it is interesting to note that the *Amb a V* protein has an extra 10 amino acids at its C terminus. These residues are absent in the mature protein isolated from the pollen (6, 7, 9), possibly due to posttranslational cleavage or proteolytic digestion during purification. There is an inframe translation termination codon, TAA (shown in bold in Fig. 2), after an ATA codon that encodes the C-terminal amino acid, Ile. In addition to the expected homologies between these two allergens (49% homology at the amino acid level), portions of the sequence are highly conserved, specifically, a part of the signal sequence, codon minus 13 to codon minus 1 (*Amb a V* numbering; underlined in Fig. 2), and the underlined portions of the 3' ends. The underlined portions of the 5' ends have 97% homology (1 mismatch in 34 nucleotides) and the 3' ends have 89.5% homology (5 mismatches in 48 nucleotides).

Amb a V genomic Southern blot

To investigate the genomic organization of *Amb a V* gene, a Southern blot analysis of short ragweed genomic DNA digested with various restriction enzymes was performed and shown in Figure 3. The *Amb a V* cDNA binds to two bands in *Bam*HI (lane 1) and *Hind*III (lane 2) digested genomic DNA and to three bands in *Rsa*I (an internal site present in *Amb a V*) digested DNA (lane 3). This heterogeneity could be due to the presence of more than one gene in a single plant or to allelic variation of the *Amb a V* gene.

Expression of r*Amb a V* in *E. coli*

r*Amb a V* was expressed in *E. coli* fused to a 26-kDa fragment of GST. After affinity purification, the fusion protein was cleaved and the r*Amb a V* and GST were separated by reverse phase HPLC. Expression of soluble, monomeric r*Amb a V* in this system was relatively inefficient (10–100 µg/liter), most likely because the large number of cysteines and high degree of folding of *Amb a V* cause it to aggregate during synthesis. Solubilization of the aggregates in guanidine.HCl, followed by dilution and dialysis did not significantly increase the recovery of monomeric r*Amb a V* (21).

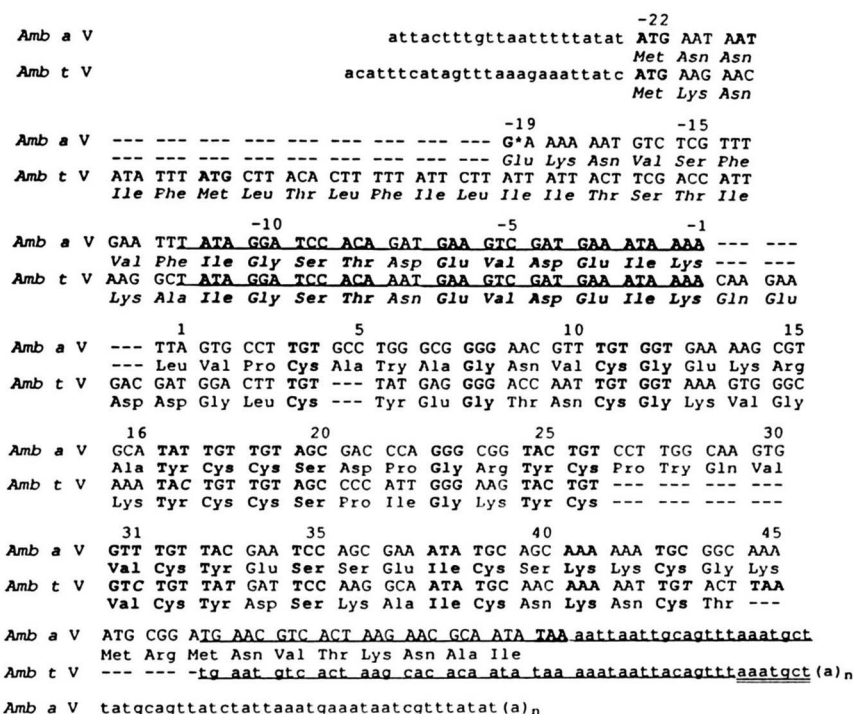


FIGURE 2. Nucleotide sequences of the cDNAs of *Amb a V* (this study) and *Amb t V* (17) with translated amino acid sequences. The identical codons and stop codons are shown in bold type. The postulated amino acid sequences of the signal peptides deduced from the nucleotide sequences are shown in italics. The highly homologous regions at the 5'- and 3'-ends are underlined (see text). The *Amb a V* translation initiation codon (ATG) at position minus 22 is putative. At the location shown by double-underlining, 50% of the clones sequenced were missing the "aaatgct" sequence and 25% of clones had a substitution of "caaatgct" for "aaatgct". The GenBank accession number for *Amb a V* is M84987; the number for the previously described *Amb t V* is M38782.

DARIA of sera from *Amb a V* allergic individuals showed a direct relationship between the amounts of IgE antibodies (18 individuals, $r = 0.888$, $p = 0.0001$) and IgG antibodies (25 individuals, $r = 0.952$, $p = 0.0001$) to native and recombinant *Amb a V*, as assessed by the amount of radioactive Ag bound (after subtracting controls) (Fig. 4). Similar correlations were obtained when "units/ml" of IgE and IgG antibodies (determined by reference to a standard) were plotted (data not shown).

Antigenic cross-reactivity between the two proteins was then assessed using a competitive DARIA where r*Amb a V* was used to inhibit the binding of human IgE and IgG to radiolabeled n*Amb a V*. In 9 of the 10 sera tested, r*Amb a V* was able to block a considerable fraction (50–80%) of the IgE and IgG antibody binding to the native protein; in one subject (DS) 100% inhibition was observed (see Fig. 5 for representative examples). The IgE and IgG inhibition data were similar for any given individual. We also performed inhibition experiments using ^{125}I -labeled r*Amb a V*. In this case, the inhibition curves for both unlabeled n*Amb a V* and r*Amb a V* were superimposable (data not shown). Taken together, the combined sets of inhibition DARIA data show that the r*Amb a V* contains a subset of the antigenic and allergenic epitopes of n*Amb a V*, except for one highly

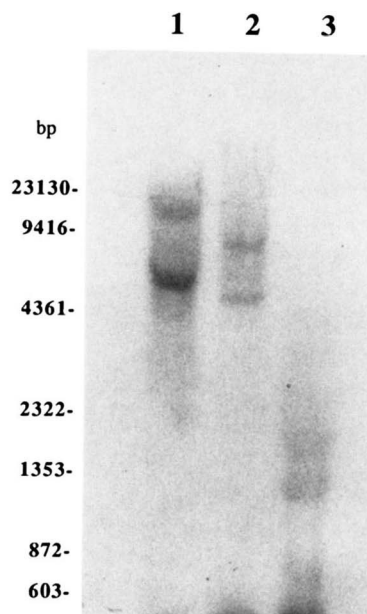
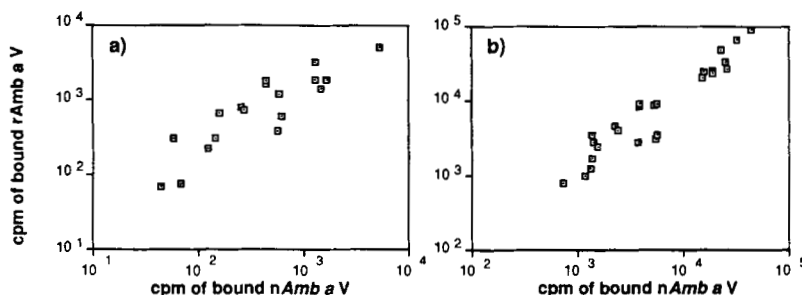


FIGURE 3. Southern blot analysis of short ragweed genomic DNA. The total genomic DNA was digested with *EcoRI* (lane 1), *BamHI* (lane 2), and *RsaI* (lane 3) and a Southern blot hybridization was performed with a ^{32}P -labeled *Amb a V* cDNA probe. Size markers are shown on the left.

FIGURE 4. Binding of *a*, human IgE and *b*, human IgG antibodies to recombinant and native *Amb a V*. The data are shown as total cpm of ^{125}I labeled protein bound by the antibodies (minus cpm for control sera).



allergic individual (Fig. 5, *c* and *d*) who was not able to distinguish between r*Amb a V* and n*Amb a V*. Finally, the ability of the recombinant and native proteins to release histamine from basophils from allergic and control individuals was tested. Basophils from two *Amb a V*-sensitive allergic individuals, AP and DS, who exhibited different inhibition curves for r*Amb a V*, released histamine as efficiently in response to r*Amb a V* as n*Amb a V* (Fig. 6). Basophils from two control individuals, TR (ragweed-allergic but not sensitive to *Amb a V*) and MEL (nonallergic), however, released no histamine in response to either r*Amb a V* or n*Amb a V* (data not shown). No histamine release was observed in any of the four subjects when the assay was performed using GST as the Ag.

Discussion

Using the anchored-PCR we have cloned and sequenced the cDNA encoding the *Amb a V* allergen from short ragweed pollen. Our PCR approach is suitable for genes that are expressed at very low levels and is much less time-consuming than conventional cDNA cloning. Initial attempts to clone this cDNA from a cDNA library constructed from short ragweed pollen was unsuccessful. The library was screened extensively by using human sera from *Amb a V*-allergic individuals and also by appropriate oligonucleotide probes including a synthetic *Amb a V* gene (kindly provided by Drs. Normand Brisson and Guy Boileau, University of Montreal, Montreal, Canada) deduced from the *Amb a V* amino acid sequence. Another attempt was made to screen the library using PCR-amplified DNA encoding a partial *Amb a V* segment as a probe, but we were still unable to pick up even a single positive clone. In contrast, Rafnar et al. (23) were able to clone cDNA encoding the *Amb a I* allergen from the same cDNA library. We, therefore, reasoned that the representation of *Amb a V* clones in the cDNA library was insignificant. This notion is further substantiated by the fact that the representation of mRNA encoding *Amb a V* was so low that a second PCR was required to find a detectable amount of amplified product in the agarose gel.

Using *Amb a V*-specific primers (BG33 and BG31), we have also amplified and sequenced *Amb a V* genomic DNA

directly from pollen. This is the first demonstration of a direct amplification and cloning of any allergen gene by PCR from genomic DNA. Genomic DNA sequencing showed that the *Amb a V* gene contains no introns. We sequenced 17 genomic clones and all of them were identical to the cDNA sequence. The absence of a known translation initiation codon inframe with the published amino acid sequence is interesting. One possible explanation is that *Taq* polymerase may skip a base during the initial elongation steps of the amplification cycles, as previously suggested (24). This possibility, however, seems unlikely because both genomic and cDNA sequencing were done from PCR products amplified in multiple independent PCR reactions. Alternatively, either an RNA-editing mechanism or a translation frame-shifting mechanism may be in operation to keep the sequence inframe for synthesis of the *Amb a V* protein. Such mechanisms have been observed in molecular genetic studies of yeast and other organisms (25–27). If RNA editing were functional, at least some of the cDNA should have an altered (inframe) sequence. However, sequencing of six full-length cDNA clones showed that all had the identical sequence that is out-of-frame by one nucleotide. This result argues against RNA editing as a possible mechanism. The usage of initiation codons other than ATG as suggested previously may be another possibility (28–30).

Comparing the nucleotide sequences of the *Amb a V* and *Amb t V* cDNA, we observed very high levels of homology at both the 5' and 3' ends (Fig. 2, underlined portions). We anticipate that these regions should also be conserved in *Amb V* homologues from other ragweed species, e.g., *Ambrosia psilostachya* (western), *Ambrosia bidentata* (southern), etc. It should then be possible to clone these *Amb V* homologues directly from pollen by PCR using conserved 5'- and 3'-end segments as primers. Initial experiments with western ragweed pollen using these primers have shown promising positive results (B. Ghosh, T. Rafnar, M. P. Perry, and D. G. Marsh, unpublished observations).

Our DNA sequencing shows the presence of codons for an "extra" 10 amino acids as compared to the published sequence at the C-terminus of the major 45 amino-acid *Amb a V* variant (7). A trypsin-like enzyme may remove the 10

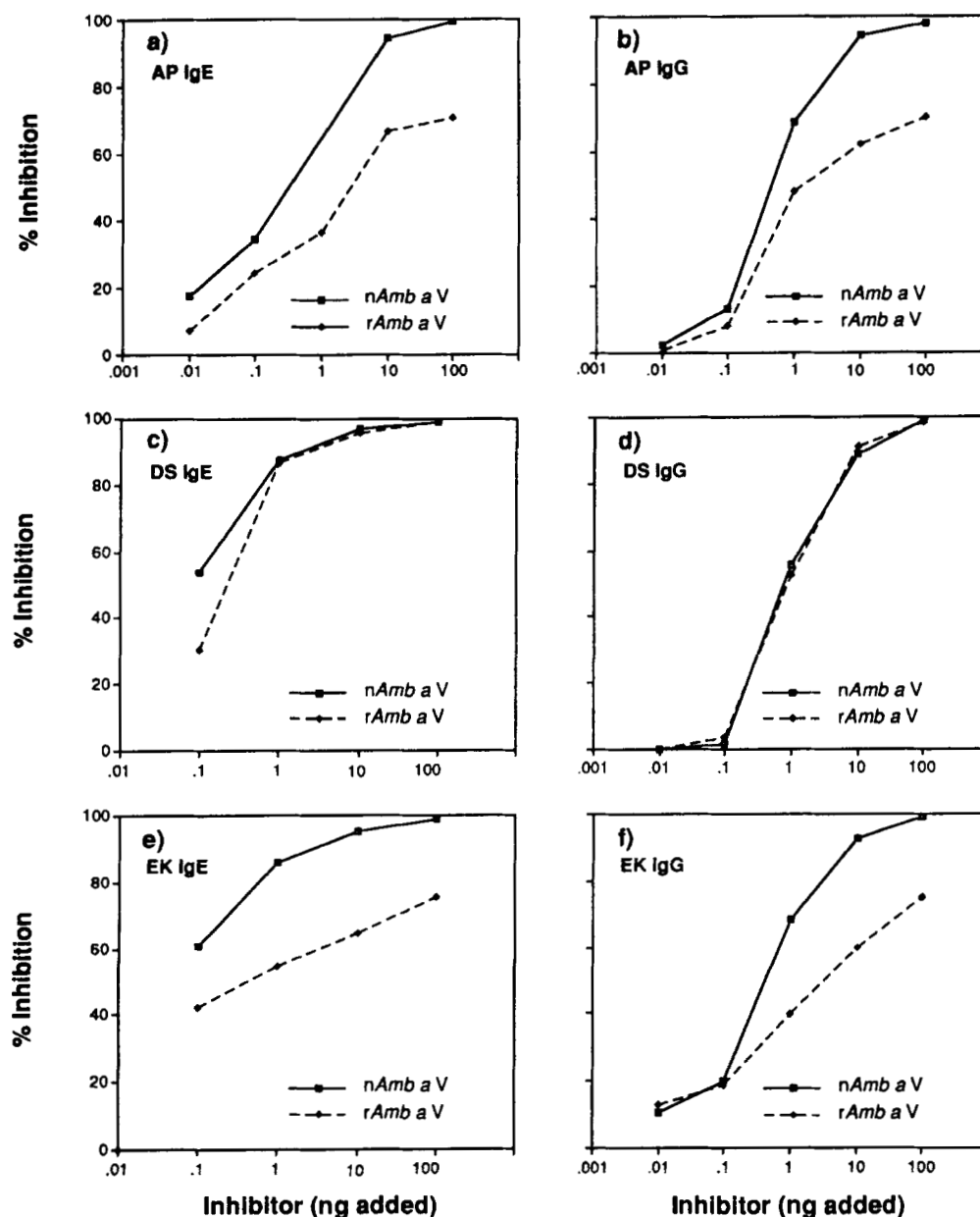


FIGURE 5. Competitive inhibition of ^{125}I -labeled *nAmb a V* binding to human IgE (a, c, and e) and IgG (b, d, and f) antibodies from three representative *Amb a V*-allergic individuals, using *nAmb a V* or *rAmb a V* as inhibitors.

amino-acid peptide during posttranslational processing or during purification of the *Amb a V* protein. It is also worth noting that in the *Amb t V* gene there is a TAA codon at this position (9, 17). In addition, a minor, 44 amino-acid *Amb a VB* variant has been isolated, which is identical to the *Amb a VA*, except that the C-terminal Lys is absent (9). We speculate that the gene encoding the *Amb a VB* form could have been generated by a mutation of the Lys codon (AAA) found in the *Amb a VA* gene to a termination codon (TAA), found in the *Amb t V* homologue.

Expression of *rAmb a V* in *E. coli* resulted in the recovery of a protein fraction that was antigenically similar, but not

identical to *nAmb a V*. This was not surprising because it is known that eukaryotic proteins, rich in Cys and Pro residues, fold inefficiently in the bacterial cytoplasm and can form insoluble aggregates (reviewed in Ref. 31). However, a fraction of the recombinant product is recovered in a soluble, monomeric form, as demonstrated by HPLC analysis after cleavage of the GST-*Amb a V* fusion protein (21). This excludes the possibility that *rAmb a V* aggregates may be interfering in the competition assays or causing nonspecific histamine release. Nevertheless, for large scale production of *rAmb a V* it may be more appropriate to express the protein in a system that secretes the products, increasing the

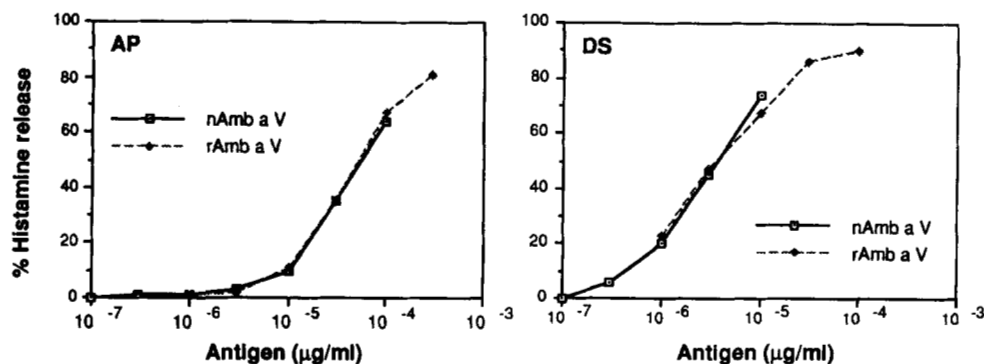


FIGURE 6. Histamine release from basophils of two *Amb a V*-allergic individuals by native and recombinant *Amb a V*.

chance of correct folding, and decreasing aggregation.

The observation that *rAmb a V* inhibits less than 100% of antibody binding to *nAmb a V* in 9 of 10 individuals tested implies that the *rAmb a V* is missing one or more immunodominant determinant(s). Despite this structural difference from *nAmb a V*, the *rAmb a V* provides a useful diagnostic tool, as demonstrated by the fact that all the *Amb a V*-sensitive subjects tested recognized epitopes common to the two molecules and that *rAmb a V* is able to release histamine from these subjects. It is worth noting that most recombinant allergens produced to date have been analyzed antigenically using Western blotting, a technique that efficiently denatures the proteins and therefore relies largely on the recognition of linear epitopes by the human antibodies. The DARIA techniques that we have used also detect the conformational epitopes. Finally, the *rAmb a V* allergen has been found to be fully active in *Amb a V*-specific T cell proliferation assays (21). The *rAmb a V* and in vitro-generated mutants of *rAmb a V* should, therefore, be useful in studies of allergen T and B cell epitopes in humans.

Acknowledgments

The authors thank Dr. David G. Klapper of the University of North Carolina for providing us with short ragweed genomic DNA and Eva Ehrlich-Kautsky and Mary Lesko for expert technical assistance.

References

- Young, P. 1980. In *Asthma and Allergies, an Optimistic Future* (National Institutes of Health Publication No. 80-388, Washington, DC), p. 14.
- Smith, J. M. 1988. Epidemiology and natural history of asthma, allergic rhinitis, and atopic dermatitis (eczema). In *Allergy: Principles and Practice*, 3rd ed. E. Middleton, ed. C. V. Mosby Co., St. Louis, p. 891.
- Lewis, W. H., P. Vinay, and V. E. Zenger. 1983. In *Airborne and Allergenic Pollen of North America*. Johns Hopkins University Press, Baltimore.
- King, T. P., P. S. Norman, and J. J. Connell. 1964. Isolation and characterization of allergens from ragweed pollen, II. *Biochemistry* 3:458.
- King, T. P., P. S. Norman, and L. M. Lichtenstein. 1967. Isolation and characterization of allergens from ragweed pollen, IV. *Biochemistry* 6:1992.
- Underdown, B. J., and L. Goodfriend. 1969. Isolation and characterization of an allergen from short ragweed pollen. *Biochemistry* 8:980.
- Mole, L. E., L. Goodfriend, C. B. Lapkoff, J. M. Kehoe, and J. D. Capra. 1975. The amino acid sequence of allergen Ra5. *Biochemistry* 14:1216.
- Roebber, M., R. Hussain, D. G. Klapper, and D. G. Marsh. 1983. Isolation and properties of a new short ragweed pollen allergen, Ra6. *J. Immunol.* 131:706.
- Roebber, M., D. G. Klapper, and D. G. Marsh. 1982. Two isoallergens of short ragweed component Ra5. *J. Immunol.* 129:120.
- Goodfriend, L., A. M. Choudhury, D. G. Klapper, K. M. Coulter, G. Dorval, J. DelCarpio, and C. K. Osterland. 1985. Ra5G, a homologue of Ra5 in giant ragweed pollen: isolation, HLA-DR-associated activity and amino acid sequence. *Mol. Immunol.* 22:899.
- Marsh, D. G., S. H. Hsu, M. Roebber, E. Ehrlich-Kautzky, L. R. Freidhoff, D. A. Meyers, M. K. Pollard, and W. B. Bias. 1982. HLA-Dw2: A genetic marker for human immune response to short ragweed allergen Ra5. I. Response resulting primarily from natural antigenic exposure. *J. Exp. Med.* 155:1439.
- Marsh, D. G., D. A. Meyers, L. R. Freidhoff, E. Ehrlich-Kautzky, M. Roebber, P. S. Norman, S. H. Hsu, and W. B. Bias. 1982. HLA-Dw2: A genetic marker for human immune response to short ragweed allergen Ra5. II. Response after ragweed immunotherapy. *J. Exp. Med.* 155:1452.
- Marsh, D. G., P. Zwollo, S.-K. Huang, B. Ghosh, and A. A. Ansari. 1990. Molecular studies of human response to allergens. *Cold Spring Harbor Symp. Quant. Biol.* 54:459.
- Zwollo, P., E. Ehrlich-Kautzky, A. A. Ansari, S. J. Scharf, H. A. Erlich, and D. G. Marsh. 1991. Molecular studies of human immune response genes for the short ragweed allergen, *Amb a V*. Sequencing of HLA-D second exons in responders and non-responders. *Immunogenetics* 33:141.
- Huang, S.-K., P. Zwollo, and D. G. Marsh. 1991. Class II MHC restriction of human T-cell responses to short ragweed allergen, *Amb a V*. *Eur. J. Immunol.* 21:1469.
- Huang, S.-K., and D. G. Marsh. 1991. Human T-cell responses to ragweed allergens: *Amb V* homologues. *Immunology* 73:363.

17. Ghosh, B., M. P. Perry, and D. G. Marsh. 1991. Cloning the cDNA encoding the *Amb t V* allergen from giant ragweed (*Ambrosia trifida*) pollen. *Gene* 101:231.
18. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239:487.
19. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
20. Smith, D. B., and K. S. Johnson. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* 67:31.
21. Rafnar, T., B. Ghosh, W. J. Metzler, S.-K. Huang, M. P. Perry, L. Mueller, and D. G. Marsh. 1992. Expression and analysis of recombinant *Amb a V* and *Amb t V* allergens: comparison with native proteins by immunological assays and NMR spectroscopy. *J. Biol. Chem.* 267:21119.
22. May, C. D., M. Lyman, R. Alberto, and J. Chang. 1970. Procedures for immunochemical assay of the histamine release from leukocytes with small volumes of blood. *J. Allergy* 46:12.
23. Rafnar, T., B. L. Rogers, J. E. Bond, M. Kuo, and D. G. Klapper. 1991. Cloning of *Amb a I* (antigen E), the major allergen family of short ragweed pollen. *J. Biol. Chem.* 266:1229.
24. Weber, J. L. 1990. *Genomic Analysis Vol. 1: Genetic and Physical Mapping*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, p. 159.
25. Brierley, I., M. E. G. Boursnell, M. M. Binns, B. Bilimoria, V. C. Blok, T. D. D. Brown, and S. C. Inglis. 1987. An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J.* 6:3779.
26. Fu, L., R. Ye, L. W. Browder, and R. N. Johnston. 1991. Translational potentiation of messenger RNA with secondary structure in *Xenopus*. *Science* 251:807.
27. Hoffman, M. 1991. RNA editing: what's in a mechanism? *Science* 253:136.
28. Kozak, M. 1990. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proc. Natl. Acad. Sci. USA* 87:8301.
29. Mehdi, H., E. Ono, and K. C. Gupta. 1990. Initiation of translation at CUG, GUG, and AUG codons in mammalian cells. *Gene* 91:173.
30. Beames, B., S. Braunagel, M. D. Summers, and R. F. Lanford. 1991. Polyhedrin initiator codon altered to AUU yields unexpected fusion protein from a Baculovirus vector. *Biotechniques* 11:378.
31. Wilkinson, D. L., and R. G. Harrison. 1991. Predicting the solubility of recombinant proteins in *Escherichia coli*. *Biotechnology* 9:443.