

Nucleotide Sequence Analysis of Three cDNAs Coding for *Poa p IX* Isoallergens of Kentucky Bluegrass Pollen*

(Received for publication, June 25, 1990)

Andre Silvanovich‡§, James Astwood‡, Lei Zhang¶, Egil Olsen¶, Fred Kisil§¶, Alec Schon§¶, Shyam Mohapatra‡¶||, and Robert Hill‡§

From the Departments of ‡Plant Science and ¶Immunology, §Medical Research Council Group for Allergy Research, University of Manitoba, Winnipeg, Manitoba R3E 0W3, Canada

Grass pollen allergens are one of the major causes of type I allergic reactions (allergic rhinoconjunctivitis, allergic bronchial asthma, and hayfever) in temperate climates afflicting 15–20% of a genetically predisposed population. Workers have found considerable physico- and immunochemical heterogeneity within the grass pollen allergens which has made them difficult to purify for both therapeutic uses and further biochemical study. We recently reported the construction of a cDNA library in λ gt11 using mRNA extracted from dehydrated Kentucky bluegrass (KBG, *Poa pratensis*). Here, we present the nucleotide and deduced amino acid sequences for three KBG pollen allergen cDNA clones, KBG 41, 60, and 31, which were isolated from the above library using a pool of six sera from grass pollen allergic patients. These clones exhibit an exceptionally high degree of sequence similarity to one another, only minor similarity to other known allergens, and no homologies to other known proteins or genes. The predicted molecular mass for the cloned proteins range from 28.3 to 37.8 kDa with pI values of 9.6–10.2. All three clones appear to possess leader peptides and lack asparagine sequons required for N-glycosylation. Therefore, the molecular mass of the post-translationally modified proteins were calculated to be 28.4–34.9 kDa, which is consistent with the size of the polypeptides revealed in Western blots of pollen proteins using an antiserum to a recombinant peptide encoded by the partial cDNA clone KBG 8.3. Northern blotting analysis indicates that expression of the genes corresponding to these clones is confined to pollen tissue. The results suggest that the clones code for a group of proteins that represent a new and previously uncharacterized group of grass pollen isoallergens, which have been hereby designated as *Poa p IX*.

Aero-allergenic proteins originate from a wide variety of sources including pollens of grasses, weeds, and trees, spores

of molds, feces of mites, and danders of pets and domestic animals (1). At very low doses (1 μ g/year), these proteins induce the synthesis of IgE antibodies which become attached to their corresponding surface receptors (*i.e.* Fc ϵ receptors) on the mast cells and basophils. Re-exposure of the individual to the same allergenic protein(s) leads to the cross-linking of the bound IgE antibodies, followed by the release of inflammatory mediators, *e.g.* histamine, arachidonic acid, and its metabolites, from these cells (2). The immunogenicity of a given allergen or antigen depends upon the following three factors: (i) the processing of the molecule to its immunogenic fragments by antigen-presenting cells, (ii) the exposure of these fragments in association with major histocompatibility molecules on the surface of these cells, and (iii) their presentation to the appropriate T-cells capable of recognizing the resulting complex (3). However, the difference between antigenic and allergenic determinants and the mechanism by which they tilt the immune system toward the preferential production of IgE antibodies in predisposed individuals is not clearly understood. Such studies have been limited due to the lack of primary sequence information of allergens and the unavailability of pure allergenic proteins or their peptides in sufficient quantities. Highly purified allergens are a prerequisite for accurate diagnosis and treatment of type I allergies.

A number of laboratories have purified and characterized allergenic proteins using standard biochemical techniques (4–12). Although such procedures were successful, they are labor intensive and yield limited amounts of pure allergens. Furthermore, the methods are inapplicable to the characterization of relatively high M_r allergens or those which constitute minor components of their corresponding sources (13). Recently, recombinant DNA technology has become the tool of choice for allergen characterization. It has resulted in rapid primary sequence determination for high molecular weight allergens (14–17), synthesis of the corresponding allergens or their peptides in virtually unlimited quantities (13, 18), and analysis of the structures (or epitopes) of the allergenic molecules responsible for binding T and B cells (13, 19).

We previously reported the construction of a λ gt11 cDNA library complementary to mRNA of KBG¹ pollen and the screening leading to the identification of several cDNA clones encoding human IgE-binding proteins. One clone, KBG 7.2, was analyzed to determine its nucleotide and deduced amino acid sequence (18) and the allergenicity of the corresponding recombinant peptide in a murine model (20). Here, we report the complete nucleotide and deduced amino acid sequences of three cDNA clones, KBG 41, 60, and 31, encoding a major group of KBG pollen allergens. These clones have a high

* This work was supported by the Medical Research Council of Canada and the Manitoba Health Research Council of Canada, Winnipeg, Manitoba, Canada. 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.

The nucleotide sequences reported in this paper for clones KBG 31, 41, and 60 have been submitted to the GenBank®/EMBL Data Bank and assigned accession numbers M38342, M38343, and M38344, respectively.

|| Recipient of a Manitoba Health Research Council Scholarship. To whom correspondence should be addressed: Dept. of Immunology, The University of Manitoba, 730 William Ave., Winnipeg, Manitoba R3W 0W3, Canada.

¹ The abbreviations used are: KBG, Kentucky bluegrass; ORF, open reading frame; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; bp, base pair.

degree of similarity to one another and only minor similarity to other known allergens. On the basis of the above results and further physicochemical and immunochemical studies,² the isoallergenic proteins which are encoded by the members of a multigene family and are expressed specifically in pollen tissue have been designated as *Poa p* IX group.

MATERIALS AND METHODS

Screening of the cDNA Library in Bacteriophage λ gt11—Plaques (250,000) from a λ gt11 library (18) were screened immunologically for the expression of pollen allergens using pooled sera of six patients allergic to KBG pollen. Patients were selected on the basis of case history, positive skin prick, and radioallergen sorbent tests but had not undergone any hyposensitization treatments. Plaques were lifted onto nitrocellulose filters (Bio-Rad, Toronto) and washed twice for 5- and 30-min intervals in 30 ml of buffer (50 mM sodium phosphate, pH 7.5, 0.5% (v/v) Tween 20, 0.5% bovine serum albumin, 0.05% sodium azide (w/v)/filter. Filters were then incubated overnight at 4 °C under slight agitation in buffer containing 10% pooled sera. The filters were then washed at 4 °C, for three 15-min intervals in 25 ml of buffer/filter and incubated overnight at 15 °C in goat-anti-human IgE conjugated with alkaline phosphatase. The filters were rewashed as previously described and developed using 5-bromo-4-chloro-3-indolylphosphate *p*-toluidine and nitroblue tetrazolium chloride. Positive clones were picked and rescreened, a total of 11 clones were isolated and purified.

KBG 41 was isolated from the cDNA library using a 371-bp nucleotide probe digested from the 5' end of KBG 4.1 with *Eco*RI and *Pst*I. Plaques (100,000) were lifted onto Hybond-N filters (Amersham Corp.), denatured in 0.5 M NaOH, 1.5 M NaCl for 30 s, neutralized in 0.5 M Tris, pH 8.0, 1.5 M NaCl for 5 min, and rinsed with 6 × SSC (where 20 × SSC is 3.0 M NaCl, 0.3 M sodium citrate, pH 7.0). Following air drying, filters were exposed to 365 nm UV light for 10 min and prehybridized in a solution consisting of 10 × Denhardt's (where 50 × Denhardt's is 1.0% (w/v) Ficoll, 1.0% polyvinyl pyrrolidone, and 1.0% bovine serum albumin (Pentex Fraction 5)) solution, containing 6 × SSC and 0.1% SDS at 65 °C for 1 h. 200 pg of probe (10⁸ cpm/μg of DNA) was nick-translated according to manufacturer's recommendations (Amersham, Toronto), added to fresh hybridizing solution. Hybridization was allowed to proceed overnight at 65 °C. Filters were washed at 65 °C for 15-min intervals, twice in 6 × SSC, 0.1% SDS and twice in 2 × SSC, 0.1% SDS. Autoradiograms were exposed for 3 days and 20 of 250 positive clones were plaque purified.

Phage DNA Preparation and Subcloning—DNA was prepared from plaque-purified phage using the liquid lysate method (21). Inserts recovered from *Eco*RI digestion were ligated into the multiple cloning site of pT7T3 19U (Pharmacia LKB Biotechnology Inc.) and used to transform *Escherichia coli* DH5α (Bethesda Research Laboratories). Further subcloning necessary for sequencing was done in the multiple cloning site of pT7T3 18U and 19U.

Sequencing of cDNA Inserts—All sequencing of cDNA inserts was done using single strand plasmid templates. These templates were prepared following the manufacturer's recommendations except that *E. coli* DH5αF' were used in place of *E. coli* NM522. Sequencing was done according to the dideoxy method (22) using M13 universal and T3 promoter primers. Due to the high GC content of the insert DNA, all sequencing reactions were done in duplicate using dGTP and dTTP labeling and termination mixes.

Structure Predictions—Potential antigenic sites were predicted through the use of hydropathy and antigenic potential plots. Hydropathy plots were calculated as described by Kyte and Doolittle (23) using a 9-residue amino acid window. Antigenic potential was calculated essentially as described by Fraga (24) with the following modifications. Values for a global profile were calculated using the following equation: AP = Hyd/(Rec - 78). Where AP is the antigenic potential and Hyd and Rec are the average hydropathy and recognition values for a 6-amino acid residue window (note: the hydropathy values used by Fraga differ from those used by Kyte and Doolittle). Regions of positive antigenic potential of 6 or more amino acid residues in length whose magnitude has not decreased relative to the same region on the Fraga hydropathy plots (not shown) are considered

potential antigenic sites. Potential membrane-bound regions were predicted as described (23, 25).

RNA Gel Blot Analysis—Total pollen and leaf RNA were electrophoresed in 1.5% agarose gels under denaturing conditions (26) and blotted onto nylon membrane. Fragments containing sequences unique to each of the full-length cDNA clones were prepared by cleavage with appropriate restriction endonucleases and nick-translated. These fragments were as follows: KBG 41, *Eco*RI-*Rsa*I fragment, bases -142 to -18; KBG 60, *Sau*3A fragment, bases 923-1068; KBG 31, *Eco*RI-*Nar*I fragment, bases 1-261.

Isolation of the KBG 8.3-β-Galactosidase Fusion Protein—*E. coli* strain JM105 was transformed with the expression vector pWR 590.1 (27) containing the insert of KBG 8.3 (8.3 is a partial clone of KBG 60 which spans the region 232-1132 bases) and was grown under appropriate conditions to stimulate fusion protein production. Total crude proteins were recovered from the host cells and the fusion protein was purified by electrophoresis on 10% SDS-PAGE followed by electroelution from the polyacrylamide gel.

Preparation of Antisera to Fusion Protein—The fusion protein was used as an immunogen in a dextran sulfate adjuvant to immunize 6-8-week-old female BDF1 mice (Jackson Laboratories, Bar Harbour, ME). The mice were immunized subcutaneously in the mucosal region, and after 23 days the mice were administered a booster of the same antigen preparation. Blood was collected 7 days later and serum was prepared.

Western Analysis—KBG proteins were isolated as previously described (18) and electrophoresed on a 15% polyacrylamide gel. Proteins were then stained with Coomassie Blue or electroblotted onto a nitrocellulose membrane. Membranes were blocked with 1% gelatin in Tris-buffered saline and probed with specific antisera. Immunoreactive bands were detected using an alkaline phosphatase-conjugated second antibody.

RESULTS

Isolation of cDNA Clones—A λ gt11 cDNA library derived from KBG mRNA which was screened with human allergic serum, yielded 11 strongly reacting clones. Of these, nine were successfully restricted with *Eco*RI, ligated into pT7T3 19U, and subcloned into *E. coli* strain DH5αF'. From partial sequencing and restriction enzyme mapping of the subcloned inserts, we deduced that five different cDNAs had been cloned. All clones except KBG 41, summarized in Table I, were identified by the above immunoscreening procedure. A full-length cDNA containing within it KBG 4.1 was identified from the KBG library using a 371-bp 5' end *Pst*I-*Eco*RI fragment of 4.1 as the probe. Twenty of 250 positive clones were plaque-purified. DNA isolated from these clones was restricted with *Eco*RI and subjected to Southern analysis using the same probe. Two clones containing 1352-bp inserts were subcloned and restriction mapped. They were found to be identical and were termed KBG 41. From Southern analy-

TABLE I
Relationships among the KBG pollen cDNA clones
selected from λ gt11 library

Class	Clones	Size of inserts ^a kbp
I	KBG 41	1.352
	KBG 4.1	0.936
II	KBG 60	1.222
	KBG 1.1	0.900
	KBG 8.3	0.800
	KBG 10	0.500
III	KBG 31	1.322
	KBG 7.2	0.560
IV	KBG 6.1	0.500
V	KBG 5.1	0.900

^a Sizes of partial cDNAs are approximations from gel electrophoresis.

² E. Olsen, R. D. Hill, F. T. Kisil, A. H. Sehon, and S. S. Mohapatra, manuscript submitted.

The KBG clones have an exceptionally high degree of similarity in the coding region with all clones having 95% identity within a consensus region of 771 bp. To the 3' end of the consensus region, KBG 60 and 31 continue to share similarity with an additional 42 bp and extending 10 bp past the stop codon. KBG 41 contains a characteristic 32-bp in-phase internal repeat (30 of 32 bases match) at the 3' end. The region begins at 929 and continues through to 960, and it is then repeated from 965 to 996. To the 5' end of the consensus region and extending back through the start codon, KBG 41 and 60 share a high degree of similarity to one another. However, they share no similarity to the corresponding region of KBG 31. Both the 3'- and 5'-untranslated regions while retaining many characteristic features, do differ substantially. Although the clones share a high degree of similarity they are readily distinguished from one another through restriction map analysis. KBG 60 has a unique *Sst*I

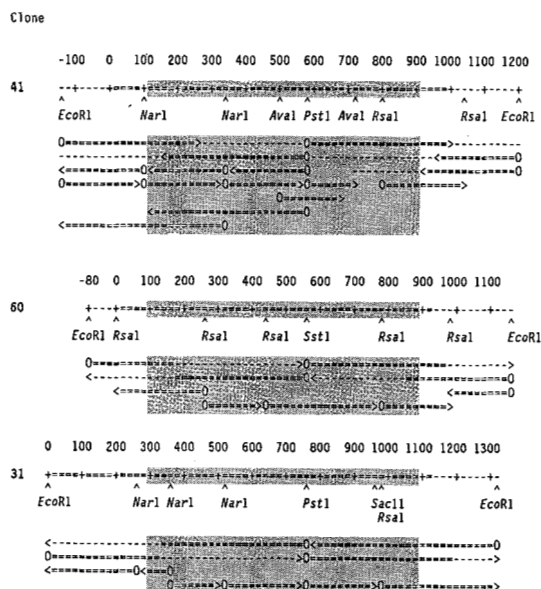
[illegible]

FIG. 2. cDNA sequences of KBG clones 41, 60, and 31. Sequences are broken into codons, have gaps inserted, and are aligned within the translated region to show maximum similarity. *Dots* represent homology with clone KBG 41. *Underlined regions* are polyadenylation signals. *Double underlined regions* on clone 41 indicate internal repeat. *Lowercase characters* are *EcoRI* linkers. *Numerical values* at the start of each line refer to clone number and the number of nucleotide residues from the A in the start codon.

site at 541 bp and KBG 31 contains a unique *Sst*II site at 953 bp (Fig. 1).

Deduced Amino Acid Sequences—The open reading frames of KBG 31, 41, and 60 consist of 373, 333, and 303 codons and yield deduced molecular masses of 37.8, 32.7, and 30.5 kDa, respectively (Fig. 3). The three derived peptides all appear to possess leader peptides on the basis of hydropathy profiles of the NH₂-terminal sequences (Fig. 4, *a* and *b*). The putative leaders for KBG 41 and 60 have all of the following features typical of a eukaryotic signal peptide (30, 31): a

```

41  -23 mavhqtvalflavalvagaas
60  -22 ...qk...V...V...
31  -28 .dkangaykta.ka.sava..ekfpvfqATFDKMKKEGLSGPDVGFAGKLDATQTSYL

41  1      YAADVEYGAPATLATPATPAAGYTPAAGCAAPKA
60  1      ...LS...A...
31  33 STKAAPKKEKDFVLSLSTEVLRFMAGYK...PASKF...K...PKV.A...

41  39 TDEQKLEIKINAGFKAFAAAGVPAVDKTKFVATFGTASNKFAEALSTEPKGAAGA
60  33 ...V...G...AN...A...VD
31  93 ...V...AS...A...V...

41  99 SSNAVLSKLDAAAYKLAYKSAEGATPEAKYDAYVATLSEALRIITAGTLEHAYKPAGEEV
60  93 ...A...D...G...A...
31  153 ...G...A...

41  159 KATPAGELQVIDKVDAAFKVAATAAANAAPANDKFTVFEEAFDAIKASTGGAYQYKFI
60  153 ...T...
31  213 ...

41  219 ALEAVKQSYAATVATPAVKYTVFETALKKATTAAMSQAQKAPAAVATATATSAVGAA
60  213 ...A...G...
31  273 ...G...

41  279 TGAVGAATGATTAAGGKYKTGAATPTAGGKYV
60  254 ...AA...
31  333 ...AA...

```

FIG. 3. Translated sequences of KBG clones 41, 60, and 31. Sequences are aligned and have gaps indicated to show maximum similarity between the clones. The alignments are the same as those in Fig. 2. Putative signal peptides are in lowercase characters. Underlined regions in KBG 31 are predicted B-cell epitopes (24). Double underlined region on clone 41 shows internal repeat. Numerical values at the start of each line refer to the clone and the number of amino acid residues from putative cleavage site of the signal peptide.

hydrophilic sequence of 5 amino acid residues at the COOH terminus, an uninterrupted hydrophobic sequence of 9 or 10 aliphatic and aromatic amino acid residues, and a hydrophilic NH₂ terminus of 7 amino acid residues. The amino acid residues include a serine at -1, the cleavage site, aliphatic alanine residues at -2 and -3, and helix breakers, proline and glycine at -4 and -5. For KBG 31, the evidence is not as clear cut due to the length of the signal peptide and the residue found at the -6 position. The signal peptide is 28 residues in length consisting of a 9-residue hydrophilic COOH terminus, an interrupted 10-residue central hydrophobic region which contains a lysine and a serine residue, and a 9-residue hydrophilic NH₂ terminus. Residues adjacent to the cleavage site include an uncharged glutamine at the cleavage site, phenylalanine at the -2 position, valine at -3, and the helix breaker proline at -4. The -5 position is phenylalanine, while the -6 position which is normally occupied by a small, neutral amino acid contains lysine. Taking into account these leader sequences, the calculated molecular masses for post-translationally modified KBG 31, 41, and 60 proteins become 34.9, 30.4, and 28.3 kDa. No asparagine sequons required for N-glycosylation (32) are found in any of the three clones. The predicted pI values for mature KBG 31, 60, and 41 are 10.08, 9.96, and 9.70.

Structure Predictions—Hydropathy (23) and antigenic probability (24) profiles calculated for KBG 41 and 31 predicted eight and 11 potential antigenic sites (Figs. 3 and 4, *a-d*) using a 9-residue window for hydropathy and a 6-residue window for antigenic potential. Hydropathy profiles were also calculated using 19 and 21 residue windows employing the hydrophilicity values of Kyte and Doolittle (23) and Eisenberg

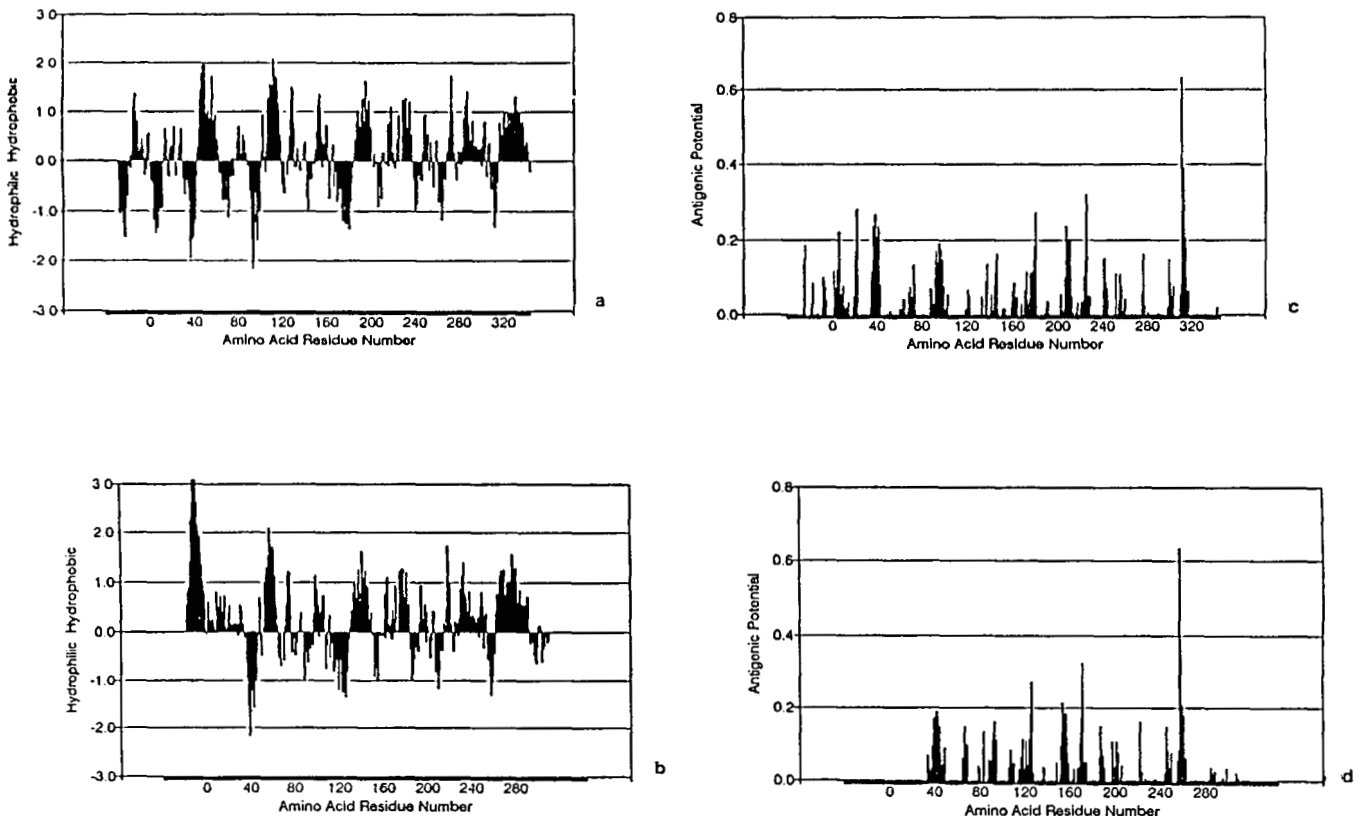


FIG. 4. Hydropathy (*a* and *b*) and antigenic potential (*c* and *d*) profiles of KBG 31 and 41. Hydropathy profiles were calculated using a 9-amino acid residue window. Antigenic potential (24) was calculated as described using a 6-residue window. Profiles are staggered to show similarity between clones. Potential antigenic site sequences are indicated in Figs. 2 and 5.

(25) in an effort to determine if any regions within KBG 31 and 41 are membrane-bound α -helical structures. While no region could be found to satisfy the criteria of Kyte (19 residue window, average hydrophobicity >1.6) or Eisenberg (21 residue window, average hydrophobicity of >0.65 for 1 window, or 1.10 total for 2 non-overlapping windows), the carboxyl-terminal of each protein had values that were quite high (19 residues, Kyte values, 1.31). Inspection of the sequence in this region showed them to be unique in several aspects that would be consistent with a membrane-associated segment. KBG 41 has a 32-residue region beginning at 264 which has 16 alanine residues and no charged residues. Likewise, KBG 31 has a 26-residue region beginning at 319 which has 13 alanine residues and no charged residues.

Sequence Comparisons—DNA sequences were also compared with those found in the Genbank data base release 61 (September 1989) using the IBI-Pustell sequence analysis package (Version 2.02). Only short and sporadic similarities were found, none of which could be considered significant. A similar search for homologous proteins in protein data banks revealed no meaningful large scale homologies.

KBG clone sequences were compared with all available data for known allergens. These include the full-length protein sequences obtained by Edman degradations for Ragweed pollen allergens *Amb a* III (12) and *Amb a* V (9) and Ryegrass pollen allergens *Lol p* II and III (4, 5), the deduced protein and cDNA sequences from the house dust mite allergen *Der p* I (15) and Birch pollen allergen *Bet v* I (14), and partial NH_2 - and/or COOH-terminal amino acid sequences for *Lol p* I (9, 33, 34), *Der f* I (35), and the cat dander allergen, *Fel d* I (6). The comparison of the consensus region and *Lol p* II and III and *Amb a* V revealed two short, but potentially important similarities within 20 residues of the methionine (Fig. 5). This region was found to be important for the immunogenic properties of the KBG 7.2 fusion protein (19). Typically the lysine residues were found to be conserved among the compared allergens. It is inferred that the cloned protein may share epitopes with other allergens.

Expression of Genes Encoding Allergens—A Southern blot analysis of genomic KBG DNA employing a 600-bp consensus sequence revealed hybridization to several fragments (data not shown) suggesting that the KBG allergenic proteins are encoded by a multigene family. Northern blot hybridization using nick-translated cDNA probes revealed that the corresponding transcripts are present in pollen but not leaf tissue (Fig. 6). Furthermore, unique fragments from each of the clones were used as probes to examine their corresponding transcripts. Probes from KBG 60 and 31 hybridized to mRNAs having a similar estimated size of 1320 nucleotides and KBG 41 hybridized to an mRNA having an estimated size of 1420 nucleotides. These results indicate that separate mRNAs exist for the cloned cDNA molecules.

For identification of the polypeptides encoded by these clones, a partial cDNA clone, KBG 8.3, which comprised sequences common to all three clones described in this study, was cloned into an expression plasmid which permitted high level expression of the corresponding recombinant peptide

<i>Lol p</i> II	37	ehgsnewl	ALKKng	dgveiKsdKPlkg	71
<i>Lol p</i> III	37	qhgsnewl	EdmTKKgn	lwevKsAKPltg	69
<i>Amb a</i> V	31	vVcYessEicsKKcgk			45
KBG 31	292	aVkytvtEtALKKaitamsqagKaAKPaaa			328

FIG. 5. Sequence similarities between KBG 31 and other known pollen allergens. Sequences of *Lol p* II and III (4 and 5) and *Amb a* V (12) are aligned with KBG 31 to show maximum similarity. Uppercase characters show amino acid residue matches. Underlined sequences are predicted B-cell epitopes. Numerical values at the beginning and end of each sequence refer to the residue number.

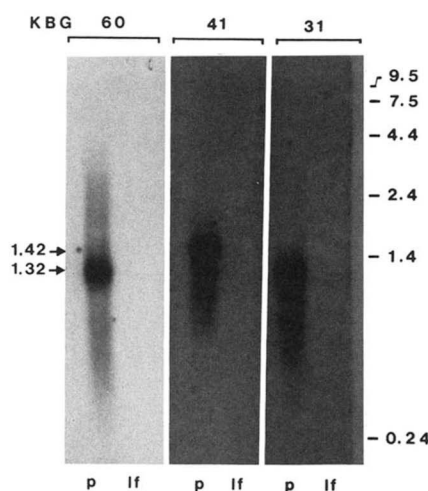


FIG. 6. RNA gel blot analysis of KBG clones 41, 60, and 31. Clones were hybridized to total RNA of KBG pollen, lanes p, and to KBG leaf tissue, lanes lf. RNA molecular size markers are indicated on the right-hand side. Transcripts are marked with arrows.

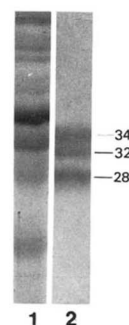


FIG. 7. Western blot analysis of KBG pollen proteins. The KBG pollen proteins were electrophoresed on SDS-PAGE and stained (lane 1) or electroblotted onto nitrocellulose and reacted with mouse antibodies to KBG 8.3 fusion protein (lane 2). The numbers on the right indicate the molecular mass estimated in kDa.

(rKBG 8.3). A Western blot employing a murine antiserum to rKBG 8.3 is shown in Fig. 7. It is clear that in the KBG pollen extract the antibodies recognized three polypeptide bands with estimated molecular masses of 28, 30, and 34 kDa. However, the 32- and 34-kDa polypeptides were barely detectable upon Coomassie staining of the gel indicating that these polypeptides are in minute amounts in the extracts. These results suggest that the three cloned cDNAs represent members of a multigene family, the expression of which is confined to pollen tissue.

DISCUSSION

In this report we present the molecular characterization of three cDNA clones, KBG 41, 60, and 31, which encode proteins that represent a new family of previously undescribed pollen isoallergens found in KBG. Considerable evidence exists to indicate the ORFs and predicted protein sequences presented here are correct. Comparison of the 6 bases preceding and the 2 following the start codons in KBG 41 and 60, 5 of 8 were found to be in agreement with consensus sequences surrounding start codons in plants (28). These start codons are in phase with the ORF of β -D-galactosidase in λ gt11. The poly(A) tails of all three clones are found from 22 to 28 bp from the canonical AATTA polyadenylation signal and 180–200 bp from the stop codons found on the predicted ORFs. The motifs TTATTTAT or ATTTA which may be associated with mRNA stability were found in the 3'-untrans-

lated region of all three cDNAs.

Based upon the deduced amino acid sequences and their hydropathy profiles we predict the presence of signal peptides of 21–28 amino acid residues in length. Sequences and the specific residues found adjacent to the predicted cleavage sites are in agreement with those seen in other signal peptides. Hydropathy profiles calculated employing both larger window sizes and alternate hydrophilicity values for amino acid residues indicated that the carboxyl-terminal regions of KBG 41 and 31 might be membrane-associated. Although evidence supporting this does not fully satisfy the criteria proposed by Kyte and Doolittle (23) and Eisenberg (25), it should be noted that their algorithms are best suited for characterizing transmembrane, α -helical structures. Exceptions to the criteria are known to exist and presumably these are due to both random coils and β -sheet structures that are able to pass through membranes (25). This highly hydrophobic region is immediately adjacent to a hydrophilic group of amino acids that represent the region with the greatest potential for allergenicity. The juxtaposition of these two regions may prove important in relation to the allergic response.

The polypeptides contain no asparagine sequons necessary for *N*-glycosylation. The calculated molecular masses for mature KBG 41, 60, and 31 proteins are 30.4, 28.3, and 34.9 kDa with pI values of 9.96, 9.70, and 10.08, respectively. These M_r values are in good agreement with those seen in Western blot analysis of KBG pollen extracts using polyclonal antibodies raised against the recombinant fusion protein of KBG 8.3. Western analysis of two-dimensional polyacrylamide gels shows a range of isoelectric points of 9–10 within this M_r range. In addition, this fusion protein was capable of binding to the IgE in sera of 90% of atopic individuals examined (20). Furthermore, by immunization of mice with the fusion protein, an IgE immune response was inducible (20). These observations indicate the three clones represent major isoallergens of KBG.

We found the deduced amino acid sequences of these KBG isoallergen proteins to be similar to one another and share a 255-residue consensus region. Within the consensus region, only slight variations were seen in the sequences, and perhaps these would do little to alter the physicochemical properties of the isoallergens. Rather, physicochemical properties are altered by the inclusion or exclusion of certain regions. Thus, KBG 31 has a unique 108-residue NH_2 -terminal region while KBG 41 has a COOH-terminal internal repeat. At present, these findings are only able to adequately explain the observed range of M_r values of the KBG isoallergens. Further sequencing and characterization of other KBG isoallergens are needed to determine (i) whether the range of observed isoelectric points is also due to such substitutions or due to altered ratios of charged amino acid residues, and (ii) what effect, if any, this will have on the allergenic properties of the proteins.

From the point of view of plant biology, the cloned sequences appear to belong to related genes of a small multigene family which are specifically expressed in pollen tissue. This, in addition to the finding that similar transcripts are present in the pollen of many grasses (18), indicating similarities in the coding region of these genes, suggests they may be important in pollen development and/or germination. These studies may help elucidate the evolution of IgE-binding domains in these and other related plant proteins and allow one to determine their biological function in pollen. Although KBG 60 and 31 identify mRNAs of similar size, the use of unique probes shows them to be different transcripts. Similarly, differential amounts of the corresponding proteins have been observed on SDS-PAGE gels. The level at which the regula-

tion of expression of these proteins occurs is not known.

The availability of primary amino acid sequence information will facilitate the identification of antigenic sites on these clinically important allergens. Computer-generated structural predictions indicate as many as 11 potential antigenic sites. Results published elsewhere for the partial clone KBG 7.2 which comprises three of 11 predicted sites (18) have indicated that at least two of these may be functional. However, it should be noted that there are reports of discrepancies from the predictions (36). Therefore, appropriate cellular and functional assays have to be performed to verify the above predictions. Clearly, the predicted sequence information would facilitate synthesis of short chemically synthesized peptides for the above studies.

Interestingly, while no large scale similarities are found to exist between our three isoallergens and other known allergens, one small but potentially important similarity was found between a COOH-terminal IgE-binding region adjacent to the methionine residue of the KBG clones (18) and a predicted B-cell epitope of *Lol p II* and III. Whether or not these small stretches of similar residues are responsible for the relatedness among allergens of a single pollen is not clear. On the basis of the extensive molecular analysis of these cDNA clones which provided information on the corresponding allergenic proteins and the studies on relationships between these and other major grass pollen allergens, these KBG allergens ought to be classified as a new group. In accord with the existing framework of IUIS International Allergen Nomenclature Subcommittee Recommendations (37), these basic isoallergens have been tentatively designated as *Poa p IX*. To our knowledge, these proteins constitute the third major group of basic allergens of grass pollens, the other two being Group III (5) and Group IV (2, 7) allergens.

Prior to the successful diagnosis and treatment of all individuals suffering from type I allergic diseases, the most basic questions regarding these diseases must be answered, for example: (i) what function if any do these proteins have in the plant, animal, or insect that is producing them? It has been found that the dust mite allergen, *Der p I*, has considerable homology with cysteine proteases (15). Likewise, the birch pollen allergen, *Bet v I*, has homology with the pea disease resistance response protein (14). Could the biological function of these proteins somehow enhance their antigenic activity? If so, why are only certain individuals afflicted? (ii) What characterizes the epitopes that are recognized by the T- and B-cells which lead to the enhanced IgE response seen in atopic individuals? It is anticipated that the identification of the epitopes will facilitate the development of novel diagnostic and therapeutic modalities for allergies and the determination of molecular mechanisms underlying the induction and regulation of allergic response in humans.

Acknowledgments—We would like to thank Andy Halayko, Andre Hamel, and Brenda Moore for excellent technical assistance. We would also like to thank Don Dubic and Dr. Mike Mowatt for assistance with the nucleotide and protein database searches.

REFERENCES

1. Matthiesen, F., Klysner, S., and Lowenstein, H. (1990) in *Epitopes of Atopic Allergens* (Sehon, A. H., Kraft, D., and Kunkel, G., eds) pp. 9–13, U. C. B. Institute of Allergy, Brussels
2. Kaplan, A. P. (ed) (1985) *Allergy*, Churchill Livingstone, New York
3. Unanue, E. R. (1985) *Nature* **317**, 359–361
4. Ansari, A. A., Shenbagamurthi, P., and Marsh, D. G. (1989) *Biochemistry* **28**, 8665–8670
5. Ansari, A. A., Shenbagamurthi, P., and Marsh, D. G. (1989) *J. Biol. Chem.* **264**, 11181–11185
6. Chapman, M. D., Aalberse, R. C., Brown, M. J., and Platts-Mills,

- T. A. E. (1988) *J. Immunol.* **140**, 812-818
7. Ekramoddouh, A. K. M., Kisil, F. T., and Sehon, A. H. (1983) *Mol. Immunol.* **20**, 465-473
8. Ipsen, H., and Lowenstein, H. (1983) *J. Allergy Clin. Immunol.* **72**, 150-159
9. Klapper, D. G., Goodfriend, L., and Capra, J. D. (1980) *Biochemistry* **19**, 5729-5734
10. Leitermann, K., and Ohman, J. L., Jr. (1984) *J. Allergy Clin. Immunol.* **74**, 147-153
11. Lin, Z., Ekramoddouh, A. K. M., Kisil, F. T., Herbert, J., and Mourad, W. (1988) *Int. Arch. Allergy Appl. Immunol.* **87**, 294-300
12. Mole, L. E., Goodfriend, L., Lapkoff, C. B., Kehoe, J. M., and Capra, J. D. (1975) *Biochemistry* **14**, 1216-1220
13. Thomas, W. R., Chua, K. Y., Greene, W. K., and Stewart, G. A. (1990) *Epitopes of Atopic Allergens* (Sehon, A. H., Kraft, D., and Kunkel, G., eds) pp. 77-81, U. C. B. Institute of Allergy, Brussels
14. Breitender, H., Pettenburger, K., Bito, A., Valenta, R., Kraft, D., Rumpold, H., Scheiner, O., and Breitenbach, M. (1989) *EMBO J.* **8**, 1935-1938
15. Chua, K. Y., Stewart, G. A., Thomas, W. R., Simpson, R. J., Dilworth, R. J., Plozza, T. M., and Turner, K. J. (1988) *J. Exp. Med.* **167**, 175-182
16. Fang, K. S. Y., Vitale, M., Fehlner, P., and King, T. P. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 895-899
17. Tovey, E. R., Jonhson, M. C., Roche, A. L., Cobon, G. S., and Baldo, B. A. (1989) *J. Exp. Med.* **170**, 1457-1462
18. Mohapatra, S., Hill, R., Astwood, J., Ekramoddouh, A. K. M., Olsen, E., Silvanovich, A., Hatton, T., Kisil, F. T., and Sehon, A. (1990) *Int. Arch. Allergy Appl. Immunol.* **91**, 362-368
19. Mohapatra, S., Hill, R., and Sehon, A. (1990) *Epitopes of Atopic Allergens* (Sehon, A. H., Kraft, D., and Kunkel, G., eds) pp. 61-64, U. C. B. Institute of Allergy, Brussels
20. Mohapatra, S., Hill, R., and Sehon, A. (1990) *Proceedings of the First Pan-American Symposium on Aerobiology and Health* (Vijay, H. M., and Comotosia, P., eds) Health and Welfare Canada, University of Montreal Press, in press
21. Kaslow, D. C. (1986) *Nucleic Acid Res.* **14**, 6767-6775
22. Tabor, S., and Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 4767-4771
23. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105-132
24. Fraga, S. (1982) *Can. J. Chem.* **60**, 2606-2610
25. Eisenberg, D. (1984) *Annu. Rev. Biochem.* **53**, 595-623
26. Alwine, J. C., Kemp, D. J., and Stark, G. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5350-5354
27. Guo, L., Stepien, P. P., Tso, J. Y., Brousseau, R., Narang, S., Thomas, D. Y., and Wu, R. (1984) *Gene (Amst.)* **29**, 251-256
28. Lutcke, H. A., Chow, K. C., Mickel, F. S., Moss, K. A., Kern, H. F., and Scheele, G. A. (1987) *EMBO J.* **6**, 43-48
29. Birnstiel, M. L., Busslinger, M., and Strub, K. (1985) *Cell* **41**, 349-359
30. Von Heijne, G. (1983) *Eur. J. Biochem.* **133**, 17-21
31. Von Heijne, G. (1985) *J. Mol. Biol.* **184**, 99-105
32. Kornfeld, R., and Kornfeld, S. (1985) *Annu. Rev. Biochem.* **54**, 631-664
33. Cottam, G. P., Moran, D. M., and Standring, R. (1986) *Biochem. J.* **234**, 305-310
34. Esch, R. E., and Klapper, D. G. (1989) *Mol. Immunol.* **26**, 557-561
35. Chua, K. Y., Stewart, G. A., Thomas, W. R., Simpson, R. J., Dilworth, R. J., Plozza, T. M., Turner, K. J., and Nisbet, A. (1988) *Int. Arch. Allergy Appl. Immunol.* **85**, 127-129
36. Attassi, A. A. (1990) *Epitopes of Atopic Allergens* (Sehon, A. H., and Kunkel, G., eds) pp. 33-40, U. C. B. Institute of Allergy, Brussels
37. Marsh, D. G., Goodfriend, L., King, T. P., Lowenstein, H., and Platts-Mills, T. A. E. (1986) *WHO Bulletin* **64**, 767-770