



CLONING, SEQUENCING AND IMMUNOLOGICAL CHARACTERIZATION OF Dac g 3, A MAJOR ALLERGEN FROM *DACTYLIS GLOMERATA* POLLEN

CLAUDINE GUÉRIN-MARCHAND,* HÉLÈNE SÉNÉCHAL,
 ANNE-PASCALE BOUIN, VIRGINIE LEDUC-BRODARD, GEORGES TAUDOU,
 ANNE WEYER, GABRIEL PELTRE and BERNARD DAVID

Unité d'Immuno-Allergie, Institut Pasteur, Paris, France

(First received 3 November 1995; accepted in revised form 17 January 1996)

Abstract—Preliminary work showed that a 14-kDa allergen with a pI of 9 was recognized by more than 60% of sera from *Dactylis glomerata* (Dac g) pollen-allergic individuals. The N-terminal amino acid sequence of this Dac g allergen was determined by Edman degradation and compared with that of Lol p 3, a major allergen of *Lolium perenne*. A sequence identity of 65% was found, suggesting that the Dac g allergen could be the homologue of Lol p 3 and therefore named Dac g 3. We report the cloning and sequence analysis of a cDNA encoding the Dac g 3 pollen allergen. The recombinant allergen (rDac g 3) expressed in plasmid vector pGEX-2T contained IgE-reactive epitopes found in its natural counterpart, and induced histamine release from basophils of Dac g-allergic individuals, confirming that the recombinant protein has biological properties similar to the pollen extracted allergen. Computer analyses showed that, in spite of a high degree of sequence homology, even closely related allergens such as Dac g 3 and Lol p 3 have dissimilar predictive secondary structures and potential different antigenicity. Because it possesses the properties of the native counterpart, rDac g 3 could be a relevant tool for molecular studies in allergy. Copyright © 1996 Elsevier Science Ltd.

Key words: Allergy, pollen, recombinant, IgE, histamine release.

INTRODUCTION

Studies performed on recombinant allergens and their corresponding peptides during the past few years have already contributed to unravelling the immunochemical nature of allergenic epitopes present on the molecules of natural allergens. Recent advances in the field of cellular immunology have led to new insights into the molecular mechanisms operating in allergy, and introduced new hypotheses for more specific immunotherapeutic approaches. Previous work revealed cross-reactivity between B epitopes of natural allergens from the same species (van Ree *et al.*, 1994) and between B epitopes of allergens from different related species of grass pollen

(Knox *et al.*, 1989; Klysner *et al.*, 1992). Cross-reactivity has also been demonstrated at the level of T cell epitopes by studies on human T cell responses (Baskar *et al.*, 1992). Furthermore, several allergens are expressed in forms differing from each other by limited sequence variations, and which have been termed “variants”. Some of these variants show distinct biochemical and antigenic properties (Perez *et al.*, 1990; Griffith *et al.*, 1991). This has led to the classification of a group nomenclature based on structural and immunological homologies. The biological importance of most pollen allergens is still unknown except for allergens related to profilins (Valenta *et al.*, 1991a) or to Ca²⁺ binding protein families (Seiberler *et al.*, 1994).

Approximately 20% of the population in developed countries suffers from IgE-mediated atopic diseases (Wüthrich, 1989). Wind-pollinated species of grasses contribute significantly to this seasonal atopic allergy (Thomas, 1991). In temperate latitudes, pollen from *Dac g* represents a major allergen contribution, together with pollen from *Phleum pratense* and *Lolium perenne* (Baldo *et al.*, 1982).

One basic allergen of *Dac g* has been characterized and partly purified in our laboratory. More than 65% of sera from *Dac g* pollen-allergic patients recognize this allergen, which can therefore be considered as a major one (Marsh *et al.*, 1986; WHO/IUIS, 1994). Comparison

*Service d'Immuno-Allergie, Institut Pasteur 25–28, rue du Docteur Roux 75724, Paris, Cedex 15 France. e-mail: clobert@pasteur.fr

Abbreviations: *Dac g*, *Dactylis glomerata*; rDac g 3, recombinant Dac g 3 protein without glutathione-S-transferase; s.p. Dac g 3, semi-purified natural Dac g 3; IEF, isoelectric focusing; GST, glutathione-S-transferase; a.p. anti-rDac g 3, affinity-purified anti-rDac g 3 antibodies; RACE, rapid amplification of cDNA ends; CSDS, cycle sequencing of double-stranded DNA system; GST-Dac g 3, recombinant fusion-protein Dac g 3 before thrombin cleavage; GCG, Genetics Computer Group; *Dac g* 3, genomic DNA of Dac g 3.

of the N-terminal amino acid sequence with that of Lol p 3 (Ansari *et al.*, 1989) suggested that this *Dac g* allergen is the homologue of Lol p 3, and thus this allergen is referred to as *Dac g* 3. The latter was cloned and characterized for its immunochemical and biological activities. It is the second cloned *Dac g* pollen allergen, after *Dac g* 2 (Roberts *et al.*, 1993).

In the present paper, cloned *Dac g* 3 is shown to contain IgE-reactive epitopes found in its natural counterpart and to elicit histamine release from basophils of *Dac g*-allergic individuals, confirming that the recombinant protein has biological properties similar to the natural protein. Predictive structural and antigenic properties of groups 2 and 3 of *Dac g* and Lol p 3 pollen allergens have also been compared.

Dac g 3 might be part of a panel of relevant recombinant allergens which will serve as future tools for the study of grass pollen allergy.

MATERIALS AND METHODS

Preparation of semi-purified Dac g 3 (s.p. Dac g 3) from water-soluble pollen extract

Preparative isoelectric focusing (IEF) was carried out on *Dac g* pollen hydrosoluble extract according to Mechéri *et al.* (1985) and the band containing *Dac g* 3 was eluted and applied to an FPLC column (Superose 12; Pharmacia Biotech, Uppsala, Sweden). Fractions were analysed by SDS-PAGE.

Selection of patient sera

Thirty grass-pollen-sensitive patients were selected as previously described (Batard *et al.*, 1993). One serum from an allergic patient without sensitivity to *Dac g* 3, and another from a non-allergic individual were selected for control experiments.

Affinity purification of rabbit anti-rDac g 3 antibodies

The purified recombinant allergen was immobilized on a 2-cm² nitrocellulose membrane strip and incubated for 2 hr in diluted sera of the rDac g 3 immunized rabbits. The bound antibodies (affinity purified or a.p. anti-rDac g 3) were eluted with 1 ml of 0.2 M glycine pH 2.5 and neutralized with Tris base 2 M.

RNA isolation from Dac g pollen

Total RNA was isolated from 5 g of *Dac g* pollen (Allergon AB; Engelholm, Sweden, batch 010989902) according to a technique previously described (Walsh *et al.*, 1989).

Cloning and sequencing of Dac g 3 cDNA

The s.p. *Dac g* 3 was separated by SDS-PAGE and transferred to a poly (vinylidene-difluoride) membrane (PVDF) for N-terminal microsequencing and enzymatic digestion. After HPLC separation, one peptide, the absorbance profile of which suggested an amino acid

composition similar to the Lol p 3 C-terminal peptide, was chosen for sequencing. The obtained sequence confirmed the C-terminal position of the chosen peptide. Then, the oligonucleotides corresponding to the N- and C-ends of the protein were synthesized. A general codon usage of the plants (Murray *et al.*, 1989) was followed to choose certain codons. For others, a mixture of oligonucleotides was used and inosine was introduced at the four-fold degeneracy positions. First strand cDNA was synthesized from 1 µg of total RNA from the *Dac g* pollen using 2 pmol of the P2 primer, 1 µl (200 units) of Superscript reverse transcriptase (Gibco Eragny, France), 2 µl of 10 × corresponding buffer, 1 µl of 10 mM of each deoxynucleoside triphosphate (dNTPs) and 2 µl of 0.1 M dithiothreitol according to the manufacturer's instructions. PCR was performed in a volume of 50 µl. The sample contained 0.4 units (0.08 µl) of GoldStar DNA polymerase from Eurogentec (Seraing, Belgium), and the appropriate buffer was supplemented with 1.5 mM MgCl₂/0.2 mM of each dNTPs, 1 µM each of the P1 and P2 primers and 10 µl of the first strand synthesis mixture. The samples were subjected to 35 cycles of amplification consisting of the following steps: 1 min 30 s at 94°C, 1 min 30 s at 55°C, 1 min at 72°C, followed by a final incubation at 72°C for 10 min. About 15 ng of the purified PCR product were used for the TA cloning procedure according to the supplier's protocol (TA cloning kit containing the pCR 2 vector, from Invitrogen, San Diego, U.S.A.). DNA sequences of two clones were determined by double-stranded sequencing carried out with Sp6 and T7 promoters and using Sequenase (U.S. Biochemicals, Cleveland, U.S.A.). Based on these sequences, a number of oligonucleotides were synthesized, some of them containing an *Eco* RI restriction site added for cloning purposes (Fig. 1). All oligomers were synthesized by Eurogentec.

3'-End amplification with 3'-RACE (rapid amplification of cDNA ends system)

Dac g 3 3'-end amplification was performed with our gene-specific primer P3 and the amplification primer provided with the 3' RACE system (Gibco BRL, Life Technologies, Gaithersburg, U.S.A.). PCR amplification of the cDNA was performed with the GoldStar DNA polymerase according to the supplier's protocol.

5'-End amplification with 5'-RACE system

First strand cDNA synthesis was initiated from our sequence-specific P2 primer. After purification as recommended by the manufacturer, an oligo-dC anchor sequence provided with the system was added to the 3'-end of the cDNA using the terminal deoxynucleotidyl transferase and dCTP. PCR amplification was carried out using our nested (internal) gene-specific P4 primer and the anchor primer provided with the system.

Isolation of DNA from pollen

Isolation of DNA was performed using the CTAB (hexadecyltrimethylammonium bromide) procedure as

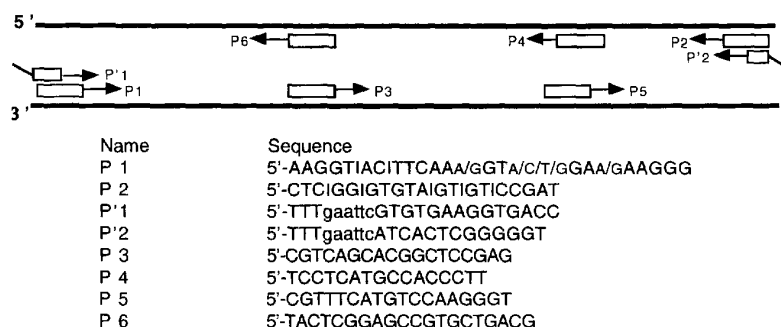


Fig. 1. Schematic location of oligonucleotide primer hybridization sites for amplification and sequencing and sequence of corresponding oligonucleotides. The degenerated nucleotides are shown in small caps. I, inosine. *Eco* RI restriction sites are shown in lower-case letters.

described (Doyle and Doyle, 1987). About 300 ng of this DNA was submitted to PCR amplification with P'1 and P'2 primers under the same conditions as described earlier.

Direct sequencing of PCR products with the CSDS system (cycle sequencing of double stranded DNA)

PCR products of 3'- and 5'-amplifications were purified after a separating step on NuSieve agarose gel, as already described. In a first step, radioactivity was transferred to the 5'-end of the sequencing primer using the T4 polynucleotide kinase and [γ - ^{32} P]ATP (DuPont de Nemours, NEN products, Les Ulis, France). Sequencing reactions were performed according to the CSDS system (Gibco BRL). The P5 primer was used for 3'-RACE PCR product sequencing, and P6 for the 5'-RACE PCR product sequencing. For genomic DNA from pollen, the sequence was determined by CSDS carried out with either the P'1 or P'2 primer.

Expression and purification of the recombinant antigen rDac g 3

The coding sequence corresponding to the Dac g 3 protein was cloned into a pGEX-2T vector (gene fusion system from Pharmacia) which offers a thrombin recognition site for cleaving the allergen from the fusion product which is GST. After cloning in the DH5, an *Escherichia coli* strain, clones were assayed for protein fusion expression. One of the expressing clones was sequenced with the 5' pGEX sequencing primer to ensure the absence of any mutation. The fusion protein was expressed and purified as described in the supplier's protocol using a column of glutathione Sepharose 4B (Pharmacia). This fusion protein (GST-Dac g 3) was used in some experiments. Thrombin (Boehringer Mannheim GmbH, Sandhofer Strasse 116, Mannheim, Germany) cleavage was performed on the fusion protein bound to the column matrix. The purity of the digestion product was evaluated on SDS-PAGE gel performed on the Phast system (Pharmacia), followed by silver staining (Blum *et al.*, 1987). Identical preparations (without the cleavage step) were done with the product of pGEX-2T without the insert to obtain GST alone.

Computing studies

These were carried out with the sequence analysis software package version 7.1 of GCG (Genetics Computer Group Inc., Madison, WI).

Immunochemical comparisons between native and rDac g 3 by competition assay

After SDS-PAGE of hydrosoluble Dac g pollen extract and blotting on nitrocellulose, a band of 10-cm length and 1-mm width, corresponding to Dac g 3 was cut out and divided into identical fragments. In fact this band was containing both Dac g 2 and Dac g 3 allergens, due to their similar apparent molecular mass.

Sera from 18 Dac g 3-sensitive patients and from two non-allergic individuals were diluted 1:5 in phosphate buffer containing 1% bovine serum albumin/0.1% Tween 20 (PBST BSA) and incubated with rDac g 3 (250 μ l of dilution and 20 μ g of rDac g 3) at 4°C overnight. The adsorbed sera were diluted 1:2 and checked for antibody depletion. The completely depleted sera were then incubated overnight with a fragment of blocked nitrocellulose. The amount of bound IgE (B) was determined with ^{125} I-labelled anti-human IgE (Pharmacia RAST kit) and counted in a gamma counter LKB Gamma master 1277. A ten-fold dilution of sera before depletion was incubated with the same corresponding piece of nitrocellulose at 4°C overnight, and the total labelled anti-human IgE (C) was counted. The percentage of inhibition was calculated using the following formula: % inhibition = $100 \times [(A-B)/A]$ where A represents the IgE bound with undepleted sera the value of which was (C-B) in cpm.

SDS-PAGE

Samples (67 μ g of rDac g 3 and an equivalent of 1.6 mg of extracted pollen) treated with SDS and β -mercaptoethanol were separated on a 15% T, 4% C acrylamide gel and transferred on a CNBr-activated nitrocellulose filter (Batard *et al.*, 1993).

Two-dimensional separation

Two-dimensional separation of hydrosoluble extract corresponding to 800 μ g of Dac g pollen and blotting

were performed according to Leduc-Brodard *et al.*, 1994.

Immunoglobulin detection

IgE detection was performed as described earlier (Battard *et al.*, 1993). Patient sera were diluted 10 times in PBST BSA a.p. Anti-rDac g 3 rabbit antibodies were used at a 1:20 dilution and anti-s.p. Dac g 3 mouse antibodies were diluted 1:100 in the same buffer.

Inhibition ELISA

The ELISA protocol followed the general procedure of Voller *et al.* (1978). Maxisorb microtitre plates (Nunc, Rockville, Denmark) coated with s.p. Dac g 3 (2 µg/ml) were incubated with a 2:1 mixture of increasing rDac g 3 or s.p. Dac g 3 or GST concentrations and patient sera. IgE antibodies bound to solid-phase s.p. Dac g 3 were measured with polyclonal equine antibody anti-human IgE (Sanofi Diagnostics Pasteur, Chaska, U.S.A.) biotinylated. Bound anti-IgE was revealed by the addition of alkaline phosphatase-conjugated streptavidin (Amersham, Little Chalfont, Buckinghamshire, England) diluted 1:5000. Bound enzyme was detected with 1 mg × ml⁻¹ of p-nitrophenyl-phosphate-disodium (Sigma, St Louis, MO, U.S.A.). Absorbance was measured using a test wavelength of 405 nm and a reference wavelength of 630 nm. The percentage of inhibition was calculated using the following formula: % inhibition = 100 × [(A-B)/A], where A represents absorbance without the inhibitor and B absorbance with the inhibitor.

Histamine release

The basophil histamine release sensitivities of grass-pollen allergic patients were assayed with GST alone, GST-Dac g 3, rDac g 3 and s.p. Dac g 3. Histamine release experiments were performed as described earlier (May *et al.*, 1970; Le Mao *et al.*, 1980). Histamine content was measured by the fluorometric method described by Lebel (1983).

RESULTS

Amplification, cloning and sequencing of cDNA coding for the Dac g 3 allergen

RT-PCR (PCR following synthesis of the first cDNA strand using reverse transcriptase) carried out on *Dac g* pollen RNA using the P1 and P2 degenerated oligonucleotides amplified a single band of 300 bp. After cloning of this PCR product in the pCR II vector, two clones were obtained, the sequences of which were identical. Based on this information, several oligonucleotides were designed, some of them having a restriction site for cloning purposes in the pGEX-2T expression vector (Fig. 1). Dac g 2 sequence data (Roberts *et al.*, 1993) were taken into account when choosing Dac g 3 internal oligonucleotides without homology with Dac g 2. High G+C content was calculated (60%) as usual for plant DNA. The Dac g 3 cDNA encoded a lysine-rich putative

protein (12.4%) containing 96 amino acid residues without cysteine, the calculated molecular mass of which was 10.925 kDa. The sequence lacked asparagine sequons (Asn-x-Ser/Thr) required for N-glycosylation (Fig. 2). Two couples of repeated sequences were found. The first was located at the 5'-end and near the 3'-end (Fig. 2, single underlined sequences), whereas the second was situated between nucleotides 115 and 125 and at the very end of the sequence (Fig. 2, double underlined sequences).

Sequencing of the 3'- and 5'-ends of the Dac g 3 cDNA

Using the P3 primer, a band containing 450 bp was amplified using the 3' RACE system.

Cycle sequencing of the PCR product using the P5 primer led to obtaining the Dac g 3 3'-end sequence, and showed the existence of an additional Glu amino-acid residue before the stop codon. This residue was lacking in the C-terminal trypsinic peptide. At several positions, the sequence following the stop codon presented two or three bands at the same level (data not shown), indicating that at least three different matrices were simultaneously sequenced. This suggested polymorphism within the polyadenylation consensus site.

A PCR product 400 bp long was obtained using the 5'-RACE system. The P6 primer chosen for the cycle sequencing of the PCR product gave a sequence at which no Met codon was shown at the 5'-end of the cDNA, suggesting the presence of a putative signal peptide before the first Val residue identified by micro-sequencing of the 5' N-terminal peptide. Several bands at the same level found in the signal peptide sequence suggested polymorphism in this region. The whole amino acid sequence was impossible to determine, because at least three matrices interfered. Nevertheless, Ala, Cys and Ala were identified without ambiguity at positions -3, -2 and -1, respectively, and it can be assumed that the prevalent structure constituted by the other amino acid residues was hydrophobic.

The oligonucleotide sequences constructed using the codon preference in plants and the real sequences obtained by 3' and 5' RACE were found to be very similar. This feature could explain why amplification of a single band was easily achieved at the first PCR amplification.

Sequence of the Dac g 3 genomic DNA

A 300-bp band was amplified using the P'1 and P'2 primers, indicating that the *Dac g* 3 gene has no intron. The sequence of the purified PCR product was carried out with the CSDS system in the two orientations. There was no evidence of polymorphism in this part of the *Dac g* 3 gene sequence.

Computer analyses

The investigation carried out using the GenBank data base provided an indication of significant homology between nucleotide sequences of Dac g 3, and four other mRNA from different plant allergens, *Lolium perenne*

```

1/1                               31/11
GTG AAG GTG ACC TTC AAG GTG GAG AAG GGT TCC GAC CCC AAG AAG CTG GTT CTC GAC ATC
V K V T F K V E K G S D P K K L V L D I
61/21                               91/31
AAG TAC ACC AGG CCA GGG GAC ACC CTT GCG GAG GTT GAG CTC CGT CAG CAC GGC TCC GAG
K Y T R P G D T L A E V E L R Q H G S E
121/41                               151/51
GAG TGG GAG CCC CTG ACG AAG AAG GGC AAC CTG TGG GAG GTG AAG AGC TCC AAG CCG CTC
E W E P L T K K G N L W E V K S S K P L
181/61                               211/71
ACT GGC CCC TTC AAC TTC CGT TTC ATG TCC AAG GGT GGC ATG AGG AAC GTC TTC GAC GAG
T G P F N F R F M S K G G M R N V F D E
241/81                               271/91
GTC ATC CCC ACC GCC TTC AAG ATC GGC ACC ACC TAC ACC CCA GAG GAG TGA
V I P T A F K I G T T Y T P E E *

```

Fig. 2. Dac g 3 cDNA nucleotides and inferred amino acid sequences. The GenBank accession number for Dac g 3 is U25343. Repeated sequences are single and double underlined.

(Lol p 2), *Phleum pratense* (Phl p 1, 2), *Holcus lanatus* (Hol l 1), and Dac g 2.

Figure 3 shows the sequences of Dac g 3, Dac g 2 and Lol p 3. With the *GAP* program which uses the algorithm of Needleman and Wunsch (1970), we found 92.6% similarity and 84.2% identity between Dac g 3 and Lol p 3. When matching Dac g 3 and Dac g 2, the percentage value was 78.7 for similarity and 65.9 for identity (the gap weight was 3.00 and the length weight was 0.1). However the *PLOTSTRUCTURE* program which plots the measurements of secondary structure in a two-dimensional squiggly representation, predicted important conformational differences between these three proteins (data not shown).

The antigenic index for the same three allergens was calculated as defined by Jameson-Wolf (Jameson and Wolf, 1988) in the *PEPTIDESTRUCTURE* program. Five main antigenic regions were predicted for Dac g 3. Dac g 2 had five similar regions plus one, and Lol p 3 lacked the fifth antigenic site, whereas it presented an important site at the C-terminal end of its molecule. These data are reported as a panel of graphs in Fig. 4, which shows the composition and location of the predicted sites.

Characterization of s.p. Dac g 3

After two different steps of purification we only obtained a semi-purified product, s.p. Dac g 3. On a SDS-PAGE gel, contaminations were undetectable by

Coomassie staining, but using the serum of mouse immunized with the s.p. Dac g 3, on Western blot, we detected an additional 13-kDa band that was, in fact, Dac g 10 (Fig. 5, lane 1, lower band). Dac g 10 is a cytochrome C-like protein with a *pI* >9 poorly recognized by IgE of Dac g 3-sensitive patients.

Antigenic/allergenic characterization of rDac g 3

Both native (upper band on lane 1, Fig. 5) and recombinant (lane 2, Fig. 5) proteins showed an apparent molecular mass of 14 kDa, and thus presented the same dramatic difference with the calculated molecular weight of 10.925.

A total of 67% of the 30 patient sera assayed contained IgE directed against native and rDac g 3 proteins (Fig. 6, A and B). The reactivity of the native protein correlates with that of the recombinant protein; for example, Dac g 3 specific IgE is missing in sera 21 and 28, as well with hydrosoluble antigen as with the rDac g 3.

Because Dac g 3 and Dac g 2 have approximately the same apparent molecular mass, the band cut from the transferred membrane and used for the competition assay was containing both proteins. Results of IgE inhibition studies using rDac g 3 vs blotted Dac g 3 are shown in Table 1. For four sera, 100% of the IgE directed against the blotted Dac g 3 were captured by the recombinant counterpart; nine sera were inhibited between 60 and 90%, and five sera showed inhibition of between 40 and

```

1                               50
Dac g 3 .VKVTRKVEK GSDKKLVLD IKYTRPGDTL AEVELRQHGS EEWEPLTKKG
Lol p 3 .TKVDLTVEK GSDAKTLVLN IKYTRPGDTL AEVELRQHGS EEWEPLTKKG
Dac g 2 EAVVTFTVEK GSDEKNLALS IKYNKEGLSM AEVELKEHGS NEWLALKKNG

51                               99
Dac g 3 N.LWEVKSSK PLTGPFNFRF MSKGGMRNVF DEVIPTAFKI GTTYTPPEE.
Lol p 3 N.LWEVKSAK PLTGPFNFRF LSKGGMRNVF DEVIPTAFIV GRITYTPFYN
Dac g 2 DGVWEIKSKK PLKGPFFNFRF VSEKGGMRNVF VDMVPADFVK GTTYKPEE.

```

Fig. 3. Homologies of Dac g 3 with Lol p 3 and Dac g 2. Alignment of deduced amino acid sequence of Dac g 3 with reported sequences of Lol p 3 (Ansari *et al.*, 1989) and Dac g 2 (Roberts *et al.*, 1993). Identical residues are indicated by boxes. Dac g 2 and Dac g 3 sequences are deduced amino acid sequences whereas Lol p 3 sequence was determined by automated Edman degradation of the protein.

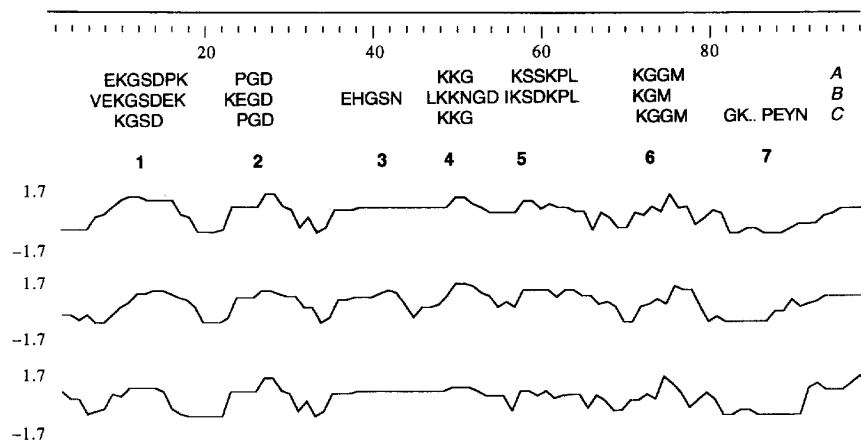


Fig. 4. Graphs of the antigenicity index of Dac g 3 (A), Dac g 2 (B) and Lol p 3 (C) allergens. The three graphs are from the PLOTSTRUCTURE study from GCG, which makes secondary structure predictions for peptide sequences and measures the antigenic index according to Jameson-Wolf. The position and sequence of the corresponding predicted antigenic sites are shown above the graphs.

60%. We demonstrated that the incomplete inhibition of sera was due to their positivity for a specific epitope of Dac g 2 on the Western-blotted band. For this purpose, three sera negative against Dac g 10 were chosen among the incompletely inhibited sera (sera 17, 10 and 19) and assayed in an ELISA inhibition test using the s.p. Dac g 3 protein (contaminated with Dac g 10) as a reference.

As expected, nearly total inhibition was obtained with the three sera (Fig. 7).

Two-dimensional analysis of immunological cross-reactivity

The a.p. anti-rDac g 3 rabbit antibodies were assayed on Dac g proteins separated by two-dimensional electrophoresis. Dac g 3 and Dac g 2 were detected as two distinct spots, whereas Dac g 1 appeared weakly (arrow, Fig. 8). These results show the cross-reactivity of Dac g 1, 2 and 3.

Histamine release

The recombinant allergen, as well as the same allergen linked to GST, elicited histamine release responses similar to that triggered by s.p. Dac g 3 (data not shown). These dose responses displayed the bell-shaped curves observed with other allergens (Lichtenstein and Osler, 1964; Valenta *et al.*, 1993). GST alone presented no significant histamine release effect. Among the three Dac g-sensitive patients studied, two displayed positive histamine release results, whereas the third presented negative results in the dose range of allergen concentrations tested. Basophils from a normal control and from an allergic individual without sensitivity to Dac g also gave negative results.

DISCUSSION

In the present paper we report the molecular cloning, sequencing and expression of a cDNA encoding a major allergen from Dac g. As we expected at the beginning of this work, the clone we isolated was homologous to the Lol p 3 pollen allergen in *Lolium perenne*, and was therefore called Dac g 3. This clone has the same sequence as genomic DNA, which is devoid of introns.

Whatever the sequencing technique used (the Sanger procedure after cloning or the DNA cycle sequencing), no variations were found between the different sequence

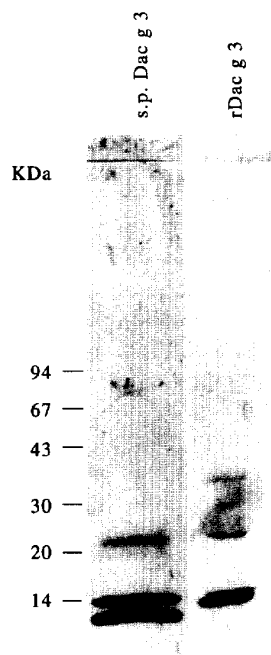


Fig. 5. Immunoblot of s.p. Dac g 3 (lane 1) and rDac g 3 (lane 2). Anti-s.p.-Dac g 3 mouse antibodies detected two principal bands corresponding to Dac g 3 and Dac g 10, and a non-identified weak band (lane 1). Lane 2: The immunodetection by a.p. anti-rDac g 3 rabbit antibodies shows that Dac g 3 (upper band in lane 1) and rDac g 3 (lane 2) have the same molecular mass.

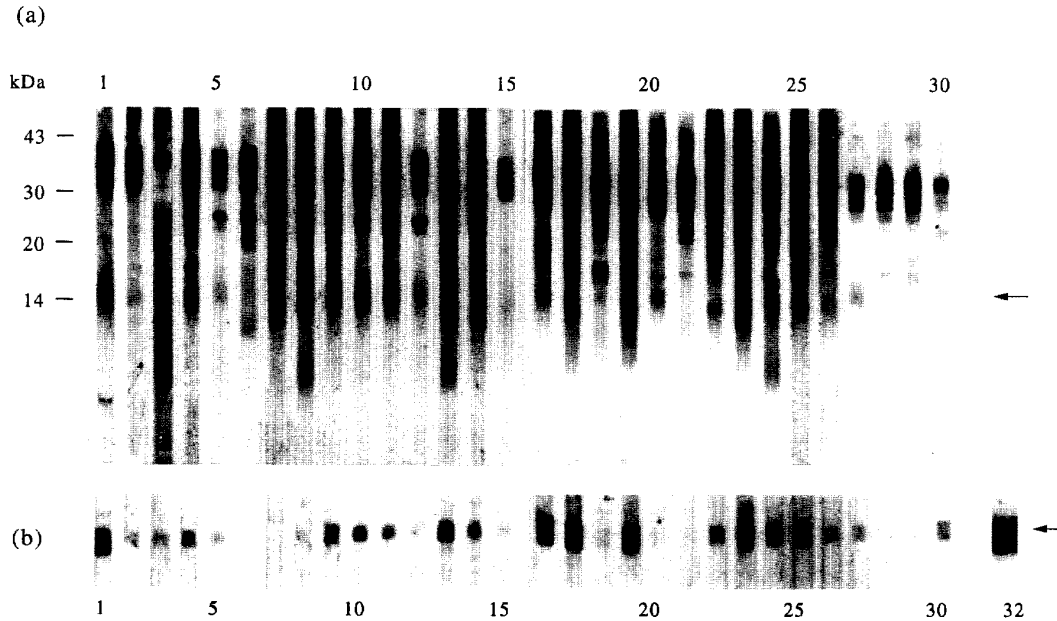


Fig. 6. Comparison of patient IgE reactivity with natural (a) and recombinant allergens (b) on Western blot. Strips 1–30 were tested with IgE from sera of patients with *Dac g*-allergy, and 31 with serum from a non-allergic patient. a.p. anti-rDac g 3 rabbit antibodies were used to identify rDac g 3 (arrow, lane 32). In (a) the arrow localizes Dac g 3 and Dac g 2 which have the same apparent molecular mass. Sera nos. 27–31 were assayed in histamine release and sera 10, 17 and 19 were tested for inhibition of IgE binding in ELISA.

Table 1. Percentage inhibition by rDac g 3 of anti-Dac g 3 IgE from sera of 18 patients

Sera	1	2	4	5	9	10	11	13	14	16	17	19	22	23	24	26	27	30
% Inhibition	100	100	79	86	78	78	77	70	100	47	53	41	100	68	88	55	67	38

Sera from Dac g 3 sensitive patients were completely depleted in rDac g 3 antibodies. The remaining IgE directed against Dac g 3 were detected on Western-blotted Dac g 3 from hydrosoluble extract, with ^{125}I -labelled anti-human IgE.

readings of the mature protein. There appears to be no variant of the Dac g 3 mature protein, but we cannot exclude the existence of a poorly represented variant.

The analysis shows two short repeated sequences (Fig. 2). However, this feature did not disturb amplification

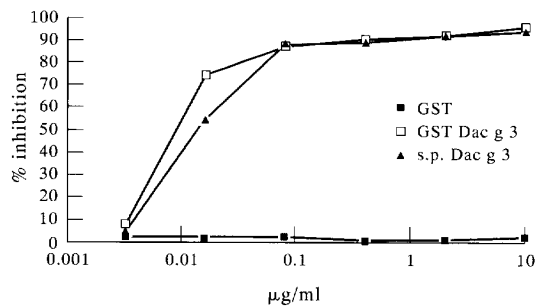


Fig. 7. The ability of GST-Dac g 3 to inhibit IgE binding on coated s.p. Dac g 3 was assayed with sera from three patients (10, 17, 19). s.p. Dac g 3 and GST were used as controls. Because the results were nearly identical for the three sera, the figure shows only the curve obtained with serum 17.

performed with the N-terminal primer because the repeated stretch was located at the 3'-end of the Dac g 3 sequence. Furthermore, CSDS using P6 primer was performed with high hybridization stringency to prevent mismatches.

As methionine was not found at the N-terminal end of the protein, we assume that amino acid residues located before Val belong to a putative signal peptide. Results obtained from sequencing of the 5'-amplification product showed several bands at the same level (data not shown) as it is obtained in simultaneous sequencing of different matrices. Under these conditions, only three amino acids were identified without ambiguity at positions –1, –2 and –3 (Ala, Cys, Ala, respectively). For the other positions, different amino acids were possible. We can only assert that the general aspect of the putative signal peptide was hydrophobic. Furthermore, the –1 and –3 rule for the prediction of a sequence cleavage site which imposes small and neutral amino acid at these positions (von Heijne, 1984) was adhered to, because Ala was identified in both positions.

Sequencing of the 3'-end amplification product of the

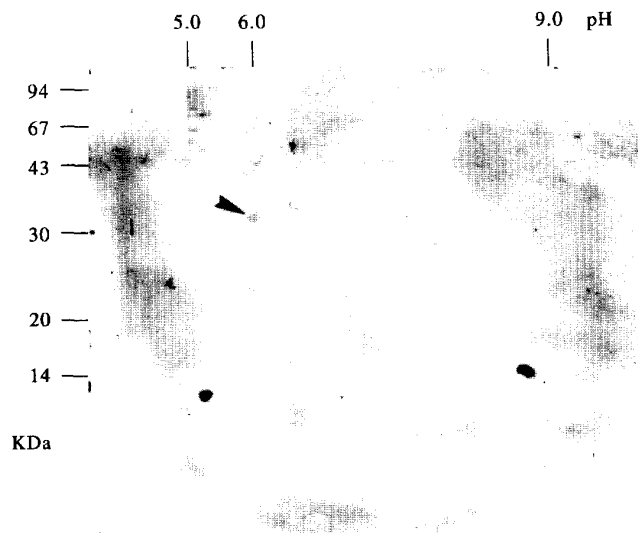


Fig. 8. Two-dimensional polyacrylamide gel electrophoresis of water-soluble extract from *Dac g* pollen. Sample corresponding to 800 μ g of pollen, was subjected to IEF (pH 3.5–10) in acrylamide in the first dimension and SDS-PAGE with the Tris-Tricine buffer system in the second dimension. Proteins were blotted onto activated nitrocellulose and immunodetected with a.p. anti-rDac g 3 rabbit antibodies. Arrow: Dac g 1 (pI=6, 35 kDa). Dac g 2 (pI=5.3, 14 kDa), Dac g 3 (pI=9, 14 kDa). The spots in the region of >43 kDa from acidic to basic region are probably artefacts due to insoluble material.

Dac g 3 cDNA revealed an additional Glu at the last position before the stop codon. This amino acid residue was missing in the sequenced C-terminal peptide and consequently, the corresponding codon could not be introduced in the P'2 primer. The recombinant Dac g 3 protein is therefore truncated from one Glu residue and contained an additional Arg residue at the N-terminal end, due to the P'1 structure (see Fig. 1). rDac g 3 is therefore more basic than its natural counterpart when compared by isoelectric focusing (data not shown). Like the 5'-end sequence, the polyadenylation consensus sequence presents several bands at the same level suggesting polymorphism in this region. Because this special feature was not shown in the sequence from P5 to stop codon, and sequence from P6 to the first Val codon, we can exclude that primer degradation and/or inappropriate hybridization stringency were involved. Then we assume that the putative signal peptide and the polyadenylation site are polymorphic.

The Dac g 3 sequence was compared with those found in the GenBank data base. Significant homologies were found with six other plant allergens (Lol p 2, Phl p 2, Phl p 1, Hol l 1, Dac g 2, and Lol p 1a). The function of these pollen allergens is not yet known. However, the finding that similar proteins are present in the pollen of several grasses suggests that they may be important for these sexual plant cells. None of the short similarities found with other sequences could be considered significant.

Computer studies showed that, in spite of 84% identity and 92.6% similarity, Lol p 3 and Dac g 3 present important differences in their predicted secondary structure (data not shown) and at the level of the predicted antigenic sites (Fig. 4). One Dac g 3 predicted site located

between amino acid positions 55 and 58 is absent in Lol p 3, and the general antigenic index of Lol p 3 is lower than that of Dac g 3. Furthermore, an additional predicted antigenic site figures at the very C-terminal of Lol p 3, just following a hydrophobic region which does not exist in Dac g 3. It is situated in the stretch of the sequence coding for the antigenic peptide described by Esch and Klapper (1989) and which van Ree *et al.* (1994) demonstrated to be responsible for the immunological cross-reactions between Lol p 1, Lol p 2 and Lol p 3.

A comparison of Dac g 2 and Dac g 3 with the *GAP* system showed 64.5% identity and 78% similarity. At the level of the Dac g 2 antigenicity, good conservation of the five antigenic sites predicted in Dac g 3 was observed, with an additional site localized between residues 39 and 41. The existence of this supplementary antigenic site may explain why some patient sera still recognized the 14-kDa band (containing Dac g 3 and Dac g 2) in Western-blot even after incubation with an excess of rDac g 3 (see Table 1).

The main difficulty encountered when comparing the immunological properties of rDac g 3 with its natural counterpart lies in the failure to obtain enough of the completely purified natural allergen due to similar molecular mass and/or isoelectric point. The protein purified by chromatography was contaminated by Dac g 10, and only a few micrograms were available. Furthermore, the band cut off from the blotted membrane and used for IgE inhibition studies was containing Dac g 2, which has the same apparent molecular mass as Dac g 3. In spite of these contaminations, the use of two complementary inhibition assays enabled us to assert that rDac g 3 is a good replica of the natural allergen and that, indeed, Dac

g 2 possesses specific additional epitope(s) as suggested by computer analyses.

Antibodies elicited in rabbits by rDac g 3 injections and immuno purified on rDac g 3 revealed a cross-reaction between Dac g 2, Dac g 3, and Dac g 1 on a two-dimensional separation of the *Dac g* proteins. In the case of the *Lol p* allergen, van Ree *et al.* (1994) demonstrated that immunological cross-reactions between *Lol p* 1, 2 and 3 were due to the presence of an antigenic site located in the Esch peptide at the C-terminal part of the molecules. Because the Dac g 1 sequence is still unknown we cannot propose any hypothesis about the cross-reactivity between Dac g 1, 2 and 3.

CONCLUSIONS

rDac g 3 is a valid recombinant allergen: it has a structure similar to the natural allergen, apparently possesses all the IgE-reactive epitopes and is well expressed in *E. coli*. This is not the case for all the recombinants which are often not glycosylated, are incorrectly folded, weakly recognized by IgE antibodies or poorly synthesized in *E. coli*. Because it is necessary to have available a wide range of allergens for diagnostic purposes, we believe that the use of combinations of recombinant allergens instead of natural crude extracts (Valenta *et al.*, 1991b; Olsen and Mohapatra, 1994) will be a difficult task. Furthermore, in spite of important intra- and inter-species similarities and cross-reactivities, it has been shown in this work that even closely related allergens such as Dac g 2 and Dac g 3 presented theoretically predicted and experimentally observed different antigenicity.

Nevertheless, recombinant allergens might be advantageous when natural allergens are present in low concentrations in extracts, or are unstable. The addition of the corresponding recombinant allergens could improve the natural extract.

Because it is possible to obtain small recombinant peptides and to introduce mutations in recombinant DNA molecules, biologically active recombinant replicas of allergens such as Dac g 3 should help to gain better insight into the cellular and molecular bases of allergy.

Acknowledgements—We thank Dr R. Griffay (Institut Pasteur) for the kind supply of the TA cloning kit, for technical help and useful advice. We also thank E. Bottius and C. Mayer for technical help. R. Peronet immunized rabbits and mice, J. D'Alayer (Institut Pasteur) microsequenced peptides. We are grateful to Eurogentec (Belgium) for gifts of GoldStar polymerase and to DuPont de Nemours (France) for giving [γ -³³P]ATP.

REFERENCES

- Ansari A. A., Shenbagamurthi P. and March D. G. (1989) Complete primary structure of a *Lolium perenne* (perennial rye grass) pollen allergen, *Lol p* III: comparison with known *Lol p* I and II sequence. *Biochemistry* **28**, 8665–8670.
- Baldo B. A., Sutton R. and Wrigley C. W. (1982) Grass pollen allergens with particular reference to cereals. *Prog. Allergy* **30**, 1–66.
- Baskar S., Parronchi P., Mohapatra S., Romagnani S. and Ansari A. A. (1992) Human T cell responses to purified pollen allergens of the grass *Lolium perenne*. *J. Immun.* **148**, 2378–2383.
- Batard T., Basuyaux B., Lambin P., Brémard-Oury C., Hamilton R. G., David B. and Peltre G. (1993) Isotypic analysis of grass pollen-specific immunoglobulins in human plasma. I. Specialization of certain classes and subclasses in the immune response. *Int. Arch. Allerg. Immun.* **100**, 68–78.
- Blum H., Beier H. and Gross H. J. (1987) Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis* **8**, 93–99.
- Doyle J. J. and Doyle J. L. (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **19**, 11–15.
- Esch R. E. and Klapper D. G. (1989) Isolation and characterization of a major cross-reactive grass group I allergic determinant. *Molec. Immun.* **26**, 557–561.
- Griffith I. J., Pollock J., Klapper D. G., Rogers B. L. and Nault A. K. (1991) Sequence polymorphism of Amb a I and Amb a II, the major allergens in *Ambrosia artemisiifolia* (short ragweed). *Int. Arch. Allerg. appl. Immun.* **96**, 296–304.
- Jameson B. A. and Wolf H. (1988) The antigenic index: a novel algorithm for the predicting antigenic determinants. *CABIOS* **4**, 181–186.
- Klysner S., Welinder K. G., Lowenstein H. and Matthiesen F. (1992) Group V allergens in grass pollens: IV. Similarities in amino acid compositions and NH₂-terminal sequences of the group V allergens from *Lolium perenne*, *Poa Pratensis* and *Dactylis glomerata*. *Clin. exp. Allergy* **22**, 491–497.
- Knox R. B., Sing M. B., Hough T. and Theerakulpisut P. (1989) The rye-grass pollen allergen, *Lol p* I. *Allerg. molec. Biol. adv. Biosci.* **74**, 161–171.
- Le Mao J., Weyer A., Pauli G., Lebel B. and David B. (1980) Studies on the *Dermatophagoides Pteronyssinus* allergens: Measurement of the relative potencies of the *D. pteronyssinus* purified extracts by *in vitro* and *in vivo* methods. *J. Allerg. clin. Immun.* **65**, 381–388.
- Lebel B. (1983) A high-sampling rate automated continuous-flow fluorometric technique for the analysis of nanogram levels of histamine in biological samples. *Analyt. Biochem.* **133**, 16–29.
- Leduc-Brodard V., David B. and Peltre G. (1994) Comparison of three horizontal two-dimensional electrophoretic techniques to separate grass pollen allergens. *Cell. molec. Biol.* **40**, 1–8.
- Lichtenstein L. M. and Osler A. G. (1964) Studies on the mechanisms of hypersensitivity phenomena. IX. Histamine release from human leukocytes by ragweed pollen antigens. *J. exp. Med.* **120**, 507–530.
- Marsh D. G., Goodfriend L., King T. P., Lowenstein H. and Platts-Mills T. A. E. (1986) Allergen nomenclature. *Bull. WHO* **64**, 767–770.
- May C. D., Lyman M., Alberto R. and Cheng J. (1970) Procedures for immunochemical study of histamine release from leukocytes with small volumes of blood. *J. Allerg.* **46**, 12–20.
- Mecheri S., Peltre G. and David B. (1985) Purification and characterization of a major allergen from *Dactylis glomerata* pollen: the Ag Dg1. *Int. Arch. Allerg. appl. Immun.* **78**, 283–289.
- Murray E. E., Lotzer J. and Eberle M. (1989) Codon usage in plant genes. *Nucl. Acids Res.* **17**, 477–498.
- Needleman S. B. and Wunsch C. D. (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. molec. Biol.* **48**, 443–455.

- Olsen E. and Mohapatra S. S. (1994) Recombinant allergens and diagnosis of grass pollen allergy. *Ann. Allergy* **72**, 499–506.
- Perez M., Ishioka G. Y., Walker L. E. and Chesnut R. W. (1990) cDNA Cloning and immunological characterization of the rye grass allergen Lol p I. *J. biol. Chem.* **7**, 16210–16215.
- Roberts A. M., Bevan L. J., Flora P. S., Jepson I. and Walker M. R. (1993) Nucleotide sequence of cDNA encoding the group II allergen of cocksfoot/orchard grass (*Dactylis glomerata*), Dac g II. *Allergy* **48**, 615–623.
- Seiberler S., Scheiner O., Kraft D., Lonsdale D. and Valenta R. (1994) Characterization of a birch pollen allergen, Bet v 3, representing a novel class of Ca^{2+} binding proteins: specific expression in mature pollen and dependance of patients' IgE binding on protein-bound Ca^{2+} . *EMBO J.* **13**, 3481–3486.
- Thomas W. R. (1991) Grass under scrutiny. *Clin. exp. Allergy* **21**, 255–257.
- Valenta R., Duchêne M., Pettenburger K., Sillaber C., Valent P., Bettelheim P., Breitenbach M., Rumpoldt H., Kraft D. and Scheiner O. (1991a) Identification of profilin as a novel pollen allergen-IgE autoreactivity in sensitized individuals. *Science* **253**, 557–560.
- Valenta R., Duchêne M., Vrtala S., Birkner T., Ebner C., Hirschwehr R., Breitenbach M. and Rumpold H. (1991b) Recombinant allergens for immunoblot diagnosis of tree-pollen allergy. *J. Allerg. clin. Immun.* **88**, 889–894.
- Valenta R., Sperr W. R., Ferreira F., Valent P., Sillaber C., Tejkl M., Duchêne M., Ebner C., Lechner K., Kraft D. and Scheiner O. (1993) Induction of specific histamine release from basophils with purified natural and recombinant birch pollen allergens. *J. Allerg. clin. Immun.* **91**, 88–97.
- van Ree R., van Leeuwen W. A., van den Berg M., Weller H. H. and Aalberse R. C. (1994) IgE and IgG cross-reactivity among Lol p I and Lol II/III. Identification of the C-termini of Lol pI, II, and III as cross-reactive structures. *Allergy* **49**, 254–261.
- Voller A., Bartlett A. and Bidwell D. E. (1978) Enzyme immunoassays with special reference to ELISA techniques. *J. clin. Pathol.* **31**, 507–520.
- von Heijne G. (1984) How signal sequences maintain cleavage specificity. *J. molec. Biol.* **173**, 243–251.
- Walsh D. J., Matthews J. A., Denmeade R. and Walker M. R. (1989) Cloning of cDNA coding for an allergen of cocksfoot grass (*Dactylis glomerata*) pollen. *Int. Arch. Allerg. appl. Immun.* **90**, 78–83.
- WHO/IUIS Allergen Nomenclature Subcommittee ACI News (1994) 6/2, pp. 38–44.
- Wüthrich B. (1989) Epidemiology of the allergic diseases: are they really on the increase? Regulatory processes in allergy and asthma. *Int. Arch. Allerg. appl. Immun.* **90**, 3–10.