

ORIGINAL PAPER

Cloning, expression and immunological characterization of full-length timothy grass pollen allergen Phl p 4, a berberine bridge enzyme-like protein with homology to celery allergen Api g 5

Å. Marknell DeWitt*, K. Andersson*, G. Peltre[†] and J. Lidholm*

*Pharmacia Diagnostics AB, Uppsala, Sweden and [†]LECA, Allergie et Environnement, Ecole Supérieure de Physique et Chimie Industrielles, Paris, France

Clinical and Experimental Allergy

Summary

Background Timothy grass pollen is a common cause of respiratory allergy in the temperate regions. The major group 4 allergen, Phl p 4, has previously been purified and studied biochemically and immunologically, but has so far not been produced and characterized as a recombinant protein.

Objective To clone and characterize timothy grass pollen allergen Phl p 4.

Methods Full-length Phl p 4 cDNA was cloned using a PCR-based strategy including 3'- and 5'-RACE. Recombinant Phl p 4 was expressed in *Escherichia coli* and purified by immobilized metal ion affinity chromatography. Its immunological activity was investigated using experimental ImmunoCAP tests, sera from Phl p 4 sensitized individuals and Phl p 4 reactive polyclonal and monoclonal animal antibodies.

Results Five full-length Phl p 4 cDNA clones were analysed. Sequence deviations between the clones were present at nine amino acid positions, and the consensus sequence comprised an open reading frame of 525 amino acids, including a predicted 25-residue signal peptide. The calculated molecular weight of the deduced mature protein was 55.6 kDa and the isoelectric point 9.9, both consistent with previously observed properties of purified nPhl p 4. Close sequence similarity was found to genomic clones from several other Pooideae grass species and to Bermuda grass pollen allergen BG60. Further, similarity was found to members of the berberine bridge enzyme (BBE) family, including celery allergen Api g 5. Recombinant Phl p 4 bound specific immunoglobulin (Ig)E from 31 of 32 nPhl p 4-reactive sera, and the IgE binding to rPhl p 4 could be inhibited by nPhl p 4 in a dose-dependent manner.

Conclusions Full-length Phl p 4 cDNA was cloned and showed sequence similarity to members of the BBE family. Recombinant Phl p 4 was produced and shared epitopes with natural Phl p 4.

Keywords allergen, allergy, Api g 5, berberine bridge enzyme, FAD binding domain, grass pollen, IgE antibody, Phl p 4, recombinant protein

Submitted 8 July 2005; revised 17 August 2005; accepted 23 September 2005

Correspondence:

Jonas Lidholm, Pharmacia Diagnostics AB, Box 6460, SE-751 37 Uppsala, Sweden.

E-mail: jonas.lidholm@diagnostics.com

Introduction

Grass pollens are a common cause of respiratory allergy in the industrialized world. Up to 20% of the general population and as much as 40% of the atopic population may be affected by grass pollen sensitization. In temperate regions of the northern hemisphere, allergenic grass species predominantly belong to the subfamily Pooideae, of which timothy grass (*Phleum pratense*) is a widely distributed and important representative [1–4].

A number of allergens contribute to the allergenicity of timothy grass pollen, as recently reviewed [4]. The most important allergens belong to groups 1 and 5, and immunoglobulin (Ig)E antibody reactivity to each of these proteins is present in 80% or more of grass pollen-sensitized individuals. Other well-characterized grass pollen allergens include profilins and EF-hand proteins, which have been referred to as pan-allergens because of their extensive cross-reactivity across diverse plant taxa [5, 6].

High molecular weight grass pollen allergens have been the subject of much fewer studies. Groups 4 and 13 allergens are both 50–60 kDa glycoproteins, of which group 13 is the most recently discovered [7]. Phl p 13 is an allergen of intermediate importance with respect to frequency of sensitization. It has been cloned and produced as a recombinant protein, which at least in part reproduced the antibody binding of natural Phl p 13. Sequence comparison revealed extensive similarity to polygalacturonase, an enzyme involved in plant cell wall biogenesis and remodelling [8].

Group 4 allergen was first reported from rye grass (*Lolium perenne*) [9] and later from a range of other grass species [10–13]. In contrast to group 13 allergens, group 4 allergens are basic proteins with an isoelectric point of 9.0–10.5. N-terminal blockage and resistance to trypsin has hampered the elucidation of the primary structure and cloning of group 4 allergens [4]. Thus, they have so far been characterized mainly by biochemical and immunological methods.

Tissue and cellular localization in the pollen grain of a group 4 allergen has been studied by immunogold electron microscopy. The binding of polyclonal anti-group 4 antibodies indicated the presence of target epitopes at both cytoplasmic and cell wall locations. In the same study, antibody binding to the cytoplasmic matrix of celery root tissue was further observed [14].

Serological studies have shown that 70–88% of grass pollen-sensitized individuals show IgE antibody binding to Phl p 4 [15, 16], although generally at a relatively low level as compared to group 1 and group 5 allergens [4]. It has been reported that 20–40% of individuals sensitized to group 4 allergens exhibit IgE antibody reactivity to carbohydrate epitopes and it is possible that glycan structures of Phl p 4 cause immunological cross-reactivity between Phl p 4 and other glycoproteins of plant origin [14, 17].

This report describes the full-length cDNA cloning and primary structure of Phl p 4, and its production as a recombinant allergen.

Materials and methods

cDNA cloning and sequence analysis of Phl p 4

Timothy grass pollen was obtained from Allergon AB (Välinge, Sweden). Total RNA was extracted using the RNAqueous extraction kit from Ambion (Austin, TX, USA), and polyadenylated RNA and cDNA were prepared using the mRNA Purification Kit and First-strand cDNA Synthesis Kit, both from Amersham Biosciences (Uppsala, Sweden). Oligonucleotide primers for cDNA synthesis, PCR and RACE experiments (Table 1) and for DNA sequencing were obtained from MedProbe (Oslo, Norway). All PCR reactions were carried out using *Pfu* and *Pfu*

Table 1. Oligonucleotides used for amplification and cloning of full-length Phl p 4

No.	Designation	Sequence
1	DG-P15/F	5'-TAC AAC TAC ATG GAG CCS TAC GT-3'
2	DG-P15/R	5'-ASG GCT CCA TGT AGT TGT AGA TGT C-3'
3	DG-P17/F	5'-GAC CCS ACC GAC TAC TTC GGC AAC-3'
4	DG-P17/R	5'-GTT GCC GAA GTA GTC GGT GGG GTC-3'
5	DG-P20/F	5'-TGG GTS GAC WSC GGC GCS CAG CT-3'
6	DG-P20/R	5'-SAG CTG SGC GCC GSW GTC SAC CCA-3'
7	DG-P22/F	5'-GTS CTS TTC AAC ATC CAG TAC GT-3'
8	DG-P22/R	5'-GTA GTT SAC GTA CTG GAT GTT GA-3'
9	PP4-3'R/GSP1	5'-CTT CAA CAT CCA GTA CGT CAA C-3'
10	PP4-3'R/GSP2	5'-CCT CTC GTG GAG CAA GGA CA-3'
11	PP4-N39	5'-CAC CGA GCC ACY GAG GAC A -3'
12	PP4-C11	5'-CGG TCC TTA TTA ATT GGC CA -3'
13	PP4/F-NdeI	5'-CGA GTC ATA TGT ACT TCC CGC CGC CTG CGG CA-3'
14	PP4/R-HindIII	5'-CCT AGA AGC TTA GTG GTG GTG GTG GTG ATG GTA CTT TTT GAT GAG CGG CGG GA-3'

Turbo DNA polymerases (Stratagene, La Jolla, CA, USA) and included 10% DMSO.

Terminal cDNA fragments were amplified by 3'-RACE according to Frohman [18] and 5'-RACE (SMART RACE cDNA Amplification Kit, Clontech, Palo Alto, CA, USA). Full-length Phl p 4 cDNA was cloned into *Sma*I digested pUC19 (New England Biolabs, Beverly, MA, USA) for DNA sequencing, and the open reading frame corresponding to the mature Phl p 4 protein was cloned into the *Nde*I/*Hind*III sites of pET-23a(+) (Novagen, Madison, WI, USA) for expression. DNA sequencing was performed using the Thermo Sequenase Radiolabeled Terminator Cycle Sequencing Kit and Redivue [α -³²P]ddNTPs Terminator Pack (Amersham Biosciences). Restriction and modifying enzymes were obtained from New England Biolabs (Beverly, MA, USA). All kits and reagents were used according to the manufacturers' recommendations.

DNA and amino acid sequence analyses and calculations were performed using programs of the GCG Wisconsin Package (Accelrys, San Diego, CA, USA). Signal peptide prediction was performed using SignalP (www.cbs.dtu.dk/services/SignalP/) [19, 20].

Production of recombinant Phl p 4 and purification of natural Phl p 4 and Dac g 4

C-terminally hexahistidine-tagged Phl p 4 was expressed in *Escherichia coli* BL21 (Invitrogen, Carlsbad, CA, USA) as previously described [21]. Cells were harvested by centrifugation, resuspended in ice-cold lysis buffer (20 mM Tris-HCl pH 8.0, 0.1 mM phenylmethanesulfonyl fluoride (PMSF), 2 µg/mL RNase A, 0.5 µg/mL DNase I) and ruptured using an Emulsiflex-C5 high-pressure

homogenizer (Avestin Inc., Ottawa, Canada). Insoluble cell material was collected by centrifugation, washed with 20 mM Tris-HCl pH 8.0, 0.1 mM PMSF, 5 mM β -mercaptoethanol (β -Me) and extracted for 1 h at room temperature with TBS (50 mM Tris-HCl pH 8.0, 0.5 M NaCl) containing 6 M guanidine hydrochloride, 5 mM β -Me and 5 mM imidazole. Following centrifugation and filtration, the extract was loaded on a HiTrap Chelating Sepharose HP column charged with 0.1 M NiSO_4 and washed with extraction buffer. After washing with TBS containing 6 M urea and 5 mM β -Me, bound protein was eluted with a 60–500 mM imidazole gradient in the same buffer. Fractions containing pure rPhl p 4 were identified by SDS-PAGE and pooled. The rPhl p 4 concentration was determined from absorbance at 280 nm, using a calculated extinction coefficient of 1.78 per mg/mL/cm. Natural Phl p 4 and Dac g 4 were purified from *P. pratense* and *Dactylis glomerata* (cocksfoot grass) pollen (Allergon AB) as previously described [17].

All chromatography equipment and separation media used were from Amersham Biosciences. SDS-PAGE was performed using precast 10% NuPAGE MES gels and the Mark 12 molecular weight markers (Invitrogen).

Analysis of immunoglobulin E antibody binding

IgE binding to rPhl p 4 was analysed using experimental ImmunoCAPTM (Pharmacia Diagnostics, Uppsala, Sweden) tests prepared as described [21] and 32 sera from grass pollen-sensitized individuals with IgE reactivity to nPhl p 4. The sera were also analysed with bromelain ImmunoCAP to assess IgE binding to cross-reactive carbohydrate determinants (CCD) [16, 22–24].

Phl p 4 epitope representation by rPhl p 4 was further studied by inhibition of IgE antibody binding by soluble nPhl p 4. Individual serum samples were pre-incubated for 3 h at room temperature with nPhl p 4 (0.001, 0.01, 0.1, 1, 10 and 100 $\mu\text{g/mL}$) or buffer alone, followed by measurement of IgE binding to rPhl p 4 ImmunoCAP tests. As a positive control, an nPhl p 4 self-inhibition experiment was performed. As a control for inhibitor specificity, a serum sample from a dog dander-sensitized individual was tested on dog-dander ImmunoCAP after pre-incubation with either buffer or the highest concentration of nPhl p 4. Pre-incubation with 100 $\mu\text{g/mL}$ of purified *E. coli* maltose binding protein (MBP) was used as a control for unspecific procedural effects. All IgE measurements were carried out in duplicate and performed according to the manufacturer's instructions, applying the recommended cut-off level of 0.35 $\text{kU}_\text{A}/\text{L}$.

Analysis of rabbit and mouse immunoglobulin G antibody binding

To study epitope similarity among nDac g 4, nPhl p 4 and rPhl p 4 and the presence of group 4-related antigenic

determinants in grass pollen and celery extracts, a polyclonal rabbit antiserum raised against nPhl p 4 [17] and four mouse monoclonal antibodies raised against nDac g 4 (mAb A, B, C and F) [11] were used. For each analysis, 40 μL of rabbit antiserum diluted 1 : 5000 or mAb solution diluted 1 : 25 was applied in an experimental ImmunoCAP rabbit IgG or mouse IgG assay, respectively, and the amount of antibody bound to each of the different immobilized proteins or extracts was calculated. Regular ImmunoCAP tests were used to assess antibody binding to extracts of *P. pratense*, *D. glomerata*, *Cynodon dactylon* (Bermuda grass), *Phragmites communis* (common reed) and *Sorghum halepense* (Johnson grass) pollen, celery and bromelain. Unspecific antibody binding was assessed using ImmunoCAP tests carrying immobilized MBP.

Results

Mapping and cDNA cloning of Phl p 4

An approach to cDNA clone Phl p 4 by RACE based on amino acid sequence information was first considered, but prevented by N-terminal blockage and resistance of the protein to trypsin digestion. To circumvent this problem, four reported peptide sequences of the group 4 allergen of *D. glomerata* [11] were instead utilized for the same purpose, assuming that at least some of them would be sufficiently homologous to Phl p 4. Based on suitable portions of these sequences, sets of degenerated forward and reverse primers (nos. 1–8, Table 1) were designed by backtranslation, applying the preferred codon usage of genes expressed in pollen of *P. pratense*.

As the linear order of the four peptides in the Dac g 4 protein was unknown, an initial experiment was performed to establish the order of the corresponding segments in Phl p 4. In a set of PCR reactions, using first stand cDNA prepared from polyadenylated *P. pratense* pollen RNA as template, 12 relevant combinations of the eight forward and reverse primers were made. Six reactions gave rise to a distinct amplification product, ranging from approximately 100 bp to 1.2 kb in size (Fig. 1a). Based on the product sizes resulting from different primer combinations, the order P20–P22–P15–P17 of the four peptide segments could be deduced, as well as their approximate spacing (Fig. 1b).

In order to obtain Phl p 4 DNA sequence from which sets of nested RACE primers could be designed, the PCR product spanning the P20–P22 segment (Fig. 1b) was cloned and analysed. Using primers 9–10 (Table 1), based on the sequence information obtained, 3'- and 5'-RACE experiments were performed, which generated terminal Phl p 4 cDNA fragments that could be cloned. DNA sequencing of these fragments enabled the design of primers 11–12 (Table 1) for amplification and cloning of full-length Phl p 4 cDNA.

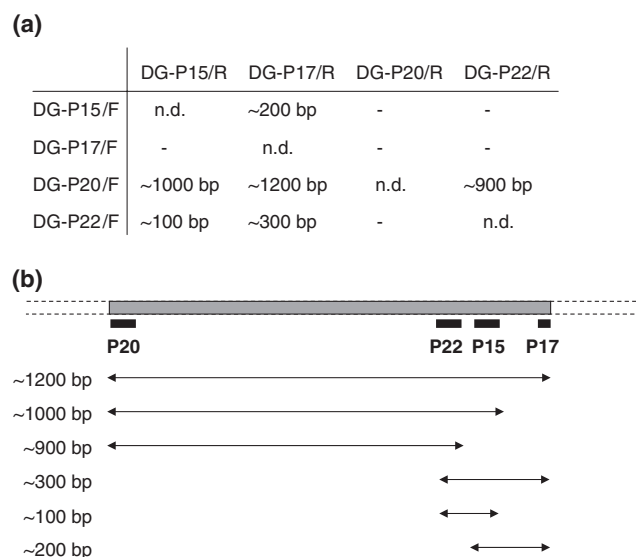


Fig. 1. Organizational mapping of Dac g 4 peptide equivalents on Phl p 4 using PCR. (a) Amplification products generated by different combinations of peptide sequence-derived primers. (b) Map showing relative positions and approximate distances of Dac g 4 peptide equivalents, consistent with the amplification results shown in (a).

Five full-length cDNA clones were isolated and sequenced. Four of the clones contained an open reading frame of 525 codons, which conformed well with the codon preference of *P. pratense* (Fig. 2). Upstream and downstream of this open reading frame, no such agreement in codon preference was present. The first 25-residue segment of the deduced Phl p 4 amino acid sequence was predicted by SignalP to be a signal peptide (Fig. 2).

Among the five clones analysed, sequence heterogeneity was present at 19 nucleotide positions. At nine of these, a change of encoded amino acid occurred, including one in the predicted signal peptide, and in one case a stop codon was formed (Table 2). One of the five clones analysed represented the dominant amino acid sequence at all variant positions and was therefore selected for predictions of protein properties and for expression of recombinant Phl p 4.

The calculated molecular weight of the predicted mature Phl p 4 protein was 55.6 kDa and its isoelectric point 9.9, both consistent with previously observed properties of purified nPhl p 4 [17]. Two consensus *N*-glycosylation sites and one putative flavin binding domain were present in the deduced sequence (Fig. 2).

A BLAST database search (www.ncbi.nlm.nih.gov/BLAST/) with the deduced amino acid sequence of Phl p 4 resulted in a 100% match over 508 residues with the translation of a genomic clone from *P. pratense* (Acc. No. AJ512488). At the nucleotide level, only three differences were found in the corresponding match. Further, a high degree of sequence identity was found to deduced amino acid sequences from genomic clones of other Pooideae

species. *L. perenne* (AJ862835): 93% over 423 aligned residues, *Secale cereale* (AJ862831): 85% over 523 residues, *Triticum aestivum* (AJ862832): 85% over 523 residues and *Hordeum vulgare* (AJ862834): 84% over 522 residues. A somewhat lower level of identity, 72% over 522 aligned residues, was found to the sequence of BG60 (AAS02108) from the non-Pooideae grass species *C. dactylon*. The database search also detected significant homology to a flavin-containing berberine bridge enzyme (BBE)-like protein from tobacco called nectarin 5 (AAP30841) [25]. In this case, the sequence identity was 48% over 418 residues.

Expression and purification of rPhl p 4

rPhl p 4 was expressed in *E. coli* as a C-terminally hexahistidine-tagged protein. The coding sequence following the segment encoding the predicted signal peptide was PCR amplified using primers 13–14 (Table 1), cloned into a pET-23a(+) vector and the construct introduced into *E. coli* strain BL21. Temperature induction of a mid-log phase culture gave rise to a prominent protein band of the expected size in SDS-PAGE (Fig. 3). Rupture and fractionation of the cells showed that the expressed protein resided entirely in the insoluble phase (not shown).

Recombinant Phl p 4 was purified by immobilized metal ion affinity chromatography (IMAC) under mildly reducing conditions in the presence of urea. The purified protein formed a single band in reducing SDS-PAGE (Fig. 3), and the final yield was approximately 10 mg/g cell pellet (fresh weight).

Analysis of IgE binding of purified rPhl p 4

The ability of rPhl p 4 to bind IgE antibodies was assessed by testing 32 sera IgE reactive to natural Phl p 4 using experimental ImmunoCAP tests carrying rPhl p 4. IgE binding by rPhl p 4 was detected in all but one of these sera (Fig. 4). Although the response levels obtained were significantly lower than to natural Phl p 4, the results indicate that recombinant rPhl p 4 at least in part reproduced the IgE binding epitopes of the natural allergen. Nineteen of 30 (63%) Phl p 4 positive sera used for the evaluation showed IgE binding to bromelain ImmunoCAP, indicating a frequent presence of CCD-reactive IgE antibodies (Fig. 4).

To further examine the sharing of epitopes between recombinant and natural Phl p 4, an IgE inhibition experiment was performed. Six sera with levels of specific IgE to rPhl p 4 ranging from 0.7 to 10 kU_A/L were selected. Each serum was pre-incubated with purified natural Phl p 4 at six concentrations between 0.001 and 100 µg/mL prior to analysis with rPhl p 4 ImmunoCAP. The assay results, shown in Fig. 5 as percentage of readout

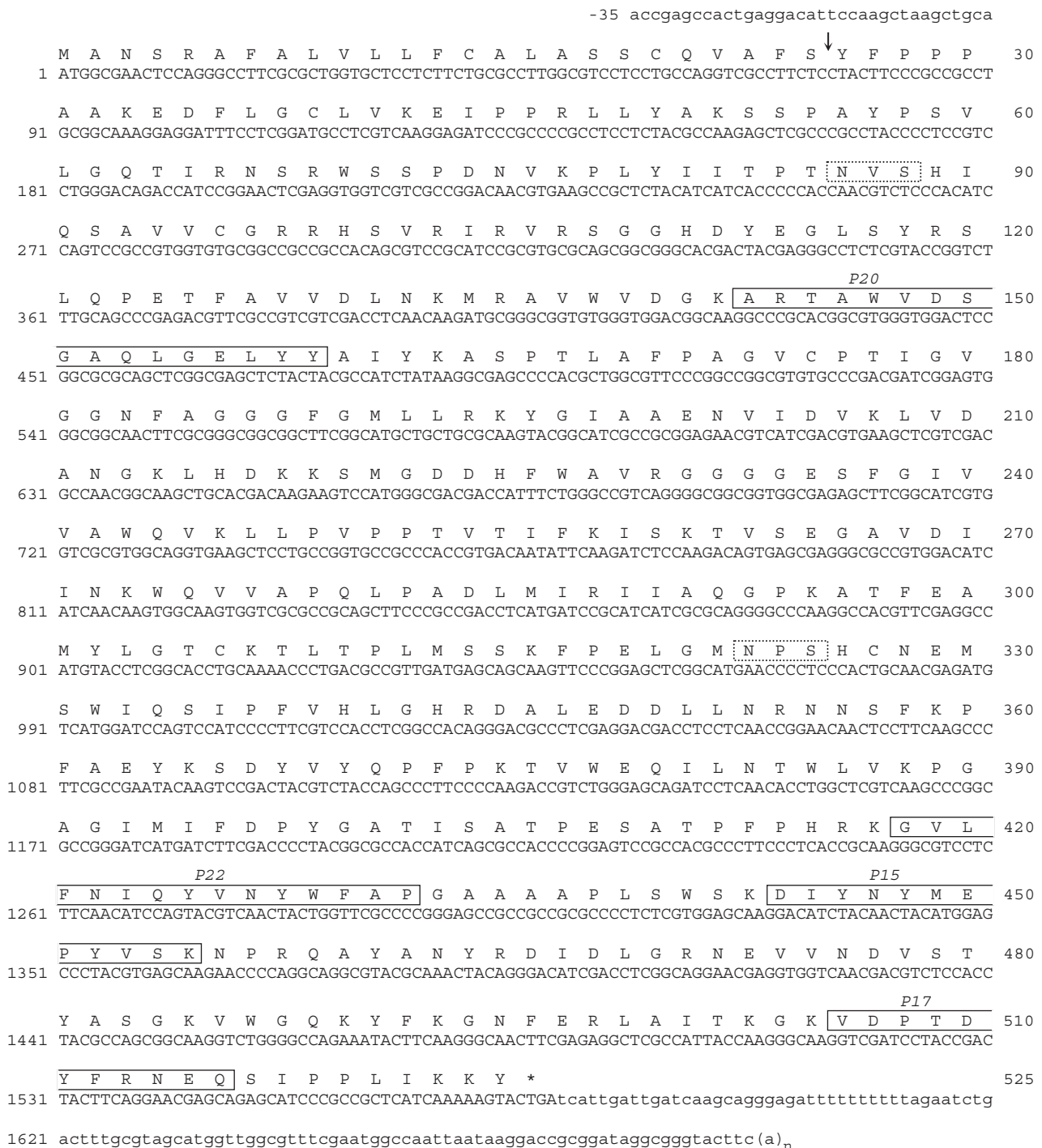


Fig. 2. DNA and deduced amino acid sequence of Phl p 4. Numbering starts at the first nucleotide of the coding sequence and the first amino acid residue. An arrowhead indicates a predicted signal peptide cleavage site. Two consensus *N*-glycosylation sites are marked by hatched boxes. Homology to a FAD-binding domain is present between residues 78 and 234 in the amino acid sequence shown. The equivalents of Dac g 4 peptides, P20, P22, P15 and P17, are boxed by solid lines. The cDNA sequence shown will be available under Accession No. DQ251188.

after pre-incubation with buffer alone, revealed a dose-dependent inhibition of IgE binding to rPhl p 4 in all sera tested. At the highest concentration of nPhl p 4 used in the experiment, 79–91% of the IgE binding to rPhl p 4 was abolished. At the same inhibitor

concentration in a self-inhibition control experiment with a subset of the sera used to study rPhl p 4, a similar degree of inhibition (83–93%) of IgE binding to immobilized nPhl p 4 by soluble nPhl p 4 occurred (data not shown).

No inhibition of IgE binding occurred after serum pre-incubation with MBP (not shown) and nPhl p 4 had no inhibitory effect on dog dander specific IgE, demonstrating that the inhibitions observed were not due to procedural or unspecific effects (Fig. 5).

Binding of polyclonal and monoclonal animal antibodies by rPhl p 4 and related natural allergens

The binding of anti-group 4 animal antibodies by recombinant and natural Phl p 4, and the presence of group 4-related epitopes among different extracts were examined using one polyclonal rabbit antiserum raised against nPhl p 4 [17] and four monoclonal mouse antibodies raised against nDac g 4 [11]. The experiment was performed using ImmunoCAP tests carrying rPhl p 4, purified

Table 2. Deduced amino acid sequence variation

Sequence position	Amino acid
24	F (4)/V (1)
34	E (4)/D (1)
101	S (3)/T (2)
211	A (4)/P (1)
252	P (4)/L (1)
423	I (4)/N (1)
459	Q (4)/Stop (1)
470	G (4)/D (1)
507	D (4)/N (1)

Number of clones out of five sequenced where the amino acid residue occurred is shown in brackets.

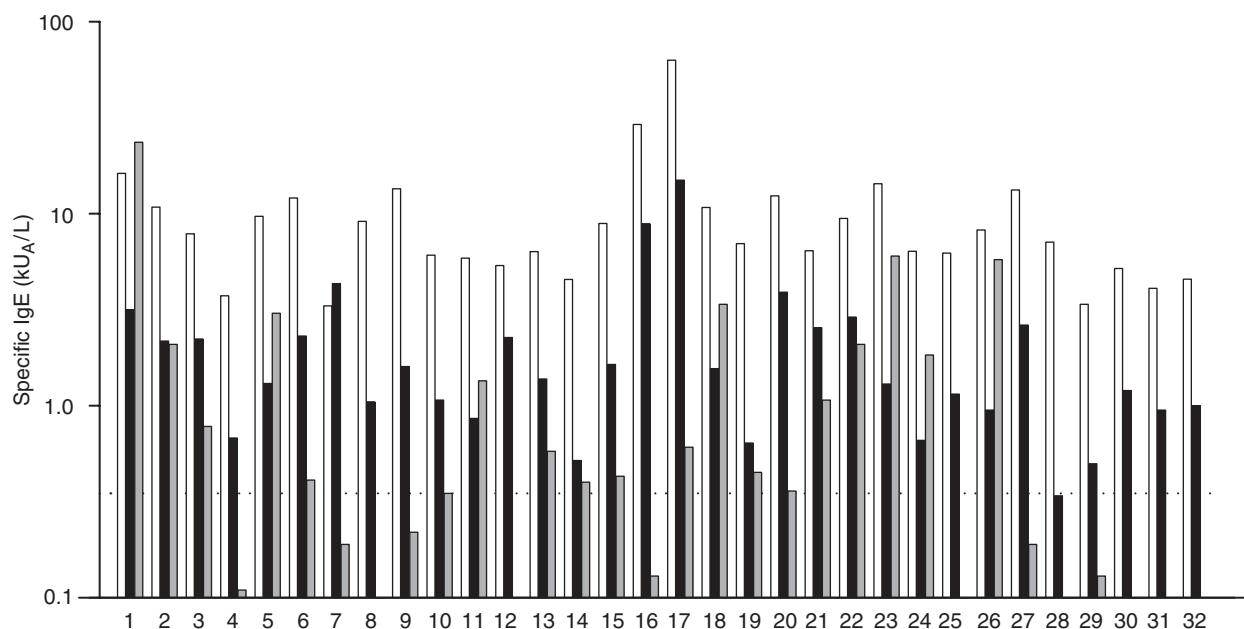


Fig. 4. Specific immunoglobulin E binding to immobilized nPhl p 4 (white bars), rPhl p 4 (black bars) and bromelain (grey bars). Serum samples of 32 grass pollen- and nPhl p 4-sensitized individuals were used for the analysis. Serum samples of subjects 31 and 32 were not tested for bromelain reactivity. Mean values of duplicate determinations are shown. The dotted line indicates a cut-off level of 0.35 kUA/L.

natural Phl p 4 and Dac g 4, pollen extracts from a range of grass species (*P. pratense*, *D. glomerata*, *C. dactylon*, *P. communis* and *S. halepense*) and celery extract. Addi-

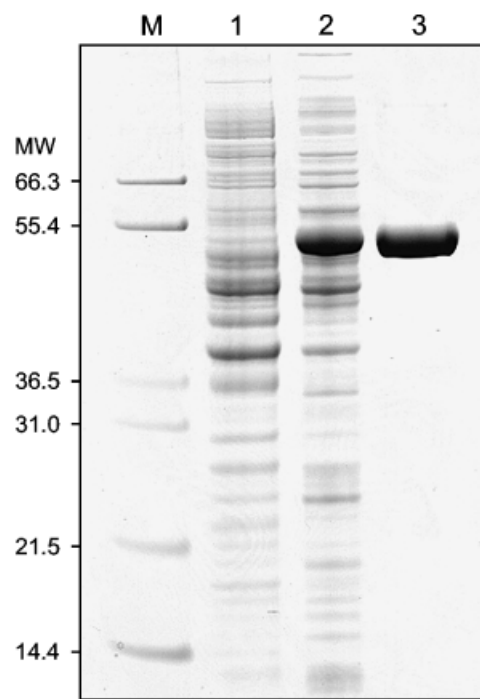


Fig. 3. Production of rPhl p 4 by expression in *Escherichia coli*. Analysis by SDS-PAGE of reduced protein samples. Lane 1: pre-induction culture sample; lane 2: culture sample at harvest; lane 3: purified rPhl p 4; M: molecular weight markers.

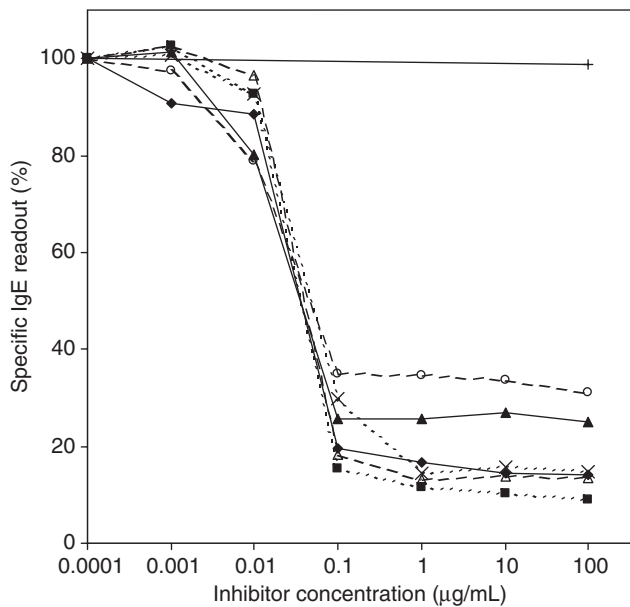


Fig. 5. Inhibition of immunoglobulin (Ig)E antibody binding to recombinant Phl p 4 by natural Phl p 4. Residual IgE binding to immobilized rPhl p 4 was measured after pre-incubation of six individual serum samples with a dilution series of purified nPhl p 4. As a negative control, a dog dander reactive serum was pre-incubated with nPhl p 4 at the highest concentration and analysed for IgE binding to dog dander extract (—○—). Results are expressed as percentage of IgE values obtained after pre-incubation with dilution buffer alone. Mean values of duplicate determinations are shown.

Table 3. Binding of polyclonal rabbit anti-Phl p 4 IgG antibodies

Solid phase	Amount of bound antibody (ng)
rPhl p 4	15.5
nPhl p 4	197
nDac g 4	110
<i>Phleum pratense</i> *	19.6
<i>Dactylis glomerata</i> *	27.1
<i>Cynodon dactylon</i> *	7.8
<i>Phragmites communis</i> *	8.1
<i>Sorghum halepense</i> *	11.8
Celery*	8.3
Bromelain	8.6
MBP†	0.19

*Natural allergen extract. †MBP was included as a negative control. IgG, immunoglobulin G; MBP, maltose binding protein.

tionally, all antibodies were tested with bromelain ImmunoCAP as an indicator of CCD binding.

IgG of the polyclonal anti-Phl p 4 antiserum bound to all grass pollen allergen-derived tests used in the experiment, with signals at least 40 times higher than to MBP (Table 3). The strongest binding occurred to purified natural Phl p 4 and Dac g 4. Among the grass pollen extracts tested, stronger binding to species belonging to

Table 4. Binding of mouse monoclonal anti-Dac g 4 IgG antibodies

Solid phase	Amount of bound antibody (ng)			
	mAb A	mAb B	mAb C	mAb F
rPhl p 4	142	77	153	220
nPhl p 4	114	48	81	127
nDac g 4	334	57	361	488
<i>Phleum pratense</i> *	84	29	33	55
<i>Dactylis glomerata</i> *	95	32	54	86
<i>Cynodon dactylon</i> *	60	23	31	50
<i>Phragmites communis</i> *	108	40	67	113
<i>Sorghum halepense</i> *	65	30	60	106
Celery*	43	27	20	46
Bromelain	64	25	61	92

*Natural allergen extract. IgG, immunoglobulin G.

the Pooideae subfamily than to the more distantly related *C. dactylon* and *P. communis* was observed. rPhl p 4 bound antibodies with an intensity intermediate among the grass pollen extracts. Rabbit anti-Phl p 4 IgG also bound celery extract; however, the antibody binding to both celery and non-Pooideae grass pollen extracts was comparable to that of bromelain. The results suggest that co-recognition of group 4 allergens from different species by the rabbit IgG antibodies is relatively limited, that antibody binding to natural grass pollen extracts was restricted by antigen availability and that most of the rabbit hyperimmune response was directed to determinants not present in rPhl p 4, potentially including glycan determinants.

The results with the anti-Dac g 4 monoclonal antibodies were partly different from those obtained with the rabbit antiserum (Table 4). Three of the four mouse anti-Dac g 4 mAbs showed much higher binding to nDac g 4 than to nPhl p 4 and, surprisingly, all showed somewhat higher binding to recombinant than to natural Phl p 4. Antibody B was particularly reactive to rPhl p 4, showing a binding to this protein that was even stronger than to nDac g 4. Further, the monoclonal antibodies did not discriminate between Pooideae and non-Pooideae grass species, and they all showed significant binding to bromelain and a somewhat weaker binding to celery extract.

Discussion

In this paper we report the full-length cDNA cloning of Phl p 4, the group 4 grass pollen allergen of *P. pratense*, and its production as a recombinant allergen. Due to its particular properties, including blockage of its N-terminal and resistance to trypsin digestion, relatively little progress has previously been made in the characterization of Phl p 4. The protein has been purified and rabbit antibodies raised against it, but amino acid sequencing or cDNA cloning of Phl p 4 has so far not been reported.

In the cloning of Phl p 4 described here, advantage was taken of the reported peptide sequences from the group 4 allergen of *D. glomerata*, a grass species closely related to *P. pratense*. Using oligonucleotide primers derived from these amino acid sequences, it was possible by PCR to establish the linear order of the segments in Phl p 4, corresponding to the four described Dac g 4 peptides. Through initial cloning of an internal cDNA fragment and subsequent RACE experiments, the entire Phl p 4 cDNA could be cloned.

Sequencing of the Phl p 4 cDNA clones revealed a 1575 bp open reading frame, conforming with the preferred codon usage of other genes expressed in *P. pratense* pollen. Of the 525 amino acid residues encoded by the open reading frame, the segment encompassing the first 25 residues was predicted to form a signal peptide, leaving a predicted mature protein of 500 residues.

Nucleotide differences between the five clones analysed were found at 19 positions. Even though amplification artefacts cannot be excluded, we find it likely that many of the sequence variations represent true polymorphisms. Contributing to this notion are the facts that a high fidelity DNA polymerase was used throughout the experimental procedures and that the nucleotide variation observed was not randomly distributed but biased towards third (silent) codon positions. One clone, which encoded the consensus amino acid sequence at all variant positions, was selected for predictions of protein properties and recombinant protein production.

The predicted properties of the deduced Phl p 4 polypeptide, excluding the predicted signal peptide, were in good agreement with previous analyses of purified natural Phl p 4. The protein has been shown to migrate, as a band of 55–60 kDa in reducing SDS-PAGE and MALDI-TOF measurements indicated an average mass of 61.3 kDa [17]. Considering the presence of glycans in natural Phl p 4 [17], the molecular mass of 55.6 kDa calculated from the sequence reported here is well within the range that would be expected. Secondly, Phl p 4 is known to be a basic protein, which is consistent with the calculated isoelectric point of 9.9. Finally and strongly compelling, the amino acid composition deduced from the sequence is in 97% agreement with the experimentally determined composition of purified natural Phl p 4 [17]. Based on these considerations, we conclude that the protein cloned in this work is the same as the Phl p 4 allergen previously purified from timothy grass pollen.

Group 4 allergens, including Phl p 4, have previously been shown to be glycosylated [9, 17, 26] and, based on the difference between the reported mass determination of nPhl p 4 and the mass calculated from the sequence reported here, it appears that approximately 6 kDa of the natural protein may be made up of carbohydrate, i.e. ~10% of its total mass. Two potential *N*-glycosylation sites were found in the Phl p 4 sequence, but it is presently

not known whether one or both of these sites carry glycan structures.

The Phl p 4 sequence contained a putative flavin binding domain and a database search revealed homology to FAD-linked oxidoreductase proteins belonging to the BBE family. BBE are enzymes that take part in the biosynthesis of sanguinarine and are believed to be involved in the plant pathogen response system, specifically against fungal infections [27]. BBE have mainly been studied in Papaveraceae species, e.g. *Eschscholzia californica*, where they occur as membrane-associated enzymes [28] predominantly present in the root, stem and leaf tissues [29]. Drawing from primary structure similarity, it is possible that Phl p 4 has a biological function analogous to that of BBE and it would then be yet another example of plant allergens involved in the pathogen response cascade [30].

The first flavin-containing allergen reported was BG60 from *C. dactylon* [31], a non-Pooideae species to which Phl p 4 showed extensive amino acid sequence identity. Interestingly, homology was also found to a recently described high molecular weight allergen from celery (*Apium graveolens*), Api g 5 [32]. Amino acid sequences from four fragments of Api g 5 (P81943) displayed 25–80% identity to Phl p 4 over stretches of 5–30 aligning residues. This glycoprotein has also been characterized as a FAD-containing oxidoreductase [33], and it has been proposed that Api g 5 only binds specific IgE by its *N*-glycan structures [33]. Previous studies have indicated a serological relationship between Phl p 4 and celery [14, 17], but evidence regarding the clinical significance of these observations is so far lacking.

A high molecular weight allergen with similarity to BBE has also been described in *Brassica napus* pollen [34]. However, no convincing homology was found between Phl p 4 and peptide sequences derived from this protein.

Phl p 4 was expressed as a hexahistidine-tagged protein at a high level in *E. coli*, purified in a single chromatographic step and tested in different ways for antibody binding. Even though the protein required the presence of urea to remain soluble, suggesting inappropriate folding, it was found to share IgE binding epitopes with natural Phl p 4 and to bind Phl p 4-reactive animal antibodies, albeit mostly with a lower intensity than the natural protein. Four alternative or complementary explanations for these results can be conceived: (i) it is possible that linear or folding-independent epitopes were present, which bound antibodies even in an unfolded or misfolded state of the protein; (ii) even if the protein was globally unfolded or misfolded, patches of correct folding may have developed in the non-denaturing environment following attachment to the ImmunoCAP solid phase and during serum contact, causing the formation of relevant epitope structures; (iii) it is possible that a small subset of the protein molecules developed an authentic folding,

causing proper epitope formation to occur, but at a low relative abundance; (iv) some of the difference in antibody binding between the natural and recombinant protein may be ascribed to the recognition of glycan structures by some antibodies, which are present in the natural allergen but lacking in the *E. coli*-expressed recombinant protein [16]. Indeed, it has previously been shown that purified nPhl p 4 loses a great proportion (20–86%) of its IgE binding activity upon glycan oxidation by periodate treatment [17], and similar results have been obtained for the group 4 allergen of *C. dactylon* [26]. Even though the validity and relative contribution of these potential mechanisms cannot be determined from the present data, an important aspect of the results is that the recombinant Phl p 4 allergen sufficed to detect specific IgE in all but one of the 32 Phl p 4 reactive sera tested.

In conclusion, we report the full-length cDNA cloning, primary structure and recombinant production of Phl p 4, a major allergen in timothy grass, previously studied only as a purified natural protein. Even though further work will be required to refine the method of production of an authentically folded and fully soluble Phl p 4 protein, recombinant Phl p 4 may become useful as a diagnostic reagent and as a model to study the role of allergen glycosylation.

Acknowledgements

We thank Ingegerd Cressy (Uppsala) for advice regarding the ImmunoCAP mouse IgG assay.

References

- Evans III R. Epidemiology and natural history of asthma, allergic rhinitis, and atopic dermatitis. In: Middleton Jr E, Reed CE, Ellis EF, Franklin Adkinson Jr N, eds. *Allergy: principles and practice*. St Louis: Mosby Year Book, 1993; 1109–36.
- Burney P, Malmberg E, Chinn S, Jarvis D, Luczynska C, Lai E. The distribution of total and specific serum ige in the European Community Respiratory Health Survey. *J Allergy Clin Immunol* 1997; 99:314–22.
- D'Amato G, Spiekma FT, Liccardi G *et al.* Pollen-related allergy in Europe. *Allergy* 1998; 53:567–78.
- Andersson K, Lidholm J. Characteristics and immunobiology of grass pollen allergens. *Int Arch Allergy Immunol* 2003; 130:87–107.
- Valenta R, Duchêne M, Ebner C *et al.* Profilins constitute a novel family of functional plant pan-allergens. *J Exp Med* 1992; 175:377–85.
- Valenta R, Hayek B, Seiberler S *et al.* Calcium-binding allergens: from plants to man. *Int Arch Allergy Immunol* 1998; 117:160–6.
- Suck R, Hagen S, Cromwell O, Fiebig H. The high molecular mass allergen fraction of timothy grass pollen (*Phleum pratense*) between 50–60 kDa is comprised of two major allergens: Phl p 4 and Phl p 13. *Clin Exp Allergy* 2000; 30:1395–402.
- Suck R, Petersen A, Hagen S, Cromwell O, Becker W, Fiebig H. Complementary DNA cloning and expression of a newly recognized high molecular mass allergen Phl p 13 from timothy grass pollen (*Phleum pratense*). *Clin Exp Allergy* 2000; 30:324–32.
- Ekrasmoddollah AK, Kisil FT, Schon AH. Immunochemical characterization of a high molecular weight basic allergen (HMBA) of rye grass (*Lolium perenne*) pollen. *Mol Immunol* 1983; 20:465–73.
- Haavik S, Paulsen BS, Wold JK. Glycoprotein allergens in pollen of timothy. II. Isolation and characterization of a basic glycoprotein allergen. *Int Arch Allergy Appl Immunol* 1985; 78:260–8.
- Leduc-Brodard V, Inacio F, Jaquinod M, Forest E, David B, Peltre G. Characterization of Dac g 4, a major basic allergen from *Dactylis glomerata* pollen. *J Allergy Clin Immunol* 1996; 98:1065–72.
- Gavrovic-Jankulovic M, Cirkovic T, Bukilica M, Fahlbusch B, Petrovic S, Jankov RM. Isolation and partial characterization of Fes p 4 allergen. *J Invest Allergol Clin Immunol* 2000; 10:361–7.
- Su SN, Lau GX, Tsai JJ, Yang SY, Shen HD, Han SH. Isolation and partial characterization of Bermuda grass pollen allergen, BG-60a. *Clin Exp Allergy* 1991; 21:449–55.
- Grote M, Stumvoll S, Reichelt R, Lidholm J, Valenta R. Identification of an allergen related to Phl p 4, a major timothy grass pollen allergen, in pollens, vegetables, and fruits by immunogold electron microscopy. *Biol Chem* 2002; 383:1441–5.
- Rossi RE, Monasterolo G, Monasterolo S. Measurement of IgE antibodies against purified grass-pollen allergens (Phl p 1, 2, 3, 4, 5, 6, 7, 11, and 12) in sera of patients allergic to grass pollen. *Allergy* 2001; 56:1180–5.
- Mari A. Skin test with a timothy grass (*Phleum pratense*) pollen extract vs. IgE to a timothy extract vs. IgE to rPhl p 1, rPhl p 2, nPhl p 4, rPhl p 5, rPhl p 6, rPhl p 7, rPhl p 11, and rPhl p 12: epidemiological and diagnostic data. *Clin Exp Allergy* 2003; 33:43–51.
- Stumvoll S, Lidholm J, Thunberg R *et al.* Purification, structural and immunological characterization of a timothy grass (*Phleum pratense*) pollen allergen, Phl p 4, with cross-reactive potential. *Biol Chem* 2002; 383:1383–96.
- Frohman M. Rapid amplification of complementary DNA ends for generation of full-length complementary DNAs: thermal RACE. *Methods Enzymol* 1993; 218:340–56.
- Nielsen H, Engelbrecht J, Brunak S, von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* 1997; 10:1–6.
- Nielsen H, Krogh A. Prediction of signal peptides and signal anchors by a hidden Markov model. *Proc Int Conf Intell Syst Mol Biol* 1998; 6:122–30.
- Marknell DeWitt Å, Niederberger V, Lehtonen P *et al.* Molecular and immunological characterization of a novel timothy grass (*Phleum pratense*) pollen allergen, Phl p 11. *Clin Exp Allergy* 2002; 32:1329–40.
- Wilson IBH, Harthill JE, Mullin NP, Ashford DA, Altmann F. Core α -1,3-fucose is a key part of the epitope recognized by antibodies reacting against plant N-linked oligosaccharides and is present in a wide variety of plant extracts. *Glycobiology* 1998; 8:651–6.
- Fötisch K, Altmann F, Haustein D, Vieths S. Involvement of carbohydrate epitopes in the IgE response of celery-allergic patients. *Int Arch Allergy Immunol* 1999; 120:30–42.
- Mari A. IgE to cross-reactive carbohydrate determinants: analysis of the distribution and appraisal of the *in vivo* and *in vitro* reactivity. *Int Arch Allergy Immunol* 2002; 129:286–95.

- 25 Carter C, Thornburg R. Tobacco nectarin V is a flavin-containing berberine bridge enzyme-like protein with glucose oxidase activity. *Plant Physiol* 2004; **134**:460–9.
- 26 Su SN, Shu P, Lau GX, Yang SY, Huang SW, Lee YC. Immunologic and physicochemical studies of Bermuda grass pollen antigen BG60. *J Allergy Clin Immunol* 1996; **98**:486–94.
- 27 Dittrich H, Kutchan TM. Molecular cloning, expression, and induction of berberine bridge enzyme, an enzyme essential to the formation of benzophenanthridine alkaloids in the response of plants to pathogenic attack. *Proc Natl Acad Sci USA* 1991; **88**:9969–73.
- 28 Hauschild K, Pauli HH, Kutchan TM. Isolation and analysis of a gene bbe1 encoding the berberine bridge enzyme from the California poppy *Eschscholzia californica*. *Plant Mol Biol* 1998; **36**:473–8.
- 29 Huang FC, Kutchan TM. Distribution of morphinan and benzo[c]-phenanthridine alkaloid gene transcript accumulation in *Papaver somniferum*. *Phytochemistry* 2000; **53**:555–64.
- 30 Hoffmann-Sommergruber K. Pathogenesis-related (PR)-proteins identified as allergens. *Biochem Soc Trans* 2002; **30**:930–5.
- 31 Liaw SH, Lee DY, Chow LP, Lau GX, Su SN. Structural characterization of the 60-kDa Bermuda grass pollen isoallergens, a covalent flavoprotein. *Biochem Biophys Res Commun* 2001; **280**:738–43.
- 32 Ganglberger E, Radauer C, Grimm R *et al*. N-terminal sequences of high molecular weight allergens from celery tuber. *Clin Exp Allergy* 2000; **30**:566–70.
- 33 Bublin M, Radauer C, Wilson IB *et al*. Cross-reactive N-glycans of Api g 5, a high molecular weight glycoprotein allergen from celery, are required for immunoglobulin E binding and activation of effector cells from allergic patients. *FASEB J* 2003; **17**:1697–9.
- 34 Chardin H, Mayer C, Senechal H, Tepfer M, Desvaux FX, Peltre G. Characterization of high-molecular-mass allergens in oilseed rape pollen. *Int Arch Allergy Immunol* 2001; **125**:128–34.