

A. Petersen
A. Buße
G. Schramm
M. Schlaak
W.-M. Becker

Division of Allergology, Forschungsinstitut
Borstel, Germany

Characterization of the Allergen Group VI in Timothy Grass Pollen (Phl p 6)

II. cDNA Cloning of Phl p 6 and Structural Comparison to Grass Group V

Key Words

cDNA cloning
Epitopes
Expression
Grass pollen allergen
IgE binding
Immunoblotting

Abstract

The complete primary structure of the major allergen Phl p 6 was determined by cDNA cloning. A cDNA library of timothy grass pollen was screened by a Phl p 6 directed mouse antiserum. Six clones were obtained. From the cDNA data we deduced a protein sequence of 110 amino acids. While the size of the leader sequences varied considerably between the different clones, only one amino acid exchange affected the mature proteins. Sequence comparison of group VI and V allergens showed a high degree of homology in the N- and C-terminal regions. The structural data strongly indicate that group VI and V are independent groups of allergens, which have derived from a common precursor gene.

Introduction

The similarity of the two major grass groups V and VI was first described by Løwenstein [1], when he isolated allergens from crude timothy grass pollen extract by ion exchange chromatography and gel filtration. Matthiesen and Løwenstein [2] noted the difficulties encountered in performing copper chelate and affinity chromatography. We observed similar characteristics in staining proteins [3]: india ink detects neither group V nor group VI components; subsequent staining by AuroDye identifies components of both groups. Recent data [4] have shown that there is a cross-reactivity between the major allergen grass groups V and VI of timothy (*Phleum pratense*). Although the N-terminal sequence of Phl p 6 showed a remarkable sequence homology to an internal part of group V allergens [5], the

connection between the two allergen groups remains uncertain. Perhaps Phl p 6 is a fragment derived by proteolysis from Phl p 5 allergens or is an independent group of allergens that has arisen by gene duplication.

Since cDNA technology was introduced in allergy research [6], the determination of complete amino acid sequences from cDNA data has become possible. In 1993, Buße et al. [7] and Buße [8] successfully elucidated the amino acid sequence of the two group V allergens in timothy grass pollen, Phl p 5a and 5b, after immunoscreening of a cDNA library using monoclonal antibodies.

In order to investigate the major allergen Phl p 6, we determined the primary structure by cDNA cloning and sequencing. Comparing sequences of Phl p 6, 5a and 5b reveals a high degree of homology which may be responsible for the similarity of their immunological reactivity.

Received:
June 22, 1994
Accepted after revision:
April 4, 1995

Correspondence to: Dr. Arnd Petersen
Division of Allergology, Forschungsinstitut Borstel
Parkallee 22
D-23845 Borstel (Germany)

© 1995 S. Karger AG, Basel
1018–2438/95/1081–0055
\$8.00/0

Materials and Methods

Patients' Pooled Serum, Monoclonal Antibodies and Polyclonal Phl p 6 Antiserum. For the identification of IgE-reactive proteins, we used a mixture of sera from 3 grass pollen sensitive patients (Pharmacia CAP System, CAP score VI). The monoclonal antibodies Bo1 [9] and IG12 [10] were raised against crude timothy grass pollen extract and are directed against allergens of grass group V and I, respectively. The hyperimmune serum against grass group VI was raised in BALB/c mice by immunization with allergen bearing nitrocellulose particles from 2D blotting [4].

Western Blotting. Proteins were separated by discontinuous SDS-PAGE using the buffer system (T=15%, C=4%) according to Laemmli [11], transferred onto nitrocellulose membrane by semi-dry blotting [12], and immunostained according to Leary et al. [13].

cDNA Cloning and Expression of the Fusion Proteins. A cDNA library of timothy grass pollen [7], which was prepared according to the method of Gubler and Hoffmann [14] (cDNA Synthesis Kit, Pharmacia, Uppsala, Sweden), was screened for grass group VI specific clones. After induction by 2 mM IPTG (isopropyl β -D-thiogalactopyranoside) an immunoscreening was performed by a mouse antiserum raised against Phl p 6 [4] and patients' pooled serum. Positive clones were rescreened, isolated and amplified.

The coinfection with the helper phage Exassist (Stratagene, La Jolla, Calif., USA) resulted in the in vivo excision of pBluescript SK-plasmid. dsDNA was sequenced by the Sanger method [15] (T7 Sequencing Kit, Pharmacia) using the following sequence specific primer (ACACTGCCAGTGGCTAA) at the 5' end of the cDNA. The primer was synthesized with the Cyclotron Plus Synthesizer (Millipore, Bedford, Tex., USA).

For expression of the fusion proteins transfected *Escherichia coli* K12 SOLR cells (Stratagene, La Jolla, Calif., USA) were induced by 2 mM IPTG. Aliquots were taken after 0 and 2 h of incubation, resuspended in SDS sample buffer, heated for 5 min at 95°C and subsequently analyzed by Western blotting.

Computer Analyses. Sequence data were analyzed on GeneWorks and PCGENE software (Intelligenetics, Geel, Belgium) and on DNASIS and PROSIS Software (Pharmacia LKB Biotechnology, Uppsala, Sweden). The antigenic binding sites were predicted according to Jameson and Wolf [16] by determination of the antigenic index.

Nucleotide Sequence Accession Number. The nucleotide sequence of clone pPHLPVI.D reported in this paper has been submitted to the Gene Bank/EMBL Data Bank with the accession number Z27090.

Results

cDNA Sequence Analysis

A cDNA library of timothy grass pollen was screened using a Phl p 6 directed mouse antiserum. Six clones were obtained (pPHLPVI.C, D, F, G, H, L). All bound patients' IgE. We identified inserts of 550–700 bp.

Figure 1 shows the nucleotide sequence and the deduced protein sequence of clone pPHLPVI.F. The N-terminal region which has already been determined by protein sequencing is underlined. The 5' upstream region consists of 104 nucleotides. Three ATG triplets in this leader sequence

are possible initiation sites for translation. The deduced amino acid sequence of the mature protein consists of 110 amino acid residues. No cysteine residues were detected.

The six cDNA clones of Phl p 6 were screened for differences in their nucleotide sequences. Size variations are a result of differing lengths of non-coding 3'-regions. Nucleotide substitution at position 388 gave rise to an amino acid change from histidine (in pPHLPVI.F, H, L) to arginine (in pPHLPVI.C, D, G). Other substitutions at positions 137, 188, 362, 389 and 413 had no effect at the protein level. The cDNA segments 5' of the sequence coding for the mature proteins differ in length between the six clones: pPHLPVI.F has 104 nucleotides, both pPHLPVI.C and G have 80 nucleotides, pPHLPVI.D 77 nucleotides, pPHLPVI.H and L 14 nucleotides. A nucleotide substitution at position 77 of the pPHLPVI.F clone (fig. 1) causes clones pPHLPVI.C, F and G to have a cytosine residue at this point and clone pPHLPVI.D to have an adenine residue. This region is missing from clones pPHLPVI.H and L.

Figure 2 depicts the reaction of patients' IgE and the Phl p 6 directed antiserum to timothy pollen extract and to cell lysates of pPHLPVI.H transfected SOLR (SOLR/pVI.H). Expression analysis samples of the bacterial cultures were taken 2 h after induction by IPTG. Patients' serum and Phl p 6 antiserum demonstrated two proteins, an intensely staining 17-kD protein band and a weaker 14-kD band. Additional fainter bands were observed at 35 and 47 kD. These bands were also observed in cultures without Phl p 6 expression (cell lysate 0 h) which would seem to indicate a cross-reactivity of the sera to the bacterial proteins. For comparison we performed Western blotting of timothy pollen extract. Using Phl p 6 antiserum and patients' IgE we were able to demonstrate that a 13-kD component was Phl p 6. The monoclonal antibodies Bo1 and IG12 were used as controls. They only bound to the separated timothy pollen extract identifying group V (38 and 32 kD) and group I (37 and 35 kD) allergens, respectively.

Comparison of the Primary Structures Phl p 6, 5a and 5b

Structural similarity between group V and VI allergens of timothy grass pollen was studied by comparing their protein sequences deduced from cDNA data.

Figure 3 shows that parts of Phl p 6, especially the N- and C-termini exhibit high sequence homology to internal parts of Phl p 5a and Phl p 5b. Phl p 6 has a 54% (61%) match of sequence identity (homology) to Phl p 5a and a 51% (56%) match to Phl p 5b. The average sequence homology of Phl p 6 to both subgroups was determined as 42%. The Phl p 5 subgroups differ in a few amino acid residues. Phl p 6 has

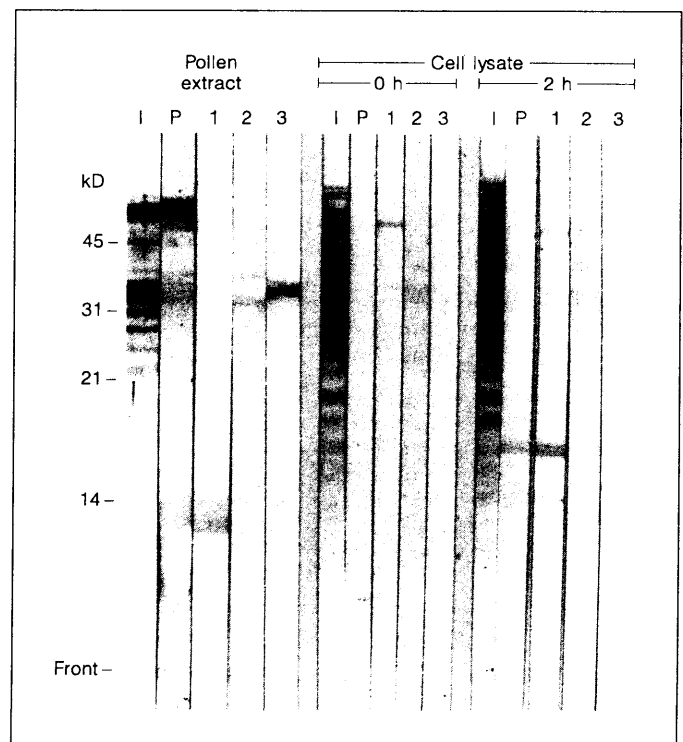
1	AA	ACC	AAC	GCA	CGA	GTA	GCA	ATG	GCA	GCG	CAC	AAG	TTC	ATG	GTG	GCG
1								M	A	A	H	K	F	M	V	A
48	ATG	TTC	CTC	GCT	GTT	GCC	GTT	GTG	TTG	GGC	TTG	GCC	ACA	TCC	CCA	ACT
10	M	F	L	A	V	A	V	V	L	G	L	A	T	S	P	T
96	GCA	GAG	GGA	GGG	AAG	GCC	ACG	ACC	GAG	GAG	CAA	AAA	TTG	ATC	GAG	GAC
26	A	E	G	G	K	A	T	T	E	E	Q	K	L	I	E	D
144	GTC	AAT	GCC	AGC	TTT	AGG	GCG	GCC	ATG	GCC	ACC	ACT	GCT	AAC	GTG	CCT
42	V	N	A	S	F	R	A	A	M	A	T	T	A	N	V	P
192	CCA	GCA	GAC	AAG	TAT	AAG	ACA	TTC	GAA	GCC	GCC	TTC	ACG	GTG	TCC	TCA
58	P	A	D	K	Y	K	T	F	E	A	A	F	T	V	S	S
240	AAG	AGA	AAC	CTC	GCT	GAC	GCC	GTT	TCA	AAG	GCG	CCC	CAG	CTG	GTG	CCC
74	K	R	N	L	A	D	A	V	S	K	A	P	Q	L	V	P
288	AAG	CTC	GAT	GAA	GTC	TAC	AAC	GCC	GCC	TAC	AAT	GCT	GCC	GAT	CAT	GCC
90	K	L	D	E	V	Y	N	A	A	Y	N	A	A	D	H	A
336	GCC	CCA	GAA	GAC	AAG	TAT	GAA	GCC	TTC	GTC	CTT	CAC	TTT	TCC	GAG	GCT
106	A	P	E	D	K	Y	E	A	F	V	L	H	F	S	E	A
384	CTC	CAC	ATC	ATC	GCC	GGT	ACC	CCC	GAG	GTG	CAC	GCT	GTG	AAG	CCC	GGC
122	L	H	I	I	A	G	T	P	E	V	H	A	V	K	P	G
432	GCG	TAG	TTG	TTC	AGC	ACG	GTC	AAG	ATC	CTT	GAC	AGC	GTG	GCT	GCC	ACC
138	A	*														
480	GGC	GCT	GCA	GCC	AAC	ACT	GCC	AGT	GGC	TAA	AAA	ATT	CGA	CTA	GCT	CCT
528	TCA	TAC	AAT	GAA	TAC	ACA	TGT	ATC	ATT	CAA	ACA	TAC	TAC	TGT	ACA	GTA
576	TGT	GCA	TGA	CCT	AGC	GGC	GAG	CAT	TTT	TTT	TAT	GAT	TAA	TCT	TTT	ATA
624	CAT	GGG	CGT	GAT	CGA	GCG	TGT	GCA	TAT	GTG	TAA	TAA	TTA	ATT	TTT	TAT
672	TTT	GCG	CCC	GGC	ACA	T										

Fig. 1. cDNA and deduced amino acid sequence of the clone pPHLPVI.F. Differences among the six pPHLPVI clones are highlighted.

identical amino acids to Phl p 5a at 15 positions (16, 28, 30, 34, 41, 45, 50, 54, 57, 83, 85, 91, 96, 100, 110) and to Phl p 5b at 10 positions (13, 35, 38, 40, 47, 49, 60, 61, 73, 86).

Predictions for antigenic binding sites on the group VI allergen were performed by the Jameson and Wolf algorithm. Predicted binding sites were correlated to sequence identities to the group V allergens. Antigenic indexes were limited to >1.2 (indicated in fig. 3). Two putative binding sites were deduced between amino acids 30–34 and 46–47. Figure 3 demonstrates the first binding site to be identical to the Phl p 5a sequences and differs by two amino acid residues to Phl p 5b. The second binding site differs between the three proteins except for a common arginine residue at position 47 shared by Phl p 6 and Phl p 5b.

Fig. 2. Expression of clones pPHLPVI.H in *E. coli* SOLR after IPTG treatment. Western blotting of timothy grass pollen extract (e) and SOLR/PVI.H cell lysates after induction for 0 and 2 h. I = Protein staining by india ink; P = patients' pooled serum; 1 = mouse Phl p 6 antiserum; 2 = monoclonal antibody Bo1 directed against group V; 3 = monoclonal antibody IG 12 directed against group I.



PHLPVI	-----GKATTEEQKLIEDVNASFR	19
PHLPVA	ADLGYPATPAAPAAGYTPATPAAPAGADAAGKATTEEQKLIKINAGFK	50
PHLPVB	ADAGYAPATPAAAGA-----AAGKATTEEQKLIEDINVGFK	36
PHLPVI	AAMATTANVPPADK ^{***} YKTFEAAFTVSS ^{***} KRN ^{***} LADAVSKAPQ-----L	59
PHLPVA	AALAGAG-VQPADKYRTFVATFGPASNKAF ^{***} AEGLSGEPKGAESSSKAAL	99
PHLPVB	AAVAARQ-RPAADK ^{***} YKTFEAA-SFRHPRPLRQ-----AGL	70
PHLPVI	VPKLEEVNAAYNADHAAPEDKYEAFVLHFSEALRIIAGTPEVHAVKPG	109
PHLPVA	TSKLDAA ^{***} YKLAYKTAEGATPEAKYDAYVATLSEALRIIAGTLEVHAVKPA	149
PHLPVB	VEKLDAA ^{***} YSVAYKA ^{***} AVGATPEAKFDSFVASLTEALRVITAGALEVHAVKPV	120
PHLPVI	A-----	110
PHLPVA	AEEVKV--IPAGELQVIEKVDAAFKVAATAANAAPANDKFTVFEAAFNDE	197
PHLPVB	TEEPGMAKIPAGELQIIDKIDA ^{***} AFKVAATAAATAPADDKFTVFEAAFNKA	170
PHLPVI	-----	110
PHLPVA	IKASTGGAYESYKFIPALEAAVKQAYAAATVATAPEVKYTVFETALKKAIT	247
PHLPVB	IKESTGGAYDTYKCIPSLAAVKQAYAAATVAAAAEVKYAVFEAA ^{***} LTKAIT	220
PHLPVI	-----	110
PHLPVA	AMSEAQKAAKPPPLPP-----PPQPPPLAATGAATAATGGYKV	285
PHLPVB	AMSEVQKVSQPATGAATVAAGAATTAAGAASGAATVAAGGYKV	263

Fig. 3. Comparison of deduced amino acid sequences of grass group VI (clone pPHLPVI.F), grass group Va (clone pPHLPVA, 7) and Vb (clone pPHLPVB, 8). Regions with high antigenic index (>1.2) are marked by asterisks.

Discussion

The primary sequence of Phl p 6 was determined by cDNA cloning. Further insight into its relationship with grass group V was obtained by comparing the deduced amino acid sequences.

Screening of the cDNA library of timothy grass pollen was performed using a Phl p 6 directed mouse antiserum. Screening by oligonucleotide primers of the known N-terminal sequence was ruled out because of the high structural similarity to the internal protein sequence of group V [5]. The fact that we obtained six Phl p 6 clones but no Phl p 5 clone demonstrates the suitability of the technique.

Besides the microheterogeneity at the protein level [4], we were also able to identify isoallergenic variants at the cDNA level as well. Only one substitution resulted in a different amino acid residue. This produced a change from arginine to the 'dissimilar' histidine in three of the clones. We could not detect any effect upon the binding of patients' IgE or Phl p 6 antiserum as a result of this. However, we noticed a remarkable decrease in pPHLPVI.F expression compared to pPHLPVI.H (data not shown). This may be a result of pPHLPVI.F containing three ATG codons as probable initiation sites for translation in the leader sequence, while pPHLPVI.H does not contain any ATG codon 5' of the cod-

ing region in the mature protein. Therefore, the existence of leader sequences might interfere with the rate of the fusion protein production.

At the protein level we have identified a mixture of isoforms in one Phl p 6 protein spot (13 kD, pI 5.2) [4], one of them possesses two additional amino acid residues at the N-terminus (glutamic acid and glycine). This corresponds to the protein sequence deduced from the cDNA data 5' of the protein coding region. Perhaps the different ATG positions for initiation have an effect upon the processing of the mature proteins. Phl p 6 is a small 13-kD protein that does not undergo posttranslational modification such as carbohydrate moieties or disulfide bonds. A potential carbohydrate attachment site was predicted by computer algorithm in position 17 (N-A-S) [17]. However, we could not identify any carbohydrate structures by glycan detection [18] nor did we get indications of glycosylation from protein sequencing [4]. Disulfide bonds require the presence of cysteine residues of which Phl p 6 is devoid.

Internal parts of the Phl p 5 proteins showed high sequence homologies to the N- and C-terminus of Phl p 6. The average sequence identities (homologies) match 51% (56%) for Phl p 5a and 54% (61%) for Phl p 5b. It is interesting to note that the amino acid sequence differences between Phl p 6 and the two Phl p 5 subgroups are distributed evenly

throughout the molecule. Our work suggests that Phl p 5 and Phl p 6 have derived from a common precursor gene and that Phl p 5 is the precursor for 5a and 5b (fig. 4). This is supported by the different leader sequences determined from Phl p 6, 5a and 5b [8, 19].

The first predicted antigen binding site (amino acids 30–34) has a high degree of homology between the three proteins and is therefore a strong candidate for a cross-reactive region. We feel that there is too much sequence diversity in the second candidate region (amino acids 46–47) between Phl p 6 and the Phl p 5 subgroups, for it to be considered as a cross-reactive region. The Jameson and Wolf algorithm has an antigenic site prediction reliability of between 50 and 60% [20]. The authors feel this is suggestive enough to warrant further investigation by epitope mapping using overlapping peptides [21].

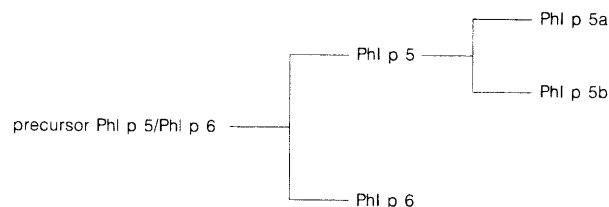


Fig. 4. Probable phylogenetical relationship between the allergen grass groups V and VI.

Acknowledgment

We would like to thank Dr. Michael Duchow (Medizinische Hochschule Lübeck, Germany) for preparing the antigenic site predictions according to the algorithm of Jameson and Wolf. We gratefully acknowledge Mrs. Maren Blümke for her excellent technical assistance.

References

- 1 Löwenstein H: Immunological partial identity and in vitro inhibitory effect of two major timothy pollen allergens to whole pollen extract of four grasses. *Int Arch Allergy Appl Immunol* 1978;57:379–383.
- 2 Matthiesen F, Löwenstein H: Group V allergens in grass pollens. I. Purification and characterization of group V allergen from *Phleum pratense* pollen, *Phl p V*. *Clin Exp Allergy* 1991;21:297–307.
- 3 Petersen A, Schramm G, Becker WM, Schlaak M: Comparison of four grass pollen species concerning their allergens of grass group V by 2D immunoblotting and microsequencing. *Biol Chem Hoppe Seyler* 1993;374:855–861.
- 4 Petersen A, Buße A, Schlaak M, Becker WM: Characterization on the major allergen group VI in timothy grass pollen (Phl p 6). I. Immunological and biochemical studies. *Int Arch Allergy Immunol* 108:49–54.
- 5 Matthiesen F, Friberg L, Olsen M, Löwenstein H: Purification and characterization of the *Phleum pratense* (timothy) pollen allergen *Phl p VI*; in Kraft D, Schon A (eds): *Molecular Biology and Immunology of Allergens*. Boca Raton, CRC Press, 1993, pp 189–191.
- 6 Donovan GR, Baldo BA: Recombinant DNA approaches to the study of allergens and allergenic determinants. *Monogr Allergy* 1990;28:52–83.
- 7 Buße A, Becker WM, Schramm G, Petersen A, Mamat U, Schlaak M: Major allergen *Phl p Va* (timothy grass) bears at least two different IgE reactive epitopes. *J Allergy Clin Immunol* 1994;94:173–181.
- 8 Buße A: Nucleotide sequence of pPHLP5B, Genbank/EMBL Data Library, accession number Z27083 (1992).
- 9 Becker WM, Schaubschläger W, Westphal W, Schlaak M: Purification of grass pollen allergens by means of immobilized monoclonal antibodies; in Chmiel H, Hammes WP, Baily JE (eds): *Biochemical Engineering*. Stuttgart, Fischer, 1988, pp 440–442.
- 10 Petersen A, Schramm G, Buße A, Schlaak M, Becker WM: Structural investigations of the major allergen *Phl p I* on the cDNA and protein level. *J Allergy Clin Immunol*, in press.
- 11 Laemmli UK: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970;227:680–685.
- 12 Kyhse-Andersen J: Electrophoretic transfer of multiple gels: A simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. *Biophys Methods* 1984;10:203–209.
- 13 Leary JJ, Brigati DJ, Ward DC: Rapid and sensitive colorimetric method for visualizing biotin-labeled DNA probes hybridized to DNA or RNA immobilized on nitrocellulose: Bio-blot. *Proc Natl Acad Sci USA* 1983;80:4045–4049.
- 14 Gubler U, Hoffmann BJ: A simple and very efficient method for generating cDNA libraries. *Gene* 1983;25:263–268.
- 15 Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* 1977;74:5463–5469.
- 16 Jameson BA, Wolf H: The antigenic index: A novel algorithm for predicting antigenic determinants. *CABIOS* 1988;4:181–186.
- 17 Marshall RD: Glycoproteins. *Annu Rev Biochem* 1972;41:673–702.
- 18 Petersen A, Becker WM, Schlaak M: Epitope mapping of allergens and antigens of timothy pollen extract. *Appl Theor Electrophor* 1991;2:129–133.
- 19 Vrtala S, Sperr WR, Reimitzer I, van Ree R, Laffer S, Müller WD, Valent P, Lechner K, Rumpold H, Kraft D, Scheiner O, Valenta R: cDNA cloning of a major allergen from timothy grass (*Phleum pratense*) pollen; characterization of the recombinant *Phl p V* allergen. *J Immunol* 1992;151:4773–4781.
- 20 Pellequer JL, Westhof E, van Regenmortel MHV: Predicting locations of continuous epitopes in proteins from their primary structure. *Methods Enzymol* 1991;203:176–201.
- 21 Walsh BJ, Howden MEH: A method for the detection of IgE binding sequences of allergens based on a modification of epitope mapping. *J Immunol Methods* 1989;121:275–280.