J Sci Food Agric 1996, **71**, 475–482

Apple Allergy: The cDNA Sequence of the Major Allergen of Apple, Determined by Performing PCR with a Primer Based on the *N*-Terminal Amino Acid Sequence, is Highly Homologous to the Sequence of the Major Birch Pollen Allergen

Britta Schöning,[a] Wolfgang H Ziegler,[b] Stefan Vieths[c]* and Werner Baltes[a]

[a] Institute of Food Chemistry, Technische Universität Berlin, Germany
[b] Franz-Volhard-Klinik of UKRV at MDC, Berlin Buch, Germany
[c] Paul-Ehrlich-Institute, Department of Allergology, Langen, Germany

(Received 23 August 1995; revised version received 19 January 1996; accepted 23 February 1996)

Abstract: Considering the known N-terminal amino acid sequence of the major apple allergen, a polymerase chain reaction (PCR) primer was selected to amplify cDNA encoding this protein. A single PCR product was obtained, cloned into *Escherichia coli* and subsequently sequenced. The missing 5′-end of the apple cDNA sequence was obtained by a 5′-RACE method. The cDNA sequence showed 72% identity with the coding region of one of the known isoforms of Bet v 1, the major allergen of birch pollen. The deduced amino acid sequence resulted in a 158-residue protein with a calculated molecular mass of 17·5 kDa and 63% amino acid sequence identity to Bet v 1. In addition, further protein alignments showed a high degree of identity with allergens from other tree pollens and some 'pathogenesis-related proteins' from food plants. According to international regulations the allergen was termed Mal d 1 for this protein, it being the first major allergen discovered and characterised in fruits of apple (*Malus domestica*).

Key words: Bet v 1, birch pollen related food allergy, cDNA sequence, cross-reactivity, major apple allergen, Mal d 1, pathogenesis related proteins, PCR, 5′-RACE.

INTRODUCTION

Allergy to fresh apples as well as to other fruits, nuts and vegetables (for example, cherry, plum, peach, kiwi, celery, carrot) is highly associated with type I allergy to birch pollen (Calkhoven *et al* 1987; Dreborg 1988; Helbling *et al* 1993). Ingestion of these plant foods causes mainly oropharyngeal symptoms (Dreborg 1988; Ortolani *et al* 1988; Helbling *et al* 1993), mediated by the patients' allergen specific IgE antibodies. Approximately 70% of all birch pollen allergic patients suffer from these so-called 'birch pollen related food allergies' (Eriksson *et al* 1982; Dreborg and Foucard 1983; Dreborg 1988; Möller 1989). The observed clustering of hypersensitivities is due to cross-reactions of specific IgE antibodies with corresponding epitopes on the allergens of food and pollen (Björkstén *et al* 1980; Halmepuro *et al* 1984; Halmepuro and Løwenstein 1985; Calkhoven *et al* 1987; Ebner *et al* 1991; Vieths *et al* 1992a).

Besides profilins, a family of cross-reacting plant pan-allergens (Valenta *et al* 1992; Van Ree *et al* 1992), and 'cross-reactive carbohydrate determinants' (Aalberse *et al* 1981; Calkhoven *et al* 1987; Vieths *et al* 1994b), most of the birch pollen related food hypersensitivities are caused by a family of plant proteins sharing common epitopes with the major allergen of birch pollen, Bet v 1

* Author to whom correspondence should be addressed at Paul-Ehrlich-Institut, Bundesamt für Sera und Impfstoffe, Abteilung für Allergologie, Paul-Ehrlich-Str. 51-59, D-63225 Langen, Germany.

J Sci Food Agric 0022-5142/96/$09.00

(Ebner *et al* 1991; Vieths *et al* 1993b, 1994c; Schöning *et al* 1995). Studies with patients' sera and monoclonal antibodies specific to Bet v 1 revealed common structures of Bet v 1 and an 18 kDa allergen from apple (Ebner *et al* 1991; Vieths *et al* 1993b, 1994c; Schöning *et al* 1995). IgE specific for this protein has been detected in 65–83% of sera from apple allergic patients (Ebner *et al* 1991; Vieths *et al* 1995a). It has therefore been concluded that this protein represents the major allergen of apple.

Recently, we further confirmed the presumed relationship between Bet v 1 and the corresponding allergens in fruits and vegetables by two-dimensional immunoblotting and N-terminal microsequencing with allergens of apple, cherry, carrot and celery (Vieths *et al* 1994c; Schöning *et al* 1995). Based on the first 15 N-terminal amino acid residues, the apple allergen, for example, revealed 53% sequence identity with Bet v 1 and cross-reacted with different Bet v 1 specific monoclonal antibodies and rabbit antisera as well as patients' sera in IgE immunoblotting experiments.

Using a Bet v 1 encoding cDNA clone as a probe, northern blotting experiments with an apple RNA preparation suggested a significant homology of the two allergens at the nucleic acid level (Ebner *et al* 1991).

In order to confirm these results, and because apple hypersensitivity has been studied as an immunopathologic model representing the typical situation of the whole phenomenon (Björkstén *et al* 1980; Ebner *et al* 1991; Vieths *et al* 1993a,b), cDNA encoding the major part of the 18 kDa apple allergen was amplified by means of polymerase chain reaction (PCR), cloned and sequenced. This paper describes the selection of suitable PCR primers, the application of a 5'-RACE method to achieve the missing 5'-end of the apple cDNA, and presents a complete cDNA sequence from which the amino acid sequence of the major apple allergen was deduced.

MATERIALS AND METHODS

RNA preparation from apple and cDNA synthesis

Apples (Granny Smith) were purchased at the local market. Total RNA was prepared by grinding iced apple slices in a guanidium thiocyanate buffer followed by fractional precipitation (MacDonald *et al* 1987). Single-stranded cDNA was prepared with the GeneAmp RNA PCR kit (Perkin Elmer Cetus, Norwalk, CT, USA) using the degenerated oligo dT primer dTdeg; TTTTTTTTTTTTTTTNN-3' (purchased from TIB MOLBIOL, Berlin, Germany). All reactions were carried out with 0·45 μg of total RNA in a final volume of 20·0 μl. The probes were denatured for 10 min at 65°C and then put on ice for 5 min. After adding reverse transcriptase (2·5 units) and RNase inhibitor (1·0 unit) reverse transcription was carried out for 60 min at 42°C followed by an inactivation step for 5 min at 99°C. Controls omitting reverse transcriptase were included.

Polymerase chain reaction (PCR) primers

The design of a specific 5' primer (upstream) was based on the N-terminal sequence of the 18 kDa apple allergen (Vieths *et al* 1994c; Schöning *et al* 1995) to amplify and isolate cDNA encoding this protein. Back translation from peptide sequence to the corresponding cDNA is difficult due to the degeneracy of the genetic code. Therefore the primer was directed against a region rich in amino acids encoded only by one or two different codons. Taking advantage of the high homology between the known N-terminal amino acid sequences of the apple allergen and Bet v 1 (Ebner *et al* 1991; Vieths *et al* 1994c; Schöning *et al* 1995), the final oligonucleotide sequence was further adapted by comparing all possible codon combinations with the known cDNA sequence of Bet v 1 (Breiteneder *et al* 1989). The combination with the lowest number of base exchanges was chosen to be tested as a specific primer for the apple allergen cDNA and termed primer Mald1/1: CGAAAATGAGTTCACCTCTGA-3' (purchased from TIB MOLBIOL, Berlin, Germany). For the reverse direction the primer dTdeg (see above) was used.

Polymerase chain reaction (PCR)

After optimising, PCR reactions were carried out in a final reaction volume of 100 μl using a Perkin Elmer 9600 (Perkin Elmer Cetus) cycler. Each reaction mix contained 1·0 × PCR-buffer II (Perkin Elmer Cetus), 1·0 mM $MgCl_2$, 0·2 μmol of dNTP (Promega, Madison, WI, USA), 100 pmol of primer dTdeg (see above), 20·0 pmol of primer Mald1/1 (see above), 2·0 μl of the reverse transcription reaction solution (see above) and 2·5 units of Taq polymerase (*AmpliTaq* DNA Polymerase, Perkin Elmer Cetus). The tubes were initially denatured for 40 s at 95°C, then annealed for 40 s at 47°C, and extended for 60 s at 72°C. In the following 39 cycles denaturation time was reduced to 20 s. A final extension reaction of 10 min was added, to ensure that all of the amplified products were full length. Negative controls, without cDNA template, and controls of genomic DNA traces in the RNA preparation were included.

PCR product cloning

Standard protocols were used for DNA manipulations (Sambrook *et al* 1989). Fifteen microlitres of the PCR samples were analysed on a 20 g kg^{-1} agarose gel. The remainder was purified using a spin column (Clon Tech)

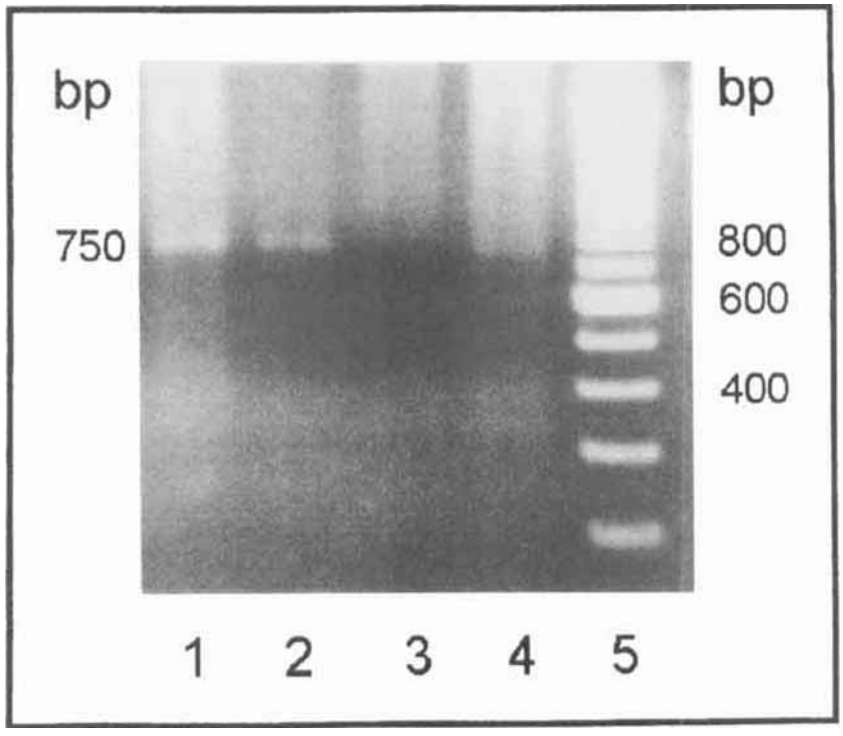

Fig 1. Agarose gel (20 g kg^{-1}, 0·5 × TBE) of the amplified apple cDNA after PCR using primers Mald1/1 and dTdeg (15 µl each). Variation of $MgCl_2$ concentration: lane 1, 1·0 mM $MgCl_2$ (best results); lane 2, 1·5 mM $MgCl_2$; lane 3, negative control; lane 4, 0·5 mM $MgCl_2$; lane 5, marker (100 bp ladder, Gibco BRL).

to remove primers. One microlitre of purified PCR-product was directly ligated into pGEM-T-vector (pGEM-T Vector System I, Promega, Madison, USA) and then transfected into competent cells of the *E coli* strain DH5α (GIBCO BRL, Gaithersburg, USA). Cells were plated and grown overnight on Ampicillin, blue/white selection dishes. White colonies were picked and used to prepare stocks and purified plasmids. Minor parts of the plasmid preparations were digested simultaneously with Aat II (USB, USA) and Sac I (Boehringer, Mannheim, Germany) and analysed on a 20 g kg^{-1} agarose gel. Positive clones carrying a plasmid with an insert of approximately 750 basepairs (bp) were selected for the sequencing procedure.

5'-RACE

Rapid Amplification of cDNA Ends (RACE) was used for the amplification of the unknown 5'-end of the mRNA of the apple allergen. After DNA sequencing of the plasmids two different gene specific antisense primer were selected for the application in an 'anchored' PCR method, Mald1/2: CCCAAAACACCATAGATACTT-3' and Mald1/3: GGCCTTCTCTTTTCCAACCTT-3' (purchased from TIB MOLBIOL, Berlin, Germany). First strand cDNA synthesis was primed with Mald1/2. A terminal deoxynucleotidyl transferase was used to add a homopolymeric dC tail to the 3'-end of the cDNA. Tailed cDNA was amplified by PCR using a mixture of three primers, the nested gene specific primer Mald1/3 and a combination of a complementary homopolymer-containing adapter primer and a corresponding universal primer. A single 5'-RACE product with approximately 550 bp was obtained using a 5'-RACE system (GIBCO BRL, Gaithersburg, USA) according to the manufacturers' instructions. After removal of primers the product was directly sequenced two times with two different primers.

DNA sequence analysis

DNA sequence analysis of the plasmids was performed on a Vistra 725 automatic fluorescent DNA sequencer (Molecular Dynamics, Amersham International, Little Chalfont, UK) using a cycle sequencing kit (ΔTaq fluorescent dye-primer cycle sequencing kit, Amersham International) for dideoxy sequencing. The primers used for sequencing, M13 forward (−21) and M13 reverse (−26), were 5' Texas Red labelled 18mers. The insert was sequenced four times from both directions.

DNA sequence analysis of the 5'-RACE product was performed with an ABI 373A automatic fluorescent DNA sequencer (Applied Biosystems Inc) using a dye-terminator cycle sequencing kit (PRISM™ Ready Reaction Dye Deoxy™ Terminator Cycle Sequencing Kit, Applied Biosyss Inc). Based on the DNA sequence obtained previously, two different primers were selected for sequencing,

Mald1/4: CGATGGTGTCTGTCAAAG-3′ and Mald1/5: GGATTTCAGCTTGCTTGATTG-3′ (purchased from TIB MOLBIOL, Berlin, Germany).

RESULTS

The yield of RNA prepared from Granny Smith apples was very low (approximately 3·0 μg total RNA per 1·0 g fresh apple), but total RNA could be used to synthesise single stranded cDNA with the poly-(A)-tail primer dTdeg. The primers Mald1/1 and dTdeg used in PCR reactions with apple cDNA produced one single band approximately 750 basepairs (bp) in length (see Fig 1). No bands were found in the controls.

After removal of primers a small amount of the crude PCR-product was ligated into pGEM-T-vector and transferred into *E coli*. Clones containing an insert of approximately 750 bp were found and selected for DNA sequencing. The application of a 5′-RACE-system to the apple mRNA resulted in one single product of approximately 550 bp. After removal of primers this product was sequenced directly.

DNA sequence analysis of both the cloned PCR product and the 5′-RACE product resulted in an overlap of 219 bases, which matched completely. The entire sequence of the cDNA is presented in Fig 2 together with the deduced amino acid sequence. The cDNA sequence showed 72% identity with the coding region of BV1SC3 (unpublished, EMBL AC X77601), one of the known isoforms of Bet v 1. The overall information results in a 158-residue protein with a calculated molecular mass of 17·5 kDa and no N-glycosylation site. Prediction of the amino acid sequence from the cDNA data and the protein sequence achieved earlier by N-terminal microsequencing showed an overlap of 26 amino acids with 25 of them being identical (see Fig 2). The nucleotide sequence was submitted to the EMBL Nucleotide Sequence Database (Accession No: X83672 MDMALD1).

A comparison of the cDNA sequence presented here to all sequences in the EMBL Nucleotide Sequence Database revealed homologies with many different nucleotide sequences, all of them encoding tree pollen allergens or proteins belonging to a family of plant 'pathogenesis-related proteins' (Breiteneder *et al* 1989;

```
AA                                                                          NS

            ACAATACTCTCAACCCTCAATAAAACCATCATCCTTCCTTGCTCATTTTCCAACCTTTTTTTAATCATC   -1

   ATG GGT GTC TAC ACA TTT GAG AAC GAG TTC ACC TCT GAG ATT CCA CCA TCA AGA TTG TTC   60
 1 Met Gly Val Tyr Thr Phe Glu Asn Glu Phe Thr Ser Glu Ile Pro Pro Ser*Arg Leu Phe

   AAG GCC TTT GTC CTT GAT GCT GAC AAC CTC ATC CCC AAG ATT GCA CCC CAG GCA ATC AAG  120
20 Lys Ala Phe Val Leu Asp Ala Asp Asn Leu Ile Pro Lys Ile Ala Pro Gln Ala Ile Lys

   CAA GCT GAA ATC CTT GAA GGA AAC GGT GGC CCC GGA ACC ATC AAG AAG ATC ACT TTT GGT  180
40 Gln Ala Glu Ile Leu Glu Gly Asn Gly Gly Pro Gly Thr Ile Lys Lys Ile Thr Phe Gly

   GAA GGC AGC CAG TAC GGC TAC GTG AAG CAC AGG ATT GAC TCG ATT GAC GAA GCA AGC TAC  240
60 Glu Gly Ser Gln Tyr Gly Tyr Val Lys His Arg Ile Asp Ser Ile Asp Glu Ala Ser Tyr

   TCA TAC TCC TAC ACT TTG ATT GAA GGA GAT GCT TTG ACA GAC ACC ATC GAG AAA ATA TCT  300
80 Ser Tyr Ser Tyr Thr Leu Ile Glu Gly Asp Ala Leu Thr Asp Thr Ile Glu Lys Ile Ser

   TAC GAG ACC AAG TTG GTG GCA TGT GGA AGT GGT TCC ACC ATC AAG AGC ATC AGT CAT TAC  360
100 Tyr Glu Thr Lys Leu Val Ala Cys Gly Ser Gly Ser Thr Ile Lys Ser Ile Ser His Tyr

   CAC ACC AAG GGA AAC ATT GAA ATC AAG GAA GAG CAC GTC AAG GTT GGA AAA GAG AAG GCC  420
120 His Thr Lys Gly Asn Ile Glu Ile Lys Glu Glu His Val Lys Val Gly Lys Glu Lys Ala

   CAT GGT TTG TTC AAG CTT ATT GAG AGC TAC CTT AAG GAC CAC CCC GAC GCA TAC AAC TAA  480
140 His Gly Leu Phe Lys Leu Ile Glu Ser Tyr Leu Lys Asp His Pro Asp Ala Tyr Asn

   ATTAATCATAAGTATCTATGGTGTTTTGGGTGTTACCGTTACTATATGGTCAGTCGAAGGTTGTGTGGCTTTTCTTTGTA 560

   TGTTTTTTTCACGTCAGCCAAAAGTAATAACCATGGTTTGGGCAGTTTGGGCTTGAAAAATAAGTGTTATGGTTGTGATG 640

   ATCTTCTTGTTGCGATGTCTTTGTTTGGAAAAAAGTTGGCAGACTGAACGGAGTGTATTATCTTCTATAAATATAATGAA 720

   ATTACTTTTTAATCAAAAAAAAAA                                                        744

   * a proline residue has been detected by N-terminal microsequencing
```

Fig 2. Overall nucleotide sequence of the cDNA and predicted amino acid sequence of the major apple allergen. Numbers indicate amino acid residues (left) and nucleotide positions (right) of the sequence. The three-letter code for amino acids is used. Amino acids identified by N-terminal microsequencing (Vieths *et al* 1994c; Schöning *et al* 1995) are underlined.

```
Mal d 1      GVYTFENEFT SEIPPSRLFK AFVLDADNLI PKIAPQAIKQ AEILEGNGGP GTIKKITFGE
BV1SC3       GVFNYEDEAT SVIAPARLFK SFVLDADNLI PKVAPENVSS AENIEGNGGP GTIKKITFPE
Bet v 1      GVFNYETETT SVIPAARLFK AFILDGDNLF PKVAPQAISS VENIEGNGGP GTIKKISFPE
Aln g 1      GVFNYEAETP SVIPAARLFK AFILDGDKLL PKVAPEAVSS VENIEGNGGP GTIKKITFPE
Cor a 1(5)   GVFNYEVETP SVIPAARLFK SYVLDGDKLI PKVAPQAITS VENVEGNGGP GTIKNITFGE
GMHA         GIFTFEDETT SPVAPATLYK ALVTDADNVI PKAV-EAFRS VENLEGNGGP GTIKKITFVE
Car b 1(2)   GVFNYEAETT SVIPAARLFK AFILDGNKLI PKVSPQAVSS VENVEGNGGP GTIKKITFSE
PEADRRA      GVFNVEDEIT SVVAPAILYK ALVTDADTLT PKVI-DAIKS IEIVEGNGGA GTIKKLTFVE

Mal d 1      GSQYGYVKHR IDSIDEASYS YSYTLIEGDA LTDTIEKISY ETKLVA-CGS GSTIKSISHY
BV1SC3       GSHFKYMKHR VDEIDHANFK YCYSIIEGGP LGDTLEKISY EIKIVAAPGG GSILKITSKY
Bet v 1      GFPFKYVKDR VDEVDHTNFK YNYSVIEGGP IGDTLEKISN EIKIVATPDG GSILKISNKY
Aln g 1      GSPFKYVKER VDEVDRVNFK YSFSVIEGGA VGDALEKVCN EIKIVAAPDG GSILKISNKF
Cor a 1(5)   GSRYKYVKER VDEVDNTNFT YSYTVIEGDV LGDKLEKVCH ELKIVAAPGG GSILKISSKF
GMHA         DGESKFVLHK IESVDEANLG YSYSVVGGVG LPDTVEKITF ECKLAAGANG GSAGKLTVKY
Car b 1(2)   GSPVKYVKER VEEIDHTNFK YNYTVIEGDV LGDKLEKVSH ELKIVAAPGG GSIVKISSKF
PEADRRA      DGETKHVLHK VELVDVANLA YNYSIVGGVG FPDTVEKISF EAKLSAGPNG GSIAKLSVKY

Mal d 1      HTKGNI-EIK EEHVKVGKEK AHGLFKLIES YLKDHPDAYN
BV1SC3       HTKGDI-SLN EEEIKAGKEK GAGLFKAVEN YLVAHPNAYN
Bet v 1      HTKGDH-EVK AEQVKASKEM GETLLRAVES YLLAHSDAYN
Aln g 1      HTKGDH-EIN AEQIKIEKEK AVGLLKAVES YLLAHSDAYN
Cor a 1(5)   HAKGDH-EIN AEEMKGAKEM AEKLLRAVET YLLAHSAEYN
GMHA         QTKGDA-QPN PDDLKIGKVK SDALFKAVEA YLLANPH-YN
Car b 1(2)   HAKGYH-EVN AEEMKGAKEM AEKLLRAVES YLLAHTAEYN
PEADRRA      YTKGDAAAPT EEQLKSDKAK GDGLFKALER YCLAHPD-YN
```

Fig 3. Sequence comparison of the major apple allergen Mal d 1 with BV1SC3 (unpublished, EMBL AC X77601) and Bet v 1 (Breiteneder *et al* 1989), two isoforms of the major allergen from birch pollen, Aln g 1, the major allergen from alder pollen (Breiteneder *et al* 1992), Cor a 1(5), an allergen of hazel pollen (Breiteneder *et al* 1993), GMH4, a stress-induced protein from soya bean (Crowell *et al* 1992), PEADRRA, a disease resistance response protein from pea (Fristensky *et al* 1988), and Car b 1(2), a pollen allergen from hornbeam (Larsen *et al* 1992). The one letter code for amino acids is used. Overall identity is indicated by bold letters.

Walter *et al* 1990; Moiseyev *et al* 1994). Seven examples out of the highly homologous sequences were chosen for an amino acid alignment. The proteins encoded by the different cDNA sequences are of nearly identical size between 158 and 160 amino acids.

TABLE 1
Degree of identity with the major apple allergen Mal d 1[a]

BV1SC3	63%
Bet v 1	56%
Aln g 1	54%
Cor a 1(5)	54%
Car b 1(2)	52%
GMH4	49%
PEADRRA	48%

[a] Degree of identity of the amino acid sequences of the major apple allergen with BV1SC3 (unpublished, EMBL AC X77601) and Bet v 1 (Breiteneder *et al* 1989), two isoforms of the major allergen from birch pollen, Aln g 1, the major allergen from alder pollen (Breiteneder *et al* 1992), Cor a 1(5), an allergen of hazel pollen (Breiteneder *et al* 1993), GMH4, a stress-induced protein from soya bean (Crowell *et al* 1992), PEADRRA, a disease-resistance response protein from pea (Fristensky *et al* 1988), and Car b 1(2), a pollen allergen from hornbeam (Larsen *et al* 1992).

Figure 3 shows the alignment of the deduced amino acid sequence (see Fig 2) with BV1SC3 (unpublished, EMBL AC X77601) and Bet v 1 (Breiteneder *et al* 1989), two isoforms of the major allergen from birch pollen, Aln g 1, the major allergen from alder pollen (Breiteneder *et al* 1992), Cor a 1(5), an allergen of hazel pollen (Breiteneder *et al* 1993), GMH4, a stress-induced protein from soya bean (Crowell *et al* 1992), PEADRRA, a disease resistance response protein from pea (Fristensky *et al* 1988), and Car b 1(2), a pollen allergen from hornbeam (Larsen *et al* 1992).

In addition to Fig 3, Table 1 presents the high degree of sequence identity of the deduced amino acid sequence (see Fig 2) with the seven proteins chosen for sequence alignment.

DISCUSSION

Strong evidence exist for the sequence data shown in Fig 2 representing the cDNA sequence of the major apple allergen. First, the total length of the deduced amino acid sequence of the putative major allergen of apple and its calculated molecular mass (MM) of 17·5 kDa correlates well with the apparent MM of 18·0 kDa observed in Disc-SDS-PAGE (Vieths *et al* 1992b, 1994a,c, 1995b; Schöning *et al* 1995). Secondly, the first 26 amino acid residues of the sequence of the

major apple allergen as obtained by N-terminal microsequencing (Vieths *et al* 1994c) are identical with the deduced amino acid sequence given in Fig 2 with the exception of pos 16. A serine is encoded by the cDNA sequence in comparison to proline analysed by microsequencing. This mismatch may be due to slightly different isoforms of the allergen in preparations of different apple sources or to an O-glycosylation site at serine pos 16. Since the major allergen of apple is Con A-reactive (own unpublished results), and has no N-glycosylation site, it may carry some O-linked sugar residues. In this case the serine residue would not be detectable by N-terminal sequencing, due probably to a glycosylation, while the proline signal could result from a memory peak from proline in pos 14 and 15.

The sequence data of the major apple allergen presented here (Fig 2) have to be evaluated in context with the results of Ebner *et al* (1991) and our own group (Vieths *et al* 1995a), documenting a frequency of IgE response in apple allergic patients of 87% and 65%, respectively, as well as the immunochemical characterisation data of this allergen (Björkstén *et al* 1980; Halmepuro *et al* 1984; Halmepuro and Løwenstein 1985; Calkhoven *et al* 1987; Ebner *et al* 1991; Vieths *et al* 1992a,b, 1993a,b, 1994a,c, 1995a,b; Schöning *et al* 1995). Furthermore, we have purified the allergen to a degree >95%, and studied the allergenic potency in comparison to Bet v 1 quantitatively (Vieths *et al* 1995b). The data obtained by all these investigations clearly fulfil the requirements of the WHO/IUIS allergen nomenclature regulations (King *et al* 1994). We propose to term this protein, being the first major allergen discovered and characterised in fruits of apple (*Malus domestica*), Mal d 1 according to the revised allergen nomenclature (King *et al* 1994).

The PCR strategy described here in combination with a 5'-RACE method provided us with the entire cDNA sequence of the major apple allergen and its deduced amino acid sequence (see Fig 2). The procedure to select a primer from the known N-terminal amino acid sequence of Mal d 1 by comparing the possible nucleotide primer sequences with the presumed homologous cDNA sequence of Bet v 1 (Breiteneder *et al* 1989) proved to be successful, firstly by resulting in only one PCR product. Secondly, the sequencing analysis of the 5'-RACE product revealed only three mismatches out of 21 bases between the cDNA sequence determined and the deduced sequence of the primer Mald1/1.

According to previous results (Vieths *et al* 1994c; Schöning *et al* 1995) a high sequence identity between the cDNA sequence of Mal d 1 and Bet v 1 was expected. The determination of the cDNA sequence presented here and the protein alignment of the deduced amino acid sequence to Bet v 1 confirmed this hypothesis (see Fig 3), and it further explains the known cross-reactivity of the patients' IgE antibodies, rabbit polyclonal antisera and some monoclonal antibodies directed against epitopes on Bet v 1 (Ebner *et al* 1991; Vieths *et al* 1993b, 1994c; Schöning *et al* 1995). Bet v 1 represents a gene family (Swoboda *et al* 1995) with some members showing a higher degree of identity with our sequence shown in Fig 2 than the Bet v 1 protein sequence published first (Breiteneder *et al* 1989). For example, BV1SC3 (unpublished, EMBL AC X77601) revealed 63% sequence identity with the major apple allergen compared to 57% with Bet v 1 according to Breiteneder *et al* (1989) (see Table 1). Individual sensitisation and cross-reactivity patterns of different allergic patients may be due to these sequence variations.

The protein alignments (see Fig 3) further suggest a possible cross-reactivity of patients' IgE antibodies directed against the major allergens of hazel, alder and hornbeam pollen and apple. This contingent enlarged cross-reactivity needs further investigations. The sequence of Mal d 1 shows a high degree of identity with many different proteins (see Table 1), all of them being tree pollen allergens or members of a family of plant 'pathogenesis-related proteins' (PRP). The latter have been discovered after treating plants and plant cell cultures with different elicitors (Fristensky *et al* 1988; Walter *et al* 1990). The high degree of amino acid sequence identity of Mal d 1, 49% with a PRP from pea for example, and the pollen allergens listed in Table 1 to these PRP is striking because little is known about the function of Mal d 1 in apple fruits or Bet v 1 in birch pollen. According to Breiteneder *et al* (1989) it is not unlikely that Bet v 1 fulfils a pathogen defence function in birch. The detailed function, in particular an enzymatic activity, has remained unknown, but latest findings suggest that these proteins may have ribonuclease activity (Moiseyev *et al* 1994; Bufe A, Borstel Research Institute, pers comm).

The deduced amino acid sequence of Mal d 1 now available helps to increase the understanding of the phenomenon of birch pollen related food hypersensitivity. Epitope mapping could be carried out by testing the reactivity of short synthetic peptides derived from the sequence. The whole cDNA of Mal d 1 may be cloned and expressed in *Escherichia coli*. Large amounts of this recombinant allergen would be helpful in T and B cell experiments, which may clarify the pathomechanism of cross-allergenicity between birch pollen and fruits. Further work on this field is in progress.

Since the degree of expression of Mal d 1 seems to be dependent on the apple strain (Vieths *et al* 1994a), the knowledge of the nucleic acid sequence might lead to the development of a PCR-based method of screening old and new apple strains and to the preparation of a phylogenetic tree for the expression of the major apple allergen. That might help food growers cultivating apple strains with low contents of Mal d 1, which may be 'low allergenic' (Vieths *et al* 1994a), and therefore may help to prevent the development of the birch pollen related apple allergy in birch pollen sensitised patients.

ACKNOWLEDGEMENT

This work was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG; Ba 412/20/1,2).

REFERENCES

Aalberse R C, Koshte V, Clemens J G J 1981 Immunoglobulin E antibodies that crossreact with vegetable foods, pollen and Hymenoptera venom. *J Allergy Clin Immunol* **68** 356–364.

Björkstén F, Halmepuro L, Hannuksela M, Lahti A 1980 Extraction and properties of apple allergens. *Allergy* **35** 671–677.

Breiteneder H, Pettenburger K, Bito A, Valenta R, Kraft D, Rumpold H, Scheiner O, Breitenbach M 1989 The gene coding for the major birch pollen allergen *Bet v* I, is highly homologous to a pea disease resistance response gene. *EMBO J* **8** 1935–1938.

Breiteneder H, Ferreira F, Reikerstorfer A, Duchene M, Valenta R, Hoffmann-Sommergruber K, Ebner C, Breitenbach M, Kraft D, Scheiner O 1992 Complementary DNA cloning and expression in Escherichia coli of Aln g I, the major allergen in pollen of alder (*Alnus glutinosa*). *J Allergy Clin Immunol* **90** 909–917.

Breiteneder H, Ferreira F, Hoffmann-Sommergruber K, Ebner C, Breitenbach M, Rumpold H, Kraft D, Scheiner O 1993 Four recombinant isoforms of Cor a I, the major allergen of hazel pollen, show different IgE binding properties. *Eur J Biochem* **212** 355–362.

Calkhoven P G Aalbers M, Koshte V L, Pos O, Oei H D, Aalberse R C 1987 Cross-reactivity among birch pollen, vegetables and fruits as detected by IgE antibodies is due to at least three distinct cross-reactive structures. *Allergy* **42** 382–390.

Crowell D, John M E, Russell D, Amasino R M 1992 Characterization of a stress-induced developmentally regulated gene family from soya bean. *Plant Molec Biol* **18** 459–466.

Dreborg S 1988 Food allergy in pollen-sensitive patients. *Ann Allergy* **61** 41–46.

Dreborg S, Foucard T 1983 Allergy to apple, carrot and potato in children with birch pollen allergy. *Allergy* **38** 167–172.

Ebner C, Birkner T, Valenta R, Rumpold H, Breitenbach M, Scheiner O, Kraft D 1991 Common epitopes of birch pollen and apples-studies by Western and northern blot. *J Allergy Clin Immunol* **88** 588–594.

Eriksson N E, Formgren H, Svenonius E 1982 Food hypersensitivity in patients with pollen allergy. *Allergy* **37** 437–443.

Fristensky B, Horovitz D, Hadwiger L A 1988 cDNA sequences for pea disease resistance response genes. *Plant Molec Biol* **11** 713–715.

Halmepuro L, Løwenstein H 1985 Immunological investigation of possible structural similarities between pollen antigens and antigens in apple, carrot and celery tuber. *Allergy* **40** 264–272.

Halmepuro L, Vuontela K, Kalimo K, Björkstén F 1984 Cross-reactivity of IgE antibodies with allergens in birch pollen, fruits and vegetables. *Int Archs Allergy Appl Immunol* **74** 235–240.

Helbling A, Lopez M, Schwartz H J, Lehrer S B 1993 Reactivity of carrot-specific IgE antibodies with celery, apiaceous spices and birch pollen. *Ann Allergy* **70** 495–499.

King T P, Hoffman D, Løwenstein H, Marsh D G, Platts-Mills T A E, Thomas W 1994 Allergen Nomenclature. *Int Arch Allergy Immunol* **105** 224–233.

Larsen J N, Stroeman P, Ipsen H 1992 PCR based cloning and sequencing of isogenes encoding the tree pollen major allergen Car b I from Carpinus betulus hornbeam. *Molec Immunol* **29** 703–711.

MacDonald R J, Swift G H, Przybyla A E, Chirgwin J M 1987 Isolation of RNA using guanidinium salts. *Meth Enzymol* **152** 119–227.

Moiseyev G P, Beintema J J, Fedoreyeva L I, Yakovlev G I 1994 High sequence similarity between a ribonuclease from *ginseng* calluses and fungus-elicted proteins from parsley indicates that intracellular pathogenesis-related proteins are ribonucleases. *Planta* **193** 470–472.

Möller C 1989 Effect of pollen immunotherapy on food hypersensitivity in children with birch pollinosis. *Ann Allergy* **62** 343–345.

Ortolani C, Ispano M, Pastorello E, Bigi A, Ansaloni R 1988 The oral allergy syndrome. *Ann Allergy* **61** 47–52.

Sambrook J, Fritsch E F, Maniatis T 1989 *Molecular Cloning. A Laboratory Manual* (2nd edn). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA.

Schöning B, Vieths S, Petersen A, Baltes W 1995 Identification and characterization of allergens related to *Bet v* I, the major birch pollen allergen, in apple, cherry, celery and carrot by two-dimensional immunoblotting and N-terminal microsequencing. *J Sci Food Agric* **67** 431–440.

Swoboda I, Jilek A, Ferreira F, Engel E, Hoffmann-Sommergruber K, Scheiner O, Kraft D, Breiteneder H, Pittenauer E, Schmid E, Vicente O, Heberle-Bors E, Ahorn H, Breitenbach M 1995 Isoforms of Bet v 1, the major birch pollen allergen, analyzed by liquid chromatography, mass spectrometry and cDNA cloning. *J Biol Chem* **270** 2607–2613.

Valenta R, Duchene M, Vrtala S, Valent P, Sillaber C, Ferreira F, Tejkl M, Hirschwehr R, Ebner C, Kraft D, Scheiner O 1992 Profilin, a novel plant pan-allergen. *Int Arch Allergy Immunol* **99** 271–273.

Van Ree R, Voitenko V, van Leeuwen W A, Aalberse R C 1992 Profilin is a cross-reactive allergen in pollen and vegetable food. *Int Arch Allergy Immunol* **98** 97–104.

Vieths S, Brockmann S, Schöning B 1992a Nahrungsmittelallergie gegen Obst und Gemüse: Eine aktuelle serologische Untersuchung zur Sensibilisierung von Pollenallergikern. *Allergologie* **15** 367–379.

Vieths S, Schöning B, Baltes W 1992b Allergy to fruits and vegetables in pollen-sensitive patients: allergen characterization by IgE immunoblotting and peroxidase staining. *Food Argic Immunol* **4** 181–197.

Vieths S, Schöning B, Jankiewicz A 1993a Occurrence of IgE binding allergens during ripening of apple fruits. *Food Agric Immunol* **5** 93–105.

Vieths S, Schöning B, Aulepp H, Baltes W 1993b Identifizierung kreuzreagierender Allergene in Pollen und pflanzlichen Lebensmitteln. *Lebensmittelchemie* **47** 49–53.

Vieths S, Jankiewicz A, Schöning B, Aulepp H 1994a Apple allergy: the IgE binding potency of apple strains is related to the occurrence of the 18 kDa allergen. *Allergy* **49** 262–271.

Vieths S, Mayer M, Baumgart M 1994b Food Allergy: specific binding of IgE antibodies from plant food sensitizised individuals to carbohydrate epitopes. *Food Agric Immunol* **6** 453–463.

Vieths S, Schöning B, Petersen A 1994c Characterization of the 18 kDa apple allergen by two-dimensional immunoblotting and microsequencing. *Int Arch Allergy Immunol* **104** 399–404.

Vieths S, Aulepp H, Schöning B, Tschirnich R 1995a Untersuchungen zur Apfelallergie bei Birkenpollenallergikern: Hauttestergebnisse und Immunoblot-Analysen mit

Tieftemperaturextrakten unterschiedlicher Apfelsorten. *Allergologie* **18** 89–97.

Vieths S, Janek K, Aulepp H, Petersen A 1995b Isolation and characterisation of the 18 kDa major apple allergen and comparism with the major birch pollen allergen (Bet v 1). *Allergy* **50** 421–430.

Walter M H, Liu J W, Grand C, Lamb C J, Hess D 1990 Bean pathogenesis-related (PR) proteins deduced from elicitor-induced transcripts are members of a ubiquitous new class of conserved PR proteins including pollen allergens. *Molec Gen Genet* **222** 353–360.