

26475

Michael W. De Vouge<sup>a</sup>

Ajay J. Thaker<sup>a</sup>

Ian H. A. Curran<sup>a</sup>

Lei Zhang<sup>a</sup>

Gauri Muradia<sup>a</sup>

Harold Rode<sup>b</sup>

Hari M. Vijay<sup>a</sup>

## Isolation and Expression of a cDNA Clone Encoding an *Alternaria alternata* Alt a 1 Subunit

<sup>a</sup> Life Sciences Division,

Bureau of Drug Research, and

<sup>b</sup> Bureau of Biologics, Drugs Directorate,

Health Protection Branch, Health Canada,

Ottawa, Canada

### Abstract

*Alternaria alternata* is recognized as an important source of fungal aeroallergens. Alt a 1, the major allergen of this mold, is a dimer of disulfide-linked subunits that migrate in SDS-PAGE under reducing conditions at apparent  $M_r$ s of 14,500 and 16,000. IgE antibodies to this protein are present in the sera of > 90% of *A. alternata*-sensitive individuals. Previous studies from this laboratory showed that the N-termini twenty amino acids of the purified subunits are nearly identical. We now report the isolation of clones from an *A. alternata* (strain 34-016) cDNA library constructed in  $\lambda$ gt11, using rabbit IgG antiserum against partially purified Alt a 1. One of nineteen clones selected from screens totalling 305,000 pfu (rb51) was sequenced, and determined to harbor an insert of 660 bp. An in-frame open reading frame within the cloned insert encodes a peptide of  $M_r$  16,960 that bears no significant homology to known allergens or proteins. The size of the rb51 transcript was determined to be  $\approx 0.7$  kb by Northern analysis of *A. alternata* total RNA. The largely hydrophobic N-terminal region of the peptide contains an  $\alpha$ -helical domain and other features characteristic of membrane targeting or secretory signals. The peptide sequence downstream of this region matches previously sequenced Alt a 1 N-terminal from two independent sources at 17 of 20, and 24 of 26 positions. Recombinant Alt a 1 expressed as a secreted protein in *Pichia pastoris* exists as a dimer in conditioned medium, as shown by immunoblotting under nonreducing conditions. Recombinant Alt a 1, like the natural allergen in *A. alternata* extracts, is also reactive with serum IgE from *A. alternata*-sensitive individuals.

### Key Words

Allergen

Fungi

*Alternaria alternata*

Alt a 1

cDNA cloning

*Pichia pastoris*

IgE immunoblot

### Introduction

Allergies to airborne antigens are estimated to afflict  $\approx 10$ –20% of Europeans and North Americans [1]. Examination of sera from afflicted individuals has determined that *Alternaria alternata* is a prominent source of fungal aeroallergens [2, 3]. Initial attempts to characterize the major allergenic components of *A. alternata* extracts yielded

protein fractions whose components bound IgE in a majority of *A. alternata*-allergic patients [4–6]. One such fraction (ALT-I), a carbohydrate-rich Sephadex G-100 fraction eluted in the  $M_r$  25,000–50,000 range, has been recognized as the source of the major allergenic activity of *A. alternata* [4]. However, this protein fraction was shown to be of heterogeneous composition, consisting of several IgE-binding components [7].

### KARGER

E-Mail karger@karger.ch  
Fax + 41 61 306 12 34  
http://www.karger.ch

© 1996 S. Karger AG, Basel  
1018-2438/96/1114-0385\$10.00/0

This article is also accessible online at:  
http://BioMedNet.com/karger

Correspondence to: Dr. H. M. Vijay  
Life Sciences Division, Bureau of Drug Research  
Health Protection Branch, Health Canada  
Postal Locator 2201C  
Ottawa, ON K1A 0L2 (Canada)

Received:  
October 25, 1995  
Accepted after revision:  
June 26, 1996

Recent immunoblot analyses of *A. alternata* extracts probed with pooled atopic sera have enabled more precise characterization of individual *A. alternata* allergens within fractionated extracts. The most commonly observed allergens in such extracts are five proteins migrating within ranges of 85–95, 55–60, 40–45, 25–30 and 15–20 kDa in SDS-polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions [7–11]. Of these, >90% of atopic sera containing *Alternaria*-specific IgE react positively against a major allergen isolated by several groups and originally designated *Alt a*-29 [12], *Alt a* 1 [13, 14], or *Alt a* Bd 29K [15]. As characterized in our laboratory, this allergen was isolated as an  $M_r$  29,000 protein by gel filtration of acetone-defatted crude extract, followed by DEAE-cellulose ion exchange chromatography [16]. Fractions were obtained that strongly inhibited IgE binding to *A. alternata* extracts in the radioallergosorbent test (RAST), but which failed to generate a passive cutaneous anaphylaxis (PCA) response [17]. Further purification and immunoblot analyses demonstrated that this allergen is a disulfide-linked dimer of subunits with apparent  $M_r$  of 14,500 and 16,000 [12] and nearly identical N-terminal sequences that are also 85% homologous to the corresponding N-terminal residues of a similar isolate obtained by Matthiesen et al. [14]. Based on the striking N-terminal homology between these two protein isolates, this allergen is now collectively termed *Alt a* 1 according to the recently revised system of allergen nomenclature [18].

Progress in the characterization and purification of fungal aeroallergens is complicated by high variability with respect to allergen profile between strains [19–21], as well as difficulties in the recovery of specific spore populations [19]. As a result, immunotherapeutic allergen extracts prepared from fungal species are often of inconsistent composition and quality, containing contaminants which may be ineffective or potentially capable of generating an anaphylactic response. By using recombinant DNA techniques to isolate cDNA clones corresponding to major fungal allergens, these allergens may be produced in organisms engineered for large scale protein expression, and conveniently purified with greater fidelity than in the fungal source. In this study, we have used rabbit polyclonal IgG antibodies against partially purified *Alt a* 1 to isolate and characterize cDNA clones corresponding to this allergen. The expression of this protein in the yeast, *Pichia pastoris*, and its reactivity with serum IgE from *A. alternata*-sensitive individuals is also described.

## Materials and Methods

### Antisera and Immunoscreening

Purification of mRNA and construction of a cDNA library in  $\lambda$ gt11 was carried out by a commercial firm (Clontech) using *A. alternata* total RNA isolated in our laboratory. Screening of the library was carried out using standard protocols [22]. Phage from the unamplified library were plated at 10,000 pfu/100 mm plate and grown for 3.5 h at 42°C. Dry nitrocellulose membranes previously impregnated with 10 mM isopropyl- $\beta$ -D-galactopyranoside were overlaid on the plates and incubated at 37°C for 1.5 h. Duplicate membranes were also applied and incubated for an additional 2 h. Filters were blocked with 1% bovine serum albumin (BSA) in Tris-buffered saline (50 mM Tris-HCl pH 7.5, 150 mM NaCl; TBS) overnight at 4°C.

Polyclonal rabbit antiserum was prepared against *Alt a* 1 that was partially purified by Whatman DE-52 ion exchange chromatography and Sephadex G-100 gel filtration [12], and used in immunoscreening and immunoblots at 1:800 dilution in TBST (TBS containing 0.05% Tween-20). This antiserum was preabsorbed by incubation with dry nitrocellulose membranes impregnated with lysates of isopropyl- $\beta$ -D-galactopyranoside-induced *Escherichia coli* Y1090. Membranes were washed for 3  $\times$  5 min each in TBST and incubated for 1 h with alkaline phosphatase (AP)-conjugated goat anti-rabbit IgG antibody (diluted 2,000 $\times$  with TBST) at room temperature. After washing as before, staining was carried out using nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate in AP buffer (100 mM Tris-HCl pH 9.5, 100 mM NaCl, 5 mM  $MgCl_2$ ).

### Polymerase Chain Reaction and DNA Sequencing of Isolated Clones

Immunoselected phage plaques were resuspended and purified by several rounds of sparse plating and immunoscreening until all plaques on a membrane stained positively. Purified phage plaque lysates (10  $\mu$ l) heated to 99°C for 5 min were subjected to polymerase chain reaction (PCR) using 10 pmol each of  $\lambda$ gt11 forward and reverse primers, 0.2 mM dNTPs, and 5 U of Taq polymerase in a 50- $\mu$ l volume of PCR buffer (10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM  $MgCl_2$ ). The reactions were heated to 95°C for 5 min and subjected to 30 PCR cycles of 95°C for 36 s, 51°C for 36 s and 72°C for 90 s. Aliquots of PCR products were ligated into the plasmid pCR II (Invitrogen) and plasmid mini-prep DNAs were sequenced using the ALF laser fluorescence system and fluorescein-conjugated M13-40 and reverse primers (Pharmacia).

### Southern and Northern Hybridizations

To assess whether immunopositive clones contained sequences encoding the *Alt a* 1 N-terminus, digoxigenin (DIG)-labeled ICHA7 probe was prepared using the DIG Oligo Labeling kit (Boehringer-Mannheim). ICHA7 is an 8-fold degenerate oligonucleotide (GAAGGTGAYTACGTGTGTYAAGAT) designed from the N-terminal peptide sequence of our *Alt a* 1 protein isolate [12]. PCR-amplified inserts from immunoselected positives were run in 0.8% agarose gels, transferred to nylon membrane, and incubated with ICHA7 probe at 5 pmol/ml in hybridization solution (1% blocking agent, 5 $\times$  SSC, 0.1% N-lauroyl sarcosine, 0.02% SDS) for 3 h at 35°C, after a 1-hour pre-hybridization. Washes were for 2 $\times$  5 min with (5 $\times$  SSC, 0.1% SDS) at room temperature, followed by 2 $\times$  10 min washes using (2 $\times$  SSC, 0.1% SDS) also at room temperature. Membranes were blocked for 30 min at room temperature with 1% Blocking Reagent (Boehringer-Mannheim) in TBS. AP-conjugated anti-DIG Fab fragments (Boeh-

ringer-Mannheim) were applied at 1:10,000 dilution in blocking solution, followed by washing for 2×15 min each with TBST containing 0.2% Tween-20, a brief incubation in AP buffer containing 50 mM MgCl<sub>2</sub> and an incubation with Lumi-Phos 530 chemiluminescent detection reagent (Boehringer-Mannheim) for 1 min. After incubating between transparency pages for 30 min at 37°C, blots were exposed to X-ray film for 5–15 min.

DIG-labeled rb51 probe used in Northern analyses of *A. alternata* total RNA was prepared by PCR in reactions as described above, except for the replacement of 0.2 mM Thymidine 5'-Triphosphate (TTP) with 0.13 mM TTP and 0.07 mM DIG-11-dUTP. Total RNA was extracted from mycelia of *A. alternata* isolate 34-016 grown in revised tobacco medium for 3–20 days. Harvested mycelia were frozen and ground to powder in liquid nitrogen. RNA extraction was carried out by the method of Chomczynski and Sacchi [23] using 1 ml Trizol (Life Technologies) per 30 mg of mycelia. RNAs were fractionated in 1.2% formaldehyde-agarose gels and transferred to nylon membrane. Labeled rb51 probe was hybridized to the blot overnight at 50°C. Washes were for 2×5 min with (2×SSC, 0.1% SDS) at room temperature and 2×10 min using (0.1×SSC, 0.1% SDS) at 68°C. Further processing of these blots were carried out as described above.

#### *Recombinant Alt A1 Expression and Immunoblot Analyses*

Plasmid pCR II containing the rb51 insert was cleaved with *NheI* and *XbaI* and subcloned directly into *AvrII*-cleaved and dephosphorylated *P. pastoris* expression vector pPIC9. The orientation of inserts from individual transformants in pPIC9 was established by restriction digest and correctly oriented constructs were linearized with *DraI* prior to transfection of spheroplasts.

Spheroplasts were transformed with 10 µg of linearized construct and plated as described in protocols provided with the *Pichia* Expression Kit (Invitrogen). Transformants appeared 4–6 days after plating and incubating at 30°C. Individual colonies (12 for each or 4 trials) were replica-plated on MD [2% agar, 1% glucose, 1.34% yeast nitrogen base (YNB), 0.00004% biotin] and MM (2% agar, 0.5% methanol, 1.34% YNB, 0.00004% biotin) plates to establish methanol utilization (mut) phenotypes of the colonies. To assess levels of expression of Alt a 1 in individual transformants, 3-ml cultures were grown in BMGY (1% yeast extract, 2% peptone, 0.01 M potassium phosphate pH 6.0, 1% glycerol, 1.34% YNB, 0.00004% biotin) and cells were pelleted and transferred to BMMY (BMGY containing 0.5% methanol instead of glycerol) 1 day later. After 2 days of growth at 30°C, aliquots of culture were removed and the cells were pelleted. A nitrocellulose membrane was spotted with 3 µl of each culture supernatant and allowed to dry. The dot blot was then immunostained using rabbit IgG antibodies against partially purified Alt a 1 as described previously. Large cultures of a mut+ transformant expressing high levels of Alt a 1 were grown and induced in liquid MM medium supplemented with 1% casamino acids and 2.5% methanol. Culture medium was concentrated using an Amicon ultrafiltration unit with YM10 membrane, diluted with sterile water, and concentrated again. After centrifugation at 30,000 g for 20 min and passage through a 0.2 µm syringe filter, the supernatant was used in immunoblotting studies. The resulting solution was concentrated ≈40-fold from the original culture medium.

Atopic sera from *A. alternata*-sensitive individuals were obtained locally; additional plasma samples were obtained commercially (PlasmaLab International). All sera scored 19% or greater in RAST assays carried out in our laboratory. Total protein in concentrated conditioned medium from *P. pastoris* Alt a 1 transformant and *A. alternata* 34-016 crude extract was assayed using the method of Bradford [24]. Samples

corresponding to 500 µg of extract protein or 27 µg of conditioned media protein were loaded into preparative 12% SDS-PAGE gels [25] and blotted onto 0.2-µm nitrocellulose membrane in Towbin transfer buffer [26] containing 15% methanol. After blocking with 1% BSA in TBS, blots were screened in a Bio-Rad manifold. Antisera or plasma were applied to blots at 1:4 in TBST and incubated overnight at room temperature. Immunoblots were washed with TBST for 3×5 min and incubated with AP-conjugated goat antihuman IgE at 1:1,000 dilution for 2 h, also at room temperature. Washing and staining with nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate were as described above.

#### *Sequence Analyses*

Analyses of hydrophobicity [27], secondary structure [28] and amino acid content were carried out using Prosis software (Hitachi). The BLAST algorithm was used for homology searches of GenBank and EMBL databases [29]. Analyses of sequence block motifs were carried out using PROSITE [30].

## **Results**

#### *Southern Analysis of PCR-Amplified Inserts from Immunopositive Clones*

High density immunoscreening of the *A. alternata*-specific cDNA library using rabbit antiserum against partially purified Alt a 1 yielded nineteen positive clones from a total of 305,000 pfu screened. The staining intensities of these positives varied considerably, suggesting that the clones represented either Alt a 1 clones of differing sizes, or sequences corresponding to more than one distinct gene product. Surprisingly, none of the positives reacted positively when immobilized plaques were immunostained with pooled human atopic IgE. This lack of IgE reactivity has been attributed to epitopic masking in nondenatured β-galactosidase fusion proteins of the λgt11 clones, since IgE binding of the fusion proteins has been demonstrated in immunoblots.

To select clones containing the Alt a 1 N-terminus, inserts from all nineteen positive clones were amplified by PCR, and Southern blots of the PCR products were probed using digoxigenin-labeled ICHA7 oligonucleotide, a 23-mer with 8-fold degeneracy designed from the Alt a 1 N-terminus [12]. This probe was hybridized to the PCR-amplified inserts at low stringency, owing to the size of the oligo and its degeneracy. As shown in figure 1, the sizes of the immunoselected inserts ranged from ≈570 to 1,880 bp, of which a large proportion were 800–900 bp in length. The majority of the clones bound the probe, although hybridization to clones rb25, rb51 and rb61 was more intense than to any of the others. In contrast, clones rb31, rb82 and rb53 were unable to bind the probe, likely because they were too short (570–640 bp) to contain Alt a 1 N-terminal peptide

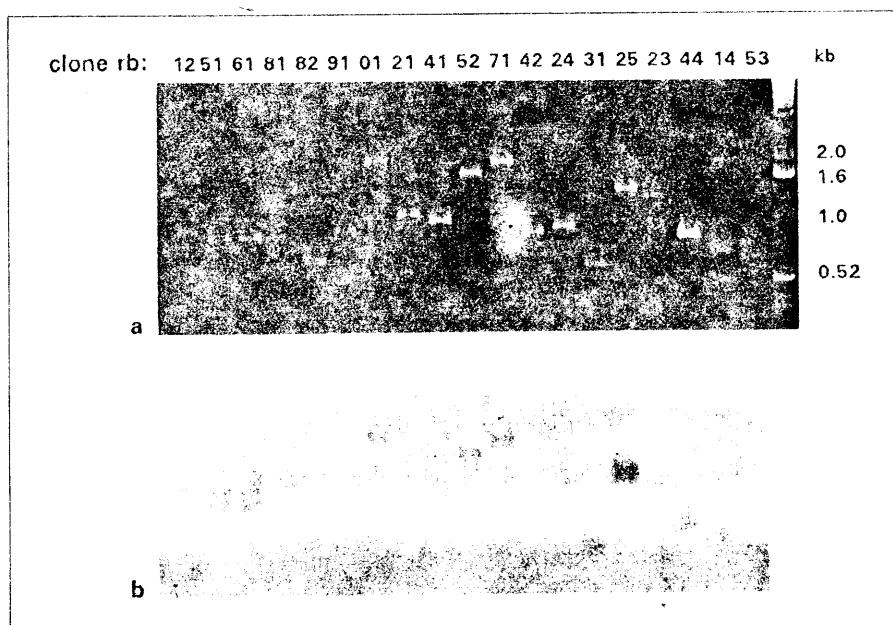


**Fig. 1.** Southern hybridization of Alt a 1 N-terminal oligonucleotide to PCR products from cDNA clones selected using Alt a 1-specific rabbit IgG. PCR products were run in 0.8% agarose gels (**a**), transferred to nylon membrane, and probed with a degenerate oligonucleotide (ICHA7, fig. 2) deduced from the Alt a 1 N-terminal peptide sequence (**b**). Positions of markers of known size are indicated in kilobase pairs.

sequences. Strongly immunoreactive clones rb91, rb24, rb21 and rb81 also failed to bind the probe despite their large sizes (840, 860, 1030 and 1060 bp, respectively). These clones may have been immunoselected by rabbit IgG specific for other proteins in the partially purified Alt a 1 preparation.

#### *DNA Sequencing of Immunopositive Clone rb51*

On the basis of its relatively strong hybridization with ICHA7 probe, clone rb51 was selected for DNA sequencing. The 660-bp insert contains an open reading frame (ORF) of 474 bp, from positions 64 to 534 (fig. 2). A sequence homologous to the oligonucleotide ICHA7, which encodes amino acids near the Alt a 1 N-terminus, is located from positions 166 to 188. The 5' terminal insert sequence of a related clone (rb42) is also shown to demonstrate that both clones encode the same gene product, differing only in 5' untranslated sequences that bridge  $\lambda$ gt11  $\beta$ -galactosidase sequences to the ORF. Searches of gene banks using the BLAST algorithm [29] failed to reveal significant sequence identity with genes encoding known proteins or allergens. Of other clones that have been partially sequenced, ten contain sequences that are common to rb51 (rb01, rb12, rb14, rb23, rb25, rb41, rb42, rb61 and rb82). Three clones (rb01, rb52 and rb71) are known to contain rRNA sequences, which likely accounts for their increased length. Seven clones have not yet been investigated (including six that did not hybridize to ICHA7 oligo). Therefore, we conclude that

most of the immunopositive clones obtained contain some, if not the entire coding sequences of Alt a 1.

#### *Translation of rb51 ORF and Comparison with Alt a 1 N-Terminal Peptide Sequences*

The translated rb51 ORF encodes a peptide of 157 amino acids, with a calculated  $M_r$  of 16,960 (fig. 3). Comparison of amino acid analyses between rb51 ORF and our Alt a 1 isolated [12] show considerable similarity (table 1), except for the amino acid proline. However, the hydrolysis conditions used may have accounted for the overestimation of proline in the amino acid analysis of Alt a 1 [12]. The predicted isoelectric point of rb51 is within 0.5 units of purified Alt a 1 (table 1). The expected N-terminal peptide sequence of Alt a 1 is located at residues 29–48 of the rb51 ORF, and matches the sequence of Curran et al. [12] (NTERM2; fig. 3) at 17 of 20 positions. Moreover, this ORF also matches the N-terminus of Alt a 1 isolated by Matthiesen et al. [14; Lowenstein, pers. commun.] at 24 of 26 positions (fig. 3), differing only at positions 30 and 40, residues whose identities remained unresolved in the latter protein. Variations in amino acid sequence between our Alt a 1 protein isolated and the translated ORF of clone rb51 are also found at these positions, and at position 29, indicating that either *A. alternata* cells express multiple isoforms of Alt a 1, or that slightly different polymorphic forms of Alt a 1 may exist in *A. alternata* isolates.

### Northern Analysis of *rb51* Transcripts in *A. alternata* Total RNA

To assess the size of *rb51* transcripts in *A. alternata*, total RNA was extracted from mycelial samples taken at various times postinoculation and analyzed by Northern blot probed with DIG-labeled *rb51*. Total RNA was extracted with little or no degradation, as is evident from the integrity of the rRNA bands in figure 4a. However, we consistently observed a substantial proportion of insoluble nucleic acid in samples from 13-, 17- and 20-day cultures, likely caused by the presence of contaminating DNA. The transcript encoding *rb51* is shown as an abundant transcript of  $\approx 0.7$  kb; the similar sizes of the transcript and *rb51* clone suggest that the ORF obtained represents a full length Alt a 1 subunit. Peak levels of *rb51* transcripts were observed in 5-day samples, which decreased rapidly from 7 to 10 days post-inoculation.

### Hydrophobicity and Secondary Structure Prediction of *rb51* ORF

Hydrophobicity and secondary structure predictions of the *rb51* ORF were carried out using the methods of Kyte and Doolittle [27], and Chou and Fasman [28]. The most hydrophobic region in the *rb51* ORF is located at the N-terminus, from amino acids 10–23 [data not shown]. This region coincides with the longest  $\alpha$ -helical domain observed in the secondary structure prediction illustrated in figure 5. As amino acids 1–28 were not previously detected in microsequenced Alt a 1 N-terminal peptides, it is likely that the first 28 amino acids represent a leader sequence that is cleaved from the protein during processing. The remaining sequence of the protein of short stretches of  $\alpha$ -helical or  $\beta$ -sheet-like domains that alternate with amino acids specifying turns. Analyses of sequence block motifs using PROSITE [30] determined that no potential N-glycosylation sites are present in the *rb51* ORF.

### Expression and IgE-Binding Capability of Recombinant Alt a 1 Protein in Immunoblots

To express the Alt a 1 protein in large quantity, *rb51* sequences were cleaved with *NheI* and *XbaI* and ligated into pPIC9, a *P. pastoris*-secreted expression vector. This vector features a *Saccharomyces cerevisiae*  $\alpha$ -factor prepro peptide sequence located upstream from its multiple cloning site. It was anticipated that the insertion of *rb51* into the *AvrII* site in this vector would be in the appropriate reading frame for correct expression of Alt a 1, although the expressed protein might contain as many as 20 additional N-terminal amino acids followed by the Alt a 1 N-terminal signal peptide. Spheroplasts were transformed with this construct, and transformants were screened for their methanol

**Table 1.** Amino acid composition and pI of Alt a 1 and translated *rb51* ORF

| Amino acid            | Alt a 1 <sup>a</sup><br>residues 100 | Rb51<br>residues 100 |
|-----------------------|--------------------------------------|----------------------|
| Aspartate, asparagine | 12.18                                | 11.4                 |
| Threonine             | 9.78                                 | 8.92                 |
| Serine                | 8.64                                 | 8.92                 |
| Glutamate, glutamine  | 7.75                                 | 7.00                 |
| Proline               | 7.20                                 | 3.82                 |
| Glycine               | 8.77                                 | 8.28                 |
| Alanine               | 9.73                                 | 10.83                |
| Cystine               | 1.01                                 | NA                   |
| Cysteine              | NA                                   | 3.18                 |
| Valine                | 6.03                                 | 4.46                 |
| Methionine            | 0.65                                 | 1.27                 |
| Isoleucine            | 3.03                                 | 3.18                 |
| Leucine               | 7.71                                 | 7.64                 |
| Tyrosine              | 4.53                                 | 5.10                 |
| Phenylalanine         | 5.37                                 | 5.73                 |
| Histidine             | 0.51                                 | 0.64                 |
| Lysine                | 5.44                                 | 5.73                 |
| Arginine              | 2.00                                 | 2.55                 |
| Tryptophan            | ND                                   | 1.27                 |

NA = Not applicable; ND = not done.

<sup>a</sup> Data obtained from Curran et al [12].

utilization phenotype and for methanol-induced production of secreted protein reactive with rabbit IgG against Alt a 1. A single transformant of mut<sup>+</sup> phenotype was isolated that secreted Alt a 1 protein into conditioned medium. Bands of  $M_r$  15–20,000 were not observed in rabbit anti-Alt a 1-probed immunoblots of media protein from methanol-induced non-transformed, or vector pPIC9-transformed *P. pastoris* [data not shown]. Aliquots of concentrate (1  $\mu$ g) were compared with an equal amount of natural allergen that was extracted from *A. alternata* and further purified by ultrafiltration and gel filtration in Sephadex G-100. The recombinant and natural Alt a 1 were compared under both nonreducing and reducing conditions in immunoblots. Surprisingly, the recombinant Alt a 1 subunit associated as a dimer in culture medium, as shown by the 29-kDa band in nonreducing immunoblots (fig. 6, top row). Under reducing conditions, the recombinant Alt a 1 migrated as a doublet of bands with slightly less mobility than those of the natural allergen. The ten N-terminal amino acids of the two recombinant Alt a 1 bands observed in reducing SDS-PAGE were analyzed; the sequences obtained revealed that the slower-migrating band begins at amino acid 22 (sequence obtained: ESRQDT-ASXP), whereas the faster-migrating begins at amino acid

|           |      |            |            |                |                   |            |      |
|-----------|------|------------|------------|----------------|-------------------|------------|------|
|           |      | 10         | 20         | 30             | 40                | 50         |      |
| RB51.SEQ  | 1    | GTACGAGCT  | ATTGAGCGTA | ASCGTCGCT      | AGCITGGTTG        | AATCTCAAT  | 50   |
| RB42.SEQ  | 1    | -----**    | -----*     | -----C         | TT*ACCC*CA        | **C*****   | 50   |
| ICHA7.SEQ | -165 | -----      | -----      | -----          | -----             | -----      | -116 |
|           |      | 60         | 70         | 80             | 90                | 100        |      |
| RB51.SEQ  | 51   | CAACTTCATC | AAAATGCACT | TCACCAACAT     | CGCCTCTCTC        | TTCCGCGCGG | 100  |
| RB42.SEQ  | 51   | *****      | *****      | *****          | *****             | *****      | 100  |
| ICHA7.SEQ | -115 | -----      | -----      | -----          | -----             | -----      | -66  |
|           |      | 110        | 120        | 130            | 140               | 150        |      |
| RB51.SEQ  | 101  | CTGGCCTTGC | CGCTGCTGCT | CCTCTCGAGT     | CTCGCCAGGA        | CACCGCATCC | 150  |
| RB42.SEQ  | 101  | *****      | *****      | *****          | *****             | *****      | 150  |
| ICHA7.SEQ | -65  | -----      | -----      | -----          | -----             | -----      | -16  |
|           |      | 160        | 170        | 180            | 190               | 200        |      |
| RB51.SEQ  | 151  | TGCCCTGTCA | CCACTGAGGG | TGACTACGTC     | TGGAAGATCT        | CGGAATTCTA | 200  |
| RB42.SEQ  | 151  | *****      | *****      | *****          | *****             | *****      | 200  |
| ICHA7.SEQ | -15  | -----      | -----**A** | -----***Y****Y | -----GT*****      | -----      | 35   |
|           |      | 210        | 220        | 230            | 240               | 250        |      |
| RB51.SEQ  | 201  | CGGACGCAAG | CCGGAAGGAA | CCTACTACAA     | CAGCCTCGGC        | TTCAACATCA | 250  |
| RB42.SEQ  | 201  | *****      | *****      | *****          | *****             | *****      | 250  |
|           |      | 260        | 270        | 280            | 290               | 300        |      |
| RB51.SEQ  | 251  | AGGCCACCAA | CGGAGGAACC | CTGACTTCA      | CCTGCTCTGC        | TCAGGCCGAT | 300  |
| RB42.SEQ  | 251  | *****      | ***G**     | -----          | -----             | -----      | 300  |
|           |      | 310        | 320        | 330            | 340               | 350        |      |
| RB51.SEQ  | 301  | AAGCTTGAGG | ACCACAAGTG | GTACTCTGT      | GGCGAGAACA        | GCTTCATGGA | 350  |
|           |      | 360        | 370        | 380            | 390               | 400        |      |
| RB51.SEQ  | 351  | CTTCTCTTTC | GACAGCGACC | GCAGCGGTCT     | GCTCTGAAG         | CAGAAGGTAA | 400  |
|           |      | 410        | 420        | 430            | 440               | 450        |      |
| RB51.SEQ  | 401  | GCGACGACAT | CACCTATGTC | GCTACCGCCA     | CTCTTCCCAA        | CTACTGCCGC | 450  |
|           |      | 460        | 470        | 480            | 490               | 500        |      |
| RB51.SEQ  | 451  | GCTGGCGGTA | ACGGCCCTAA | GGACTTTGTC     | TGCCAGGGTG        | TTGCCGACGC | 500  |
|           |      | 510        | 520        | 530            | 540               | 550        |      |
| RB51.SEQ  | 501  | CTACATCACC | CTCGTSACTC | TCCCAAGAG      | CTCTTAAGCG        | ATAGTTGGTC | 550  |
|           |      | 560        | 570        | 580            | 590               | 600        |      |
| RB51.SEQ  | 551  | GAGCATATAA | GAAGCATCAA | CTTCTCTTAC     | TGTAATTTT         | TTATTAGCGC | 600  |
|           |      | 610        | 620        | 630            | 640               | 650        |      |
| RB51.SEQ  | 601  | ATTCAAGATA | GACATTCCCT | ATTCTAGACT     | <u>TAAATTAACC</u> | ATTATCTGA  | 650  |
|           |      | 660        | 670        | 680            | 690               | 700        |      |
| RB51.SEQ  | 651  | TCAAAAAAAA | -----      | -----          | -----             | -----      | 700  |

**Fig. 2.** Nucleotide sequence of cDNA clones rb51, rb42 and oligonucleotide ICHA7. Initiation (ATG) and termination (TAA) codons are double-underlined. The polyadenylation signal (or closest approximation) is underlined. Asterisks denote bases that match those shown in clone rb51.

**Fig. 3.** Peptide sequence of rb51 ORF and comparison with previously described N-terminal peptide sequences. NTERM1.AMI and NTERM2.AMI represent alternative possible Alt 1 N-terminal peptide sequences as determined by Curran et al. [12]. NTERM.M.AMI represents Alt 1 N-terminal peptide sequences as determined by Matthiesen et al. [14; Lowenstein, pers. commun.]. Asterisks denote amino acids matching the deduced peptide sequence of rb51 ORF (RB51AA).

|             |     |            |            |            |            |            |     |
|-------------|-----|------------|------------|------------|------------|------------|-----|
|             |     | 10         | 20         | 30         | 40         | 50         |     |
| RB51.AA     | 1   | MQFTTJASLF | AAAGLAAAP  | LESROQTASC | PVTTEGDYVW | KISEFYGRKP | 50  |
| NTERM1.AMI  | -28 | -----      | -----      | -----AD    | *****V     | *****F     | 22  |
| NTERM2.AMI  | -28 | -----      | -----      | -----AD    | *****V     | *****F     | 22  |
| NTERM.M.AMI | -28 | -----      | -----      | -----*X    | *****X     | *****F     | 22  |
|             |     | 60         | 70         | 80         | 90         | 100        |     |
| RB51.AA     | 51  | EGTYYNLSGF | NIKATNGGTL | DFTCSAQADK | LEDHKWYSCG | ENSMDFDFFD | 100 |
| NTERM.M.AMI | 23  | ****       | -----      | -----      | -----      | -----      | 100 |
|             |     | 110        | 120        | 130        | 140        | 150        |     |
| RB51.AA     | 101 | SDRSGLLLKQ | KVSDDITYVA | TATLPNYCRA | GGNGPKDFVC | QGVADAYITL | 150 |
|             |     | 157        | -----      | -----      | -----      | -----      |     |
| RB51.AA     | 151 | VTLPKSS*   | -----      | -----      | -----      | -----      | 200 |



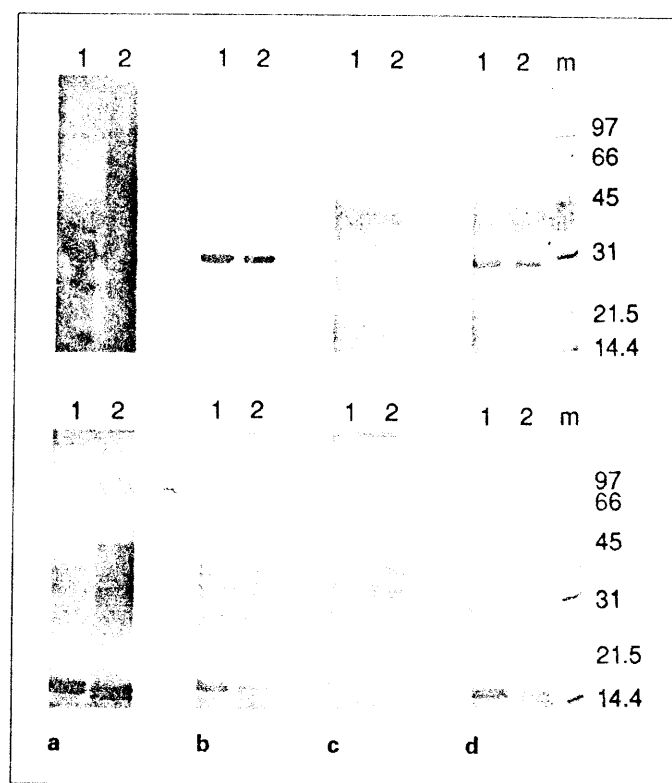


**Fig. 6.** Comparison of recombinant (lanes 1) and *A. alternata*-derived (lanes 2) Alt a 1 in immunoblots. Natural Alt a 1 allergen was purified by ultrafiltration and gel filtration on Sephadex G-100 [12]. Recombinant allergen was prepared as described in Material and Methods. Aliquots of 1  $\mu$ g were run in 10% (nonreducing, top row) or 12% (reducing, bottom row) SDS-PAGE, transferred to PVDF membrane, and immunostained with rabbit IgG against partially purified Alt a 1 (a), pooled human atopic IgE from *A. alternata*-sensitive individuals (b), or nonatopic human serum pool (c), or stained with Coomassie brilliant blue R-250 (d). Positions of  $M_r$  standards of known size (m) are indicated in thousands.

29K [12–15]. Using rabbit IgG antibodies against partially purified Alt a 1, we immunoselected a 660-bp clone (rb51), one of nineteen clones immunoselected from an *A. alternata*-specific cDNA library constructed in  $\lambda$ gt11. Northern analyses identified a transcript of  $\approx 0.7$  kb that hybridizes with labeled rb51 probe.

An in-frame ORF within the rb51 cloned insert encodes a protein with calculated  $M_r$  of 16,960 and pI of 4.59. Its peptide sequence matches previously sequenced N-termini of two independent Alt a 1 isolates at 17 of 20, and 24 of 26 positions. These findings indicate that clone rb51 encodes an Alt a 1 isoform. Despite a lack of sequence data, it is possible that Alt a Bd 29K [15] and Alt a I<sub>1563</sub> [11, 31], disulfide-linked dimers of 29–31 kDa with pIs of 4.2, may represent the same protein. However, the subunit sizes of Alt a

I<sub>1563</sub> are 19 and 20 kDa, considerably larger than those of Alt a 1, and this protein has been stated as containing 20% carbohydrate [31]. Analysis of the translated rb51 ORF using PROSITE failed to reveal potential N-linked glycosylation sites. Although there is no obvious consensus sequence for isolated O-glycosylation sites [32], proline at –1 relative to the serine or threonine acceptor, or at –6 or +3 in the absence of a charged residue at –1, has been identified as a potential, but not absolute predictor of O-linked glycosylation [32, 33]. An examination of rb51 ORF reveals 3 sites that fulfill these criteria, with acceptors at residues 44, 53 and 157 (fig. 3). However, previous attempts to deglycosylate Alt a 1 in laboratory, either enzymatically with endoglycosidase F or O-glycosidase or chemically using Trifluoromethanesulphonic Acid (TFMS), invariably failed to reduce the  $M_r$  of the protein [12], indicating that Alt a 1 is nonglycosylated. We cannot account for the observations of Paris et al. [31] suggesting that Alt a I<sub>1563</sub> is glycosylated, unless their protein is a variant form that contains a small number of glycosylation consensus sites. One other recently sequenced clone has been reported to contain the Alt a 1 N-terminus; however, this PCR-amplified sequence encodes a peptide of  $M_r$  9,500 and pI of 7.00, and is highly divergent from Alt a 1 in sequence [34]. The sequence of this clone is also distinct from the three *A. alternata* allergens recently cloned by Achatz et al. [35], and exhibits no significant homology with other known allergens or proteins in the GenBank/EMBL databases.

Differences between sequenced N-termini of various Alt a 1 isolates and homologous sequences within rb51 are confined to specific sites (residues 29, 30, 40 and 46 in rb51) [12, 14]. These ambiguities could represent amino acid positions which differ among closely related isoforms of Alt a 1. Isoelectric focusing of Alt a 1 has revealed a major component with a pI of 4.15, and minor components with pIs of 4.25 and 4.40 [12]. Minor differences between the amino acid distributions of Alt a 1 and the translated rb51 ORF may also reflect the presence of more than one isoform in the sample subjected to amino acid analysis [12]. However, if multiple isoforms of Alt a 1 were expressed in *A. alternata*, cDNAs to such isoforms should also have been selected as distinct clones using Alt a 1-specific rabbit IgG. A comparison of three clones (rb51, rb61, rb42) whose DNA sequences at this site are known failed to reveal differences that could have accounted for the amino acid alterations expected at these specific positions. It remains possible that other clones may contain isoform-specific sequence deviations at these or other sites.

Analyses of the translated rb51 ORF have revealed additional information regarding the structure of the Alt a 1 gene product. The N-terminal 28 amino acids of the translated

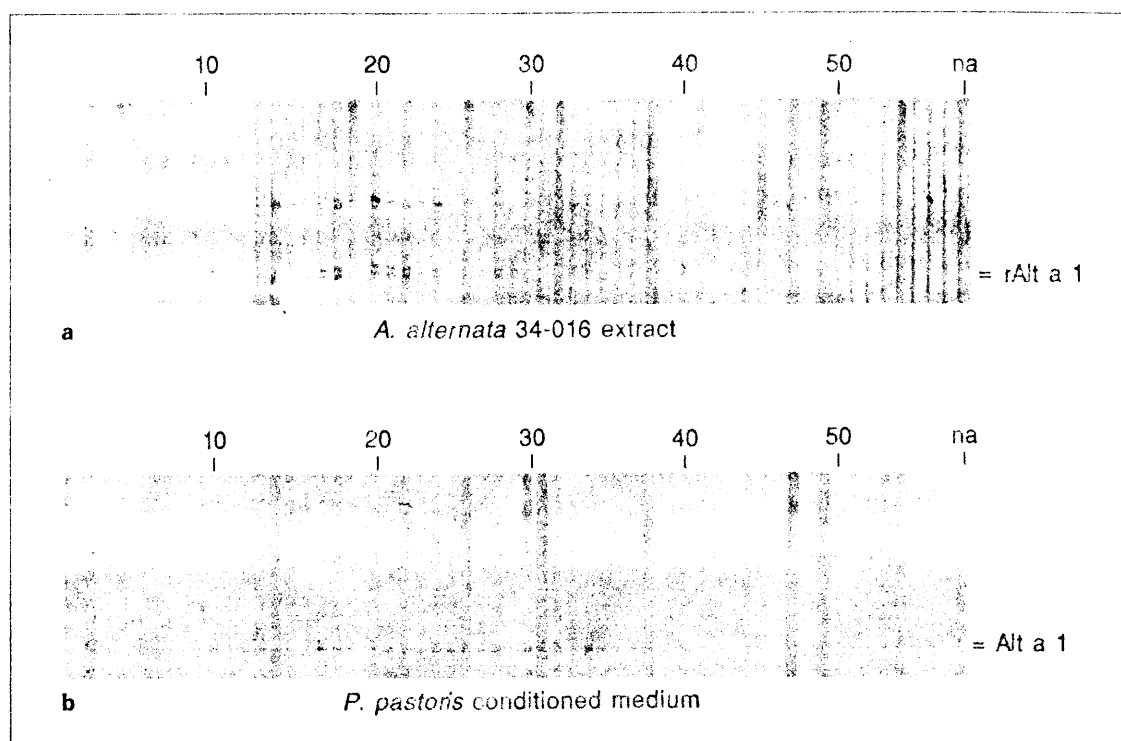


**Fig. 7.** Reactivity of serum IgEs from *A. alternata*-sensitive individuals to *A. alternata*-derived (a), or recombinant (b) Alt a 1. Preparative immunoblots of crude extract or concentrated conditioned medium were screened with 57 atopic IgEs (lanes 1–57) or nonatopic pooled serum (lane na) as described in Material and Methods.

ORF, a highly hydrophobic and strongly  $\alpha$ -helical domain, appears to serve as a cleaved signal peptide that directs Alt a 1 for export from the cell, as significant amounts of Alt a 1 are secreted into culture broth of *A. alternata* [12]. Although rb51 is lacking a signal cleavage motif common at the C-terminus of both bacterial and eukaryotic signal peptide sequences (Ala-Gly-Ala), signal peptide sequence requirements for functional cleavage in eukaryotes tend to be more relaxed than in bacterial systems [36]. The Alt a 1 N-terminal signal peptide also shares 41% homology over the first 22 amino acids with signal peptides of *Neurospora crassa* BLI-7 (also known as CCG-2) and *Aspergillus fumigatus* RODA genes that encode hydrophobins, secreted hydrophobic proteins implicated in the formation of conidial rodlets [37–39]. However, Alt a 1 hydrophobins are highly dissimilar in sequence and structure other than their limited signal peptide homology.

We have also shown that recombinant Alt a 1 secreted into the media of *P. pastoris* associates as a dimer, similar to natural allergen purified from *A. alternata* culture medium or mycelia. Under reducing conditions in SDS-PAGE, recombinant Alt a 1 appears as two distinct bands, even

though both are translated from the same Alt a 1 sequences. Sequencing of the peptide bands indicates that the difference between the two N-termini is only three amino acids not enough to account for differences in mobility observed between the bands. This observation parallels that of the natural allergen, whose subunits also migrate differently in reducing SDS-PAGE, even though their N-termini are nearly identical. These findings imply that Alt a 1 may be a homodimer of differentially processed subunits, that posttranslational processing of one or both members of the recombinant dimer would probably occur during or following their association as a dimer, and that such processing may be similar to that which occurs in the natural allergen. These processing functions may be intrinsic to the interacting subunits, or carried out by other secreted proteins such as extracellular proteases, although *P. pastoris* carries out relatively little protein secretion of its own and probably secretes a significantly different profile of proteases than would *A. alternata*. It is interesting to note that *P. pastoris* appears to be able to cleave the *A. alternata*-derived Alt a 1 N-terminal signal peptide, even though cleavage of the two recombinant Alt a 1 bands appears to be at different sites from the expected

cleavage site of the natural allergen at position 28. This has resulted in production of a recombinant allergen that is more similar to the natural allergen in terms of size and structure than was previously expected.

Serum IgEs from 82% of *A. alternata*-sensitive individuals bound *Pichia*-expressed Alt a 1 in immunoblots, confirming this protein as a major allergen in *A. alternata* and demonstrating the potential benefit of recombinant Alt a 1 as an immunodiagnostic reagent. The ease with which relatively pure Alt a 1 has been obtained will facilitate the development of a specific enzyme-linked immunosorbent assay or monoclonal antibodies. Knowledge of the Alt a 1 sequence will now permit localization of antigenic and allergenic sites on Alt a 1. Although Alt a 1 was first isolated as

an allergen that was capable of inhibiting high RAST scores without eliciting a PCA response [17], we have recently shown that sera from mice immunized with KLH-conjugated Alt a 1 N-terminal 20mer peptide can elicit a PCA response upon challenge with BSA-conjugated peptide [40]. Consequently, precise mapping of B and T-cell epitopes, including those that generate Th1 and Th2-responses, remains a priority.

## Acknowledgements

The authors thank Dr. David Watson and Dr. Makoto Yaguchi for recombinant Alt a 1 N-terminal sequencing. This study was supported by the National Biotechnology Strategy (Canada).

## References

- Gergen PJ, Turkeltaub PC, Kovar MG: The prevalence of allergic skin test reactivity to eight common aeroallergens in the US population: Results from the second National Health and Nutrition Examination Survey. *J Allergy Clin Immunol* 1987;80:669-679.
- Aas K, Leegard J, Aukrust L, Grimmer O: Immediate type hypersensitivity to common molds. *Allergy* 1980;35:443-451.
- Yee EG, Bahna SL: Skin testing and RAST to molds in patients with respiratory allergy. *J Allergy Clin Immunol* 1986;77:200.
- Yunginger JW, Jones RT, Nesheim ME, Geller M: Studies of *Alternaria* allergens. III. Isolation of a major allergenic fraction [ALT-1]. *J Allergy Clin Immunol* 1980;66:138-147.
- Vijay HM, Huang H, Laverne G, Young NM, Bernstein IL: Studies on *Alternaria* allergens. II. Presence of two related antigens with contrasting allergenic properties in *Alternaria tenuis* extracts. *Int Arch Allergy Appl Immunol* 1981;65:410-416.
- Nyholm H, Lowenstein H: Identification and characterization of allergens from *Alternaria alternata*. *Allergy* 1982;37(suppl 1):40.
- Vijay HM, Burton M, Young NM: The hypoallergen of *Alternaria alternata* carries a dominant epitope of the major allergen; in Schon AH, Kraft D, Kunkel G (eds): Epitopes of Atopic Allergens. Proc Workshop 14th Congr Eur Acad Allergol Clin Immunol, Berlin, 1989, Brussels, UCB Institute of Allergy, 1990, pp 14-17.
- Budd T, Kuo CY, Yoo T, McKenna W, Cazin J: Antigens of *Alternaria*. Isolation and partial characterization of basic peptide allergen. *J Allergy Clin Immunol* 1983;71:277-282.
- Kroutil L, Bush RK: Detection of *Alternaria* allergens by Western blotting. *J Allergy Clin Immunol* 1985;75:174.
- Bengtsson A, Karlsson Å, Rolfsen W, Einarsson R: Detection of allergens in mould and mite preparations by a nitrocellulose electroblotting technique. *Int Arch Allergy Appl Immunol* 1986;80:383-390.
- Paris S, Fitting C, Latgé JP, Herman D, Guinépain MT, David B: Comparison of conidial and mycelial allergens of *Alternaria alternata*. *Int Arch Allergy Appl Immunol* 1990;92:1-8.
- Curran IHA, Young NM, Burton M, Vijay HM: Purification and characterization of Alt a-29 from *Alternaria alternata*. *Int Arch Allergy Immunol* 1993;102:267-275.
- Matthiesen F, Larsen L, Lowenstein H: Purification and characterization of two important mold allergens, Alt a 1 and Cla h 1 (abstract 21). 14th Congr Eur Acad Allergol Clin Immunol Allergol, Berlin, 1989.
- Matthiesen F, Olsen M, Lowenstein H: Purification and partial sequencing of the major allergen of *Alternaria alternata*. *J Allergy Clin Immunol* 1992;89:241.
- Deards MJ, Montague AE: Purification and characterization of major allergen of *Alternaria alternata*. *Mol Immunol* 1991;28:409-415.
- Vijay HM, Huang H, Young NM, Bernstein IL: Studies on *Alternaria* allergens. I. Isolation of allergens and *Alternaria tenuis* and *Alternaria solani*. *Int Arch Allergy Appl Immunol* 1979;60:229-239.
- Vijay HM, Huang H, Laverne G, Young NM, Bernstein IL: Studies on *Alternaria* allergens. II. Presence of two related antigens with contrasting allergenic properties in *Alternaria tenuis* extracts. *Int Arch Allergy Appl Immunol* 1979;65:410-416.
- IUIS WHO Allergen Nomenclature Subcommittee: Allergen Nomenclature. *Bull WHO* 1995;72:797-806.
- Solvaggio JE, Burge HA, Chapman JA: Emerging concepts in mold allergy: What is the role of immunotherapy? *J Allergy Clin Immunol* 1993;92:217-222.
- Vijay HM, Huang H, Young NM, Bernstein IL: Studies on *Alternaria* allergens. IV. Comparative biochemical and immunological studies of commercial *Alternaria tenuis* batches. *Int Arch Allergy Appl Immunol* 1984;74:256-261.
- Steringer I, Aukrust L, Einarsson R: Variability of antigenicity and allergenicity in different strains of *Alternaria alternata*. *Int Arch Allergy Appl Immunol* 1987;84:190-197.
- Huynh TV, Young RA, Davis RW: Constructing and screening cDNA libraries in  $\lambda$ gt10 and  $\lambda$ gt11; in Glover D (ed): DNA Cloning: A Practical Approach. Oxford, IRL Press, 1985, vol 1, pp 49-78.
- Chomczynski P, Sacchi N: Single-step method of RNA extraction by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 1987;162:156-159.
- Bradford M: A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye-binding. *Anal Biochem* 1976;72:248-254.
- Laemmli UK: Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature* 1970;227:680-685.
- Towbin H, Staehelin T, Gordon J: Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc Natl Acad Sci USA* 1979;76:4350-4354.
- Kyte J, Doolittle RF: A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 1982;157:105-132.
- Chou PY, Fasman GD: Prediction of protein conformation. *Biochemistry* 1974;13:222-245.

- 29 Altschul SE, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. *J Mol Biol* 1990;215:403-410.
- 30 Bairoch A: The PROSITE dictionary of sites and patterns in proteins, its current status. *Nucleic Acids Res* 1993;21:3097-3103.
- 31 Paris S, Debeaupuis JP, Prévost MC, Cossetto M, Latgé JP: The 31 kDa major allergen Alt a I<sub>1</sub> of *Alternaria alternata*. *J Allergy Clin Immunol* 1991;88:902-908.
- 32 Wilson IBH, Gavel Y, von Heijne G: Amino acid distributions around O-linked glycosylation sites. *Biochem J* 1991;275:529-534.
- 33 O'Connell B, Tabak LA, Ramasubbu N: The influence of flanking sequences on O-glycosylation. *Biochem Biophys Res Commun* 1991;180:1024-1030.
- 34 Hu F, Pacheco F, Landuyt J, Rosenthal D, Barnes C, Portnoy J: IgE binding to a synthetic recombinant peptide containing the N-terminal sequence of Alt a I. *J Allergy Clin Immunol* 1995;95:2166.
- 35 Achatz G, Oberkofler H, Lechenauer E, Simon B, Unger A, Kandler D, Ebner C, Prillinger H, Kraft D, Breitenbach M: Molecular cloning of major and minor allergens of *Alternaria alternata* and *Cladosporium herbarum*. *Mol Immunol* 1995;32:213-227.
- 36 Punsley AP: Protein Targeting. San Diego, Academic Press, 1989, pp 45-111.
- 37 Fherle J, Russo VE: *Neurospora crassa* blue-light-inducible gene bli-7 encodes a short hydrophobic protein. *DNA Seq* 1992;3:131-141.
- 38 Beli-Pederson Q, Dunlap JC, Loros JJ: The *Neurospora* circadian clock-controlled gene, ccc-2, is allelic to eas and encodes a fungal hydrophobin required for formation of the conidial rodlet. *Genes Dev* 1992;6:2382-2394.
- 39 Thau N, Monod m, Crestan B, Rolland C, Tronchin G, Latgé JP, Paris S: Rodletless mutants of *Aspergillus fumigatus*. *Infect Immun* 1994;62:4380-4388.
- 40 Zhang L, Curran IHA, Muradia G, De Vouge MW, Rode H, Vijay HM: N-terminus of a major allergen, Alt a I, of *Alternaria alternata* defined to be an epitope. *Int Arch Allergy Immunol* 1995;108:254-259.