

## A Recombinant Group 1 House Dust Mite Allergen, rDer f 1, with Biological Activities Similar to Those of the Native Allergen

Elaine A. Best,<sup>1</sup> Kim E. Stedman, Cynthia M. Bozic, Shirley Wu Hunter, Lisa Vailes,\* Martin D. Chapman,\* Catherine A. McCall, and Martin J. McDermott

Heska Corporation, Fort Collins, Colorado 80525; and \*Asthma and Allergic Diseases Center, Box 225, Department of Medicine, University of Virginia, Charlottesville, Virginia 22908

Received June 16, 2000, and in revised form August 14, 2000

**Serum IgE directed against Der f 1, a protease found in the feces of *Dermatophagoides farinae*, correlates well with allergic sensitization to house dust mite in humans and is a risk factor for developing asthma. Native Der f 1 (nDer f 1) is produced as a pre-pro form and processed to an ~25-kDa mature form. We have expressed recombinant forms of Der f 1 (rDer f 1) in *Pichia pastoris* using AOX1-promoter expression vectors. Fusion of either the pro-enzyme form or the mature form to the *Saccharomyces cerevisiae*  $\alpha$  factor pre-pro sequence resulted in secretion of the mature form of the protein from *P. pastoris*. The secreted protein was heterogeneously glycosylated at a single N-glycosylation site and had an apparent molecular mass of 35–50 kDa. Both the  $\alpha$  factor signal peptide and the pro-enzyme region were efficiently processed during secretion. A version of the pro-enzyme with a mutated consensus N-linked glycosylation site was secreted from *P. pastoris* as a mature, unglycosylated, ~25-kDa protein. The IgE binding activity of this unglycosylated rDer f 1 was similar to that of glycosylated forms produced by *P. pastoris* and to nDer f 1 obtained from mites. Thus, oligosaccharides are not required for secretion from *P. pastoris* or for IgE binding *in vitro*. Recombinant and native versions of Der f 1 displayed protease activity on casein zymogram gels. The availability of a highly purified recombinant Der f 1 will facilitate experimental and clinical studies of mite allergy.** © 2000 Academic Press

Proteins found in high concentrations in the feces of house dust mites (1) (including *Dermatophagoides* spp., *Euroglyphus maynei*, *Blomia tropicalis*, and *Lepidoglyphus destructor*) are a contributing factor in atopic disease (perennial rhinitis, asthma, and atopic dermatitis) worldwide. Since the early 1900s, house dust mite (HDM)<sup>2</sup> allergy has been diagnosed and treated using crude extracts derived from mites. Although the composition and purity of allergen extracts are standardized for potency, the allergen contents of various extracts may differ by as much as sixfold (2). Accordingly, some patients may be misdiagnosed as nonallergic, and some may respond poorly to specific immunotherapy because they are treated with allergen extracts that contain an inappropriate level (either sub- or superoptimal amounts) of the relevant allergen.

During the last decade, significant progress has been made toward identifying those proteins in commercial allergen extracts that are relevant for diagnosis and treatment of HDM allergy. Although more than a dozen mite proteins have been shown to bind IgE in allergic human sera, over 80% of HDM allergic patients have IgE directed against a group of ~25- to 30-kDa allergens. These allergens have extensive homology (~80% identity) to each other and to the papain superfamily of cysteine proteases (actinidin, papain, cathepsin B, and cathepsin H) and are known collectively as group 1 HDM allergens (3). Hewitt and co-

<sup>2</sup> Abbreviations used: HDM, house dust mite; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; mAb, monoclonal antibody; AOX1, alcohol oxidase; MALDI-MS, matrix-assisted laser desorption ionization mass spectrometry; PFU, plaque-forming units; PBS, phosphate-buffered saline; TFA, trifluoroacetic acid; LB, Luria Berfani; YPD, yeast peptone dextrose; PVDF, polyvinylidene fluoride; PBS-T, phosphate-buffered saline-Tween 20; ABTS, 2,2'-azino-di-(3-ethylbenz-thiazoline-6-sulfonate).

<sup>1</sup> To whom correspondence should be addressed at Heska Corporation, 1613 Prospect Parkway, Ft. Collins, CO 80525. Fax: 970-493-7333. E-mail: beste@heska.com.

workers have speculated that the known enzymatic function of the group 1 allergens may play a role in their allergenicity (4, 5).

Although cDNA clones encoding group 1 allergens from *Dermatophagoides farinae* (Der f 1) (6) and *Dermatophagoides pteronyssinus* (Der p 1) (7) have been available for many years, attempts to produce biologically active recombinant versions of these allergens have been only moderately successful, most likely because these molecules have complex molecular features. These features include a 19-residue signal peptide, an 80-residue pro-enzyme sequence, a 222- to 223-residue mature protein, a potential N-linked glycosylation site, and three cysteine residue pairs assumed to be involved in disulfide bridges (3, 6).

Several investigators have attempted to produce recombinant versions of the group 1 HDM antigens. Recombinant allergens produced in *Escherichia coli* and *Saccharomyces cerevisiae* were less immunoreactive than the native allergen (8–10).

To date, the most promising system for producing a rDer f 1 is a baculovirus expression system developed by Shoji and coworkers (11, 12). These investigators constructed a recombinant baculovirus that included the baculovirus (*Autographica californica* nuclear polyhedrosis) polyhedrin promoter and a region encoding the 18 residue Der f 1 signal peptide, the 80 residue pro-sequence, and the entire mature Der f 1. The rDer f 1 was secreted in high yields (50 mg/L) from baculovirus-infected *Spodoptera frugiperda* Sf9 cells, but as a mixture of three forms: two pro-forms of 33 and 35 kDa with different amino termini, and a mature 28-kDa form. In tests using pooled human sera from HDM allergic patients, the pro-form had only 20% of the IgE reactivity of the native molecule. The IgE binding activity of the pro-form could be enhanced to a level comparable to the native antigen if the pro-peptide was removed in a postpurification protease digestion step. Although Shoji and coworkers' work (11, 12) is encouraging, their process and assay results have some limitations: (1) Der f 1 contains internal lysine residues. A nonspecific protease such as lysylendopeptidase cannot be easily controlled to cleave only at the engineered lysine outside the mature protein and not at any of the internal lysines in the mature protein. In addition, use of a protease adds additional steps and costs to the process. (2) Shoji and coworkers' IgE binding studies used sera pooled from a collection of human patients (11, 12). Patients with low levels of antigen-specific IgE may be under represented or omitted from pooled sera. Thus, Shoji and coworkers' results may overestimate the reactivity of their baculovirus rDer f 1.

In this communication, we report on a process for producing recombinant versions of Der f 1 in *Pichia pastoris*. In contrast to the work of others, our process produces a nonfused, mature rDer f 1 that is fully

active following recovery from the expression host. Recombinant Der f 1 is easily purified from the yeast culture supernatant by either cation exchange chromatography or monoclonal antibody affinity chromatography. No refolding or enzymatic cleavage is required to enhance the immunological activity of the protein. The allergen properties (IgE-binding activity measured in 42 individual HDM allergic patients and *in vitro* proteolytic activity) of the *P. pastoris* rDer f 1 compare very favorably with the native antigen.

## MATERIALS AND METHODS

### DNA Manipulations

Plasmids pPICZ $\alpha$ A and pPICZ $\alpha$ B, *P. pastoris* strain X33, and the antibiotic zeocin were obtained from Invitrogen Corporation (Carlsbad, CA). Plasmid DNA isolation and DNA cloning were done following standard protocols (13). DNA inserts were sequenced on both strands using the ABI PRISM dRhodamine Terminator Cycle Sequencing Ready Reaction kit (PE Biosystems, Foster City, CA) and a PE Biosystems Prism (377-96-XL) automated DNA sequencer. Cycle sequencing conditions were those specified in the manufacturer's protocol. Primers for PCR and sequencing were obtained from Life Technologies (Gaithersburg, MD). DNA sequences were assembled and evaluated using DNASIS, v2.1 (Hitachi Software Engineering Co., South San Francisco, CA).

### Plasmids and Strains for Expressing rDer f 1

Plasmid pBC128, for expressing pro-Der f 1, was constructed by overlap extension PCR (14) using a set of four primers (EB66, EB67, EB68, and EB69) (Fig. 1). Two primers, EB66 (sense primer corresponding to nucleotides CGCGTCCCTCTCGAGAAAAGAGAGGC-TAGACCAGCTTCAATCAAA) and EB67 (antisense primer corresponding to nucleotides GGGCTTTCCTTT-TGCGGCCGCTCACATGATTACAACATATGG), incorporate *Xho*I and *Not*I cloning sites (shown as underlined text) to facilitate cloning into plasmid pPICZ $\alpha$ B. Primer EB66 fuses the DNA encoding the C-terminal portion of the *S. cerevisiae*  $\alpha$  peptide (Leu-Glu), the KEX2 cleavage site (Lys-Arg), and a Glu-Ala spacer sequence to the DNA encoding the first residue (Arg) of the Der f 1 pro-peptide. Primers EB68 (sense primer corresponding to nucleotides GCAAGAGAACAAGATGCAGAAGACCA-AATTCGC) and EB69 (antisense primer corresponding to nucleotides GCGAATTTGGTCTTCTGCATCTTTGT-TCTCTTGC) overlap on opposite strands of the DNA and change rare arginine codons to the most abundant yeast codon for this residue (base changes are shown as underlined text). The ~950-bp *Xho*I-*Not*I insert in pBC128 was generated by a three-step process. In the first step, primers EB66 and EB69 were combined with a *D. farinae*

cDNA library to produce an ~600-bp fragment. The PCR reaction included the following reactants in a 50- $\mu$ l reaction volume:  $2 \times 10^7$  plaque-forming units (PFU) of the *D. farinae* cDNA library (prepared from whole-body mites in  $\lambda$ ZAP (Stratagene, La Jolla, CA),  $1 \times$  Pfu (*Pyrococcus furiosus*) DNA polymerase buffer (Stratagene), 0.2 mM each of the four dNTPs, 2.5 U Pfu DNA polymerase, and 0.5  $\mu$ M forward and reverse primers (EB66 and EB69). Amplification conditions were 1 cycle of 96°C for 3 min; 30 cycles of 96°C for 1 min, 55°C for 1 min, 72°C for 2 min; and 1 cycle of 72°C for 7 min. A second PCR reaction was performed using primers EB67 and EB68 using the aforementioned conditions to produce an ~400-bp product. The two PCR products were gel purified and used as templates for a third round of PCR amplification using the conditions specified previously. The resulting ~950-bp fragment was gel purified, digested with *Xho*I and *Not*I, and ligated into similarly digested pPICZ $\alpha$ B (15).

The ligation mixture was transformed into *E. coli* strain DH5 $\alpha$ , and clones were selected at 37°C on low-salt LB medium with zeocin (15). One plasmid with the expected DNA sequence was saved as pBC128. Plasmid pBC128 was linearized at a unique *Sac*I site in the 5' AOX1 sequence, and introduced into *P. pastoris* strain X33 by electroporation, essentially as described by Cregg and Russell (16). Recombinants were selected at 30°C on YPD medium (15) supplemented with 1000  $\mu$ g/ml zeocin. Integration of the expression cassette into the *P. pastoris* genome was confirmed by PCR using 5' AOX1 (5'-GACTGGTTC AATTGACAAGC) and 3' AOX1 (5'-GCAAATGGCATTCTGACATCC) primers and conditions available from Invitrogen Corporation. One clone was saved as strain HCY215.

Plasmid pBC138, for expressing mature Der f 1, was derived from pBC128 by PCR mutagenesis using primers SB11 (sense primer corresponding to nucleotides CGCATAGTCCCTCTCGAGAAAAGAACAAGCGCTTGCCGTATC) and EB67. Primer SB11 incorporates a 5' *Xho*I restriction site (shown in underlined text) and encodes a fusion of the first residue of the Der f 1 mature peptide in frame with the KEX2 cleavage sequence (Lys-Arg) of pPICZ $\alpha$ A. A single round of PCR amplification was performed in a 50- $\mu$ l reaction volume using 50 ng pBC128 DNA as template and primers SB11 and EB67. Reactant concentrations and PCR conditions were as described above. The Der f 1 sequence in pBC138 is identical to the sequence in pBC128, with the exception that the former plasmid lacks the region encoding the 80-residue pro-peptide. The recombinant plasmid was linearized with *Sac*I and electroporated into competent *P. pastoris* strain X33 cells, and a zeocin-resistant (1000  $\mu$ g/ml) candidate was saved as strain HCY286.

An expression plasmid for producing an unglycosylated, pro-Der f 1 was derived from pBC128 by inverse PCR mutagenesis (17). The PCR reaction included the

following components in a 50- $\mu$ l reaction volume: 25 ng pBC128 plasmid DNA,  $1 \times$  PFU polymerase buffer, 0.2 mM each of the four dNTPs, 2.5 U PFU polymerase, and 0.5  $\mu$ M forward and reverse primers (EB89 and EB90). Primers EB89 (sense primer corresponding to nucleotides GCCGCAACTGAATCAGCTTATTTGGCCTACCGTCAAACGCTCTTTGG) and EB90 (antisense primer corresponding to nucleotides GACACCAGAGAAAGCCCAACATGAACCACAGCCTCCTTG CATACG) contain 5' phosphates and are situated end-to-end on opposite strands of the DNA template. Primer EB89 changes an asparagine codon (AAC) to glutamine (CAA); the relevant bases are shown as underlined text in the above primer sequence. The amplification conditions were 1 cycle of 96°C for 3 min; 30 cycles of 96°C for 1 min, 55°C for 1 min, 72°C for 9 min; and 1 cycle of 72°C for 7 min. The amplified product was desalted, self-ligated with T4 DNA ligase (Life Technologies, Gaithersburg, MD), and transformed into competent *E. coli* DH5 $\alpha$  cells. The mutated plasmid was sequenced and shown to have the expected base changes, and one isolate was saved as pBC158. Plasmid pBC158 was linearized with *Sac*I and transformed into competent *P. pastoris* X33 cells, and transformants were selected as described above. A representative clone was saved as strain HCY288.

#### *Expression and Purification of Native and Recombinant and Proteins*

*Native Der f 1.* *Dermatophagoides farinae* mites were produced at Heska Corporation from a culture obtained from Dr. Federico Montelegre (Department of Microbiology, University of Ponce, School of Medicine, Puerto Rico). Native Der f 1 was purified from an aqueous extract of mites using a modification of the procedure described by Lombardero and coworkers (18). Ten grams of house dust mites (whole mites and feces) was extracted, without homogenization, with 20 ml of phosphate-buffered saline (PBS), pH 7.5, containing protease inhibitors (Complete Kit, Boehringer Mannheim, Indianapolis, IN) for 4 h at 4°C. The extract was concentrated to 5 ml and loaded onto a Sephacryl S-100 column (2.5  $\times$  70 cm, Amersham Pharmacia Biotech) previously equilibrated with PBS. SDS-PAGE analysis showed the presence of a major 31-kDa band in fractions eluting with a retention volume consistent with molecular masses of 25–30 kDa. Pooled fractions were applied to a monoclonal antibody affinity column, prepared as described below. Bound protein was eluted as described below and neutralized immediately with one-tenth volume of 1 M Tris, pH 7.0.

*Expression of rDer f 1.* One liter baffled flasks with porous silicon closures containing 100 ml BMG (15) were inoculated with *P. pastoris* clones. Cultures were grown in an orbital shaker incubator (30°C, 250 rpm) for approximately 48 h to an optical density of 4–6 at

600 nm. Cells were collected by centrifugation and pellets were resuspended to an optical density of approximately 1 (600 nm) in 1.5 L BMM medium (15) in 6-L baffled flasks with porous closures. Cultures were induced at 30°C, 250 rpm for 5 days; methanol (0.5% v/v) was added daily. The supernatant fraction from candidate clones was collected, and cells and debris were removed by centrifugation (20 min at ~16,000g). Recombinant proteins were purified further as described below.

**Cation exchange chromatography.** Supernatants (1–4 L) from *P. pastoris* cultures expressing rDer f 1 were concentrated 5- to 10-fold and dialyzed against 25 mM sodium acetate, pH 4.5 (Buffer A), using tangential flow diafiltration (3000 MW cutoff, AG Technologies, Needham, MA). The dialyzed samples were loaded onto an SP-Sephacrose column (1.6 × 10 cm, Amersham Pharmacia Biotech, Piscataway, NJ). Bound protein was eluted with a linear salt gradient to 100% Buffer B (25 mM sodium acetate, 1 M NaCl, pH 4.5) in 25-column volumes at a flow rate of 5.0 ml/min. Fractions (5.0 ml) were collected, analyzed, and concentrated as described below.

**Monoclonal antibody affinity chromatography.** Approximately 75.0 mg mAb 4C1 (19) was coupled to 40 ml cyanogen bromide-activated Sepharose 4B (Amersham Pharmacia Biotech, Inc.) according to the manufacturer's recommendations. Supernatants from *P. pastoris* cultures containing rDer f 1 were adjusted to pH 8.5 and loaded onto the monoclonal antibody affinity column (2.5 × 8 cm) previously equilibrated with 500 mM NaCl, 50 mM Tris, pH 8.5 (equilibration buffer), at a flow rate of 2–3 ml/min. The column was washed with 5–10 column volumes of the equilibration buffer and bound protein was eluted with 50 mM diethylamine, pH 11.6. Fractions containing rDer f 1 were immediately neutralized with 2 M Tris, pH 7.6 (1:20, vol:vol/fraction).

Fractions containing nDer f 1 and rDer f 1 isolated by either ion exchange or affinity chromatography were pooled, concentrated 5- to 10-fold using a Centriprep 10 concentrator (Amicon, Beverly, MA), and stored at –80°C. The purities of the various proteins were assessed by SDS-PAGE analysis (20) and staining with either silver nitrate or Coomassie blue. Protein concentration was determined using Bradford Coomassie blue dye reagent (Bio-Rad, Hercules, CA) according to the manufacturer's instructions.

#### *N-Terminal Sequencing*

Protein samples were separated by SDS-PAGE, blotted to a PVDF membrane, and visualized by staining with Coomassie blue. Relevant bands were excised and sequenced using standard pulsed liquid Edman degradation chemistry on either Applied Biosystems

Procise 473A or Procise HT 494-01 Sequencing System using reagents from Applied Biosystems (PE Biosystems).

#### *Endoglycosidase Digestion*

Recombinant versions of Der f 1 (1.5–3.0 µg) were incubated with 0.04 units *N*-glycosidase F (Boehringer Mannheim, Indianapolis, IN) in a 25-µl reaction containing 40 mM sodium phosphate, pH 7.0, 0.5% SDS, 1.0% octylglucoside, 10 mM EDTA. The reaction was incubated at 37°C for 24 h and stopped by addition of 5 µl Laemmli sample buffer (20). Samples were analyzed by SDS-PAGE. Proteins were visualized by staining with Coomassie blue.

#### *Mass Spectrometry*

The recombinant and native forms of Der f 1 were dialyzed against water, lyophilized, and analyzed by matrix-assisted laser desorption ionization mass spectrometry on a Voyager DE MALDI-MS (PE Biosystems, Foster City, CA). Samples were dissolved in deionized 8 M urea and diluted 1:10 in matrix. Mass analysis was performed in the presence of internal standards, insulin (5.74 kDa), thioredoxin (11.7 kDa), and apomyoglobin (16.9 kDa).

#### *Carbohydrate Analysis of Proteins*

Samples of both nDer f 1 and rDer f 1 were isolated by reverse-phase chromatography. Samples (~250–300 µg) were injected onto an RP-C1 column (2.1 × 250 mm, Phenomenex, Torrance, CA) previously equilibrated with 0.1% trifluoroacetic acid (TFA), 2% acetonitrile (CH<sub>3</sub>CN). Bound protein was eluted with 100% CH<sub>3</sub>CN, 0.085% TFA using a linear gradient (1.5% change/min). Protein samples were assayed for carbohydrate content using a colorimetric glycoprotein carbohydrate estimation kit (Pierce, Rockford, IL) according to the manufacturer's instructions.

#### *Monoclonal Antibodies and Sera*

Murine monoclonal antibodies mAb 4C1 and mAb 6A8 were produced and selected as described previously (19). The mAb 4C1 recognizes epitopes on both Der f 1 and Der p 1, while mAb 6A8 is specific for Der f 1. The two mAbs recognize distinct epitopes on Der f 1 (19, 21).

Collection of human sera for chimeric ELISAs was approved by the Human Investigation Committee of the University of Virginia. Human sera were collected from patients who were skin test positive for *D. pteronyssinus* and who were shown to have HDM-specific IgE using the CAP assay (CAP scores of 3–6, Pharmacia Upjohn, Piscataway, NJ). None of the patients was receiving immunotherapy when the serum was collected.

Pooled human sera for competition ELISA (purchased from PlasmaLab International, Everitt, WA) had IgE against *D. farinae* and RAST scores of 1.1 to 1.6.

#### Competition ELISA

Competition ELISA was performed to compare the abilities of recombinant forms of Der f 1 to inhibit binding of IgE antibodies in the sera of HDM allergic patients to native Der f 1. Varying concentrations of rDer f 1 were preincubated with pooled sera from HDM allergic patients (PlasmaLab International, Everitt, WA) diluted 1:30 for 1 h at room temperature. A recombinant flea salivary antigen, rCte f 1 (Heska Corp., Ft. Collins, CO), was used as a negative control. Unbound IgE present in the sera was detected by incubating the sera for 2 h in nDer f 1-coated plates. A biotinylated version of the recombinant subunit of the human high-affinity IgE receptor (Heska Corp., Ft. Collins, CO), FcεRIα, in combination with streptavidin-horseradish peroxidase, was used to detect antigen-specific IgE. Background absorbance was subtracted out. Percentage inhibition was calculated by comparing the relative amount of IgE bound to native Der f 1 when sera were preincubated with recombinant forms of Der f 1 to the amount of bound IgE observed using sera preincubated without recombinant allergen.

#### Chimeric IgE ELISA

The relative abilities of native and recombinant Der f 1 to bind antigen-specific IgE antibodies were determined using a chimeric IgE ELISA adapted from published protocols (22). Briefly, a standard curve for the quantitation of serum IgE was created by coating microtiter plates (Immulon2, Dynatech) with 1000 ng/well of a chimeric antibody containing portions of anti-Der p 2 mAb @ Dpx. The plates were blocked with 300 μl/well of 1% BSA in PBS-T for 30 min at room temperature and coated with 50 ng/well of native Der p 2 in PBS for 1 h at room temperature. The chimeric IgE antibody was titrated from 50 to 0.1 ng/well in duplicate, across the plate, and incubated for 1 h at room temperature. Bound chimeric antibody was detected using a 1/4000 dilution of biotinylated goat anti-human IgE (Kirkegaard & Perry, UK). Following a 1-h incubation period (room temperature) with biotinylated goat anti-human IgE, streptavidin-peroxidase (Sigma, St Louis, MO) and ABTS (Sigma, St Louis, MO) were added according to the manufacturer's instructions. The IgE standard curve was constructed by plotting absorbance at 405 nm against known concentrations of the chimeric IgE antibody.

Microtiter plates were coated with 1.0 μg/well of mAb 4C1 in 50 mM carbonate/bicarbonate buffer, pH 9.6, and incubated overnight at 4°C. The plates were blocked as described above and coated with 50 ng/well

of either native Der f 1 or rDer f 1 in PBS for 1 h at room temperature. Serum samples (100 μl in PBS-T) were added at 1:1 and 1:10 dilutions and incubated for 1 h at room temperature. Background absorbances were subtracted out. Bound IgE was detected as described above and quantitated by direct comparison to the IgE standard curve.

#### Protease Activity

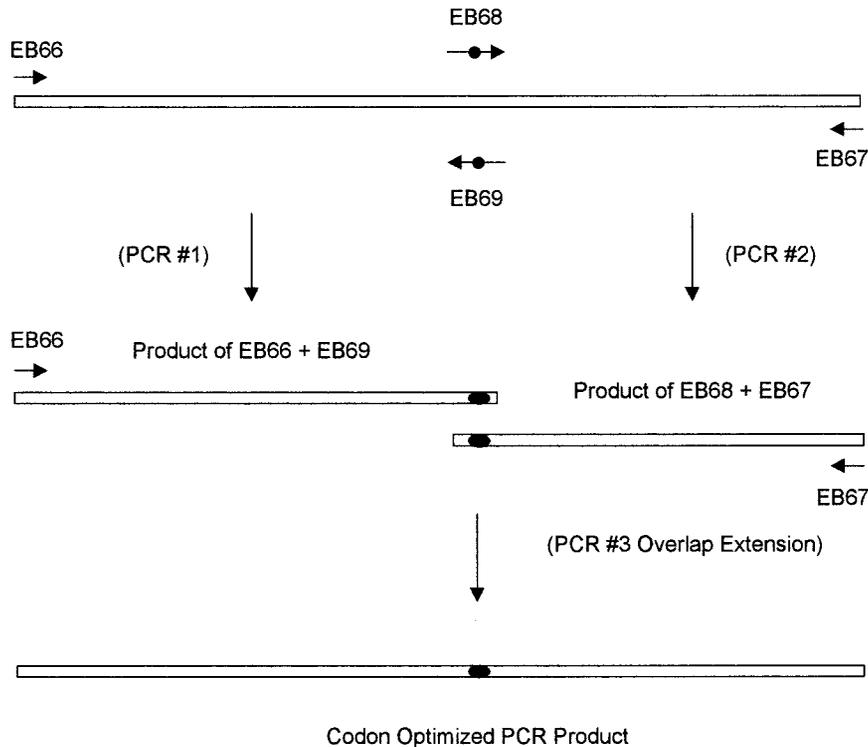
Zymogram casein gels containing 12% acrylamide were purchased from Novex (Novex, San Diego, CA) and used according to the manufacturer's instructions.

## RESULTS AND DISCUSSION

#### Molecular Biology

Our initial efforts to express the mature Der f 1 in *E. coli* using the pET expression system developed by Studier and coworkers (23) produced poor results. SDS-PAGE analysis of whole-cell extracts of induced cultures did not reveal an obvious protein band (data not shown). Inspection of the mRNA sequence encoding the mature Der f 1 revealed a cluster of rare arginine codons (24, 25) at residues 104, 106, and 107 (residue numbers based on the sequence of Dilworth and coworkers (6)). Replacement of these rare arginine codons (CGA) with synonymous, *E. coli* preferred codons (CGT) resulted in a significant improvement in pro-Der f 1 expression in *E. coli*, but the recombinant protein was insoluble (data not shown). Der f 1 contains six cysteine residues that are presumed to be involved in disulfide bonds (6). Because disulfide bonds are not typically formed in the reducing environment of the *E. coli* cytoplasm, it was not surprising to find that recombinant pro-Der f 1 was insoluble and sequestered in inclusion bodies in *E. coli*. Pro-Der f 1 isolated from *E. coli* inclusion bodies could be refolded into monomers, but these monomers bound IgE in the serum of HDM allergic patients with significantly lower affinity than nDer f 1 (data not shown). We do not know whether this reduced affinity is due to the presence of the pro-peptide, which could obscure IgE epitopes, or to misfolding of the recombinant antigen.

Illy and coworkers (26) showed that *P. pastoris* strains engineered to secrete pro-cathepsin B, another member of the papain superfamily, produced the pro-enzyme. Treatment of pro-cathepsin B with acid resulted in efficient removal of the pro-peptide (26). Based on these investigators' success, we decided to evaluate Der f 1 expression in *P. pastoris*. Because CGA is a rare codon in yeast (*S. cerevisiae*), we suspected this triplet arginine cluster would be an impediment for high-level protein production in *P. pastoris* (24). Accordingly, we developed a PCR strategy (Fig. 1) for changing the CGA arginine



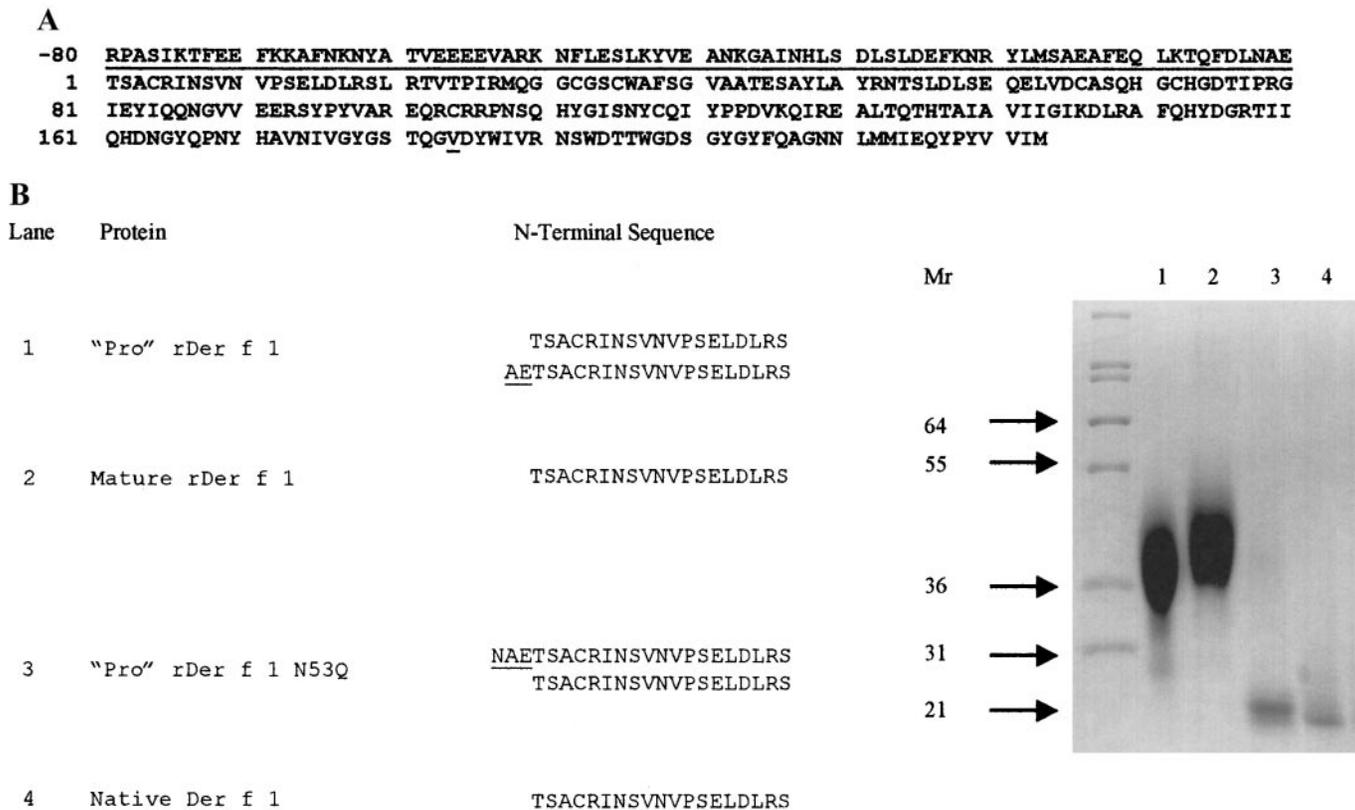
**FIG. 1.** Assembly of a codon-optimized version of rDer f 1. A three-step PCR amplification process was used to produce the *XhoI*–*NotI* fragment in plasmid pBC128. Primers EB68 and EB69 change rare arginine codons to ones preferred by yeast. Optimized codons are indicated by a filled circle in the primer arrow or in the PCR product. Separate PCR experiments were performed using primers EB66 and EB69 (PCR #1) and primers EB68 and EB67 (PCR #2). The products from PCR #1 and PCR #2, which contain optimized arginine codons and approximately 30 nucleotides of overlapping sequences, were annealed and reamplified in a third PCR experiment using primers EB66 and EB67 (PCR #3 Overlap Extension). The final, codon-optimized product was subcloned into pPICZ $\alpha$ B as described in the text.

codons at residues 101, 104, 106, and 107 to AGA, the preferred arginine codon in *S. cerevisiae*. Because it was impossible to know *a priori* whether proper folding of the mature form in *Pichia* requires the pro-peptide, we constructed two nearly identical expression plasmids: one (pBC128) that encodes a fusion of the *S. cerevisiae*  $\alpha$  factor pre-pro signal sequence to the pro-Der f 1 sequence and one (pBC138) that encodes a fusion of the *S. cerevisiae*  $\alpha$  factor pre-pro region to the first amino acid of the mature Der f 1.

Plasmid pBC128, which was designed for expressing pro-Der f 1 in *P. pastoris*, was found to contain a valine codon (GTC) in place of aspartic acid codon (GAC) at a position encoding residue 184 of the mature Der f 1 protein (Fig. 2) (6). Because our *E. coli* pro-Der f 1 expression clones also contained a valine codon at this position, we believe this substitution is a polymorphism rather than a PCR-induced error (data not shown). This substitution is not surprising. Sequence polymorphisms have been reported in cDNA clones of Der p 1 (27) and of the group 2 HDM antigens (3).

### Protein Expression

Recombinant versions of Der f 1 were greater than 90% homogeneous by SDS-PAGE (Coomassie blue staining) following elution from SP-Sephrose at 100–150 mM NaCl. *P. pastoris* strains engineered to produce either pro-Der f 1 (strain HCY215) or mature Der f 1 (strain HCY286) were found to produce proteins that appeared as diffusely migrating products with apparent molecular masses ranging from 35 to 50 kDa under reducing conditions (Fig 2). The apparent heterogeneity of the recombinant proteins suggests that they are modified with heterogeneous mixtures of glycans at the single, N-linked glycosylation site. In contrast, *P. pastoris* strain HCY288, which was engineered to produce a pro-enzyme form with an altered N-linked glycosylation site, secreted a protein that migrated as a single, sharp band, with an apparent molecular mass of ~25 kDa (slightly larger than the mass of nDer f 1 (see Fig. 2). In agreement with the findings of Lombardero and coworkers (18), we found that both the native and the unglycosylated recombinant versions of Der f 1 migrated with the same apparent



**FIG. 2.** Amino acid sequence of nDer f 1 and recombinant forms. (A) Amino acid sequence of nDer f 1. The underlined residues include the 80-residue pro-enzyme region. The mature Der f 1 consists of positively numbered amino acid residues 1 to 223. The valine polymorphism at residue 184 is also underlined. (B) Reducing SDS-PAGE analysis of recombinant forms of Der f 1. Proteins are stained with Coomassie blue. Lanes 1 and 2 contain 7  $\mu$ g of protein recovered from strains HCY215 (pro-Der f 1 construct) and HCY286 (mature Der f 1 construct), respectively. Lanes 3 and 4 contain 3  $\mu$ g each of protein recovered from strain HCY288 (unglycosylated pro-rDer f 1 N53Q construct) and native Der f 1, respectively. The designation "pro" in the figure indicates that the gene construct was the pro-enzyme; the secreted product was processed to the mature enzyme. N-terminal amino acid sequences derived from the major bands present in lanes 1–4 are also depicted.

molecular mass under nonreducing conditions (data not shown).

Surprisingly, we found that *P. pastoris* strains engineered to express either pro-Der f 1 (HCY215) or mature Der f 1 (HCY286) secreted the mature protein. In contrast to Illy and coworkers' work on cathepsin B expression in *P. pastoris* (26) and Shoji and coworkers' work on Der f 1 expression in insect cells (11, 12), no postpurification processing (acid treatment or enzymatic cleavage) was required to remove the pro-peptide from Der f 1.

Recombinant versions of Der f 1 recovered from *P. pastoris* supernatants were processed to varying extents. N-terminal sequencing of rDer f 1 recovered from strains engineered to produce pro-Der f 1 (HCY215) contained two major products: a correctly processed mature form, TSACRINSVNVPSEL... , and a slightly larger product that included the last two residues of the pro-peptide, AETSACRINSVNVPSEL... (Fig. 2). The molar ratio of these two products was approximately equivalent. Protein recovered from strains en-

gineered to express mature Der f 1 (HCY286) contained a single sequence: TSACRINSVNVPSEL... N-terminal sequence analysis of the protein recovered from strain HCY288 (designed to produce an unglycosylated pro-enzyme) revealed two major sequences: NAETSACRINSVNVPSEL... and TSACRINSVNVPSEL... Proteins with these two sequences were found in relatively equal amounts.

Protein yields from strains expressing the various forms of rDer f 1 differed considerably. The highest yield (~30–35 mg/L) was obtained from the strain engineered to express pro-Der f 1 (HCY215). Approximately 10-fold lower yields (3–4 mg/L) were obtained from strains engineered to express the mature Der f 1 (HCY286). These results suggest that the pro-peptide is not absolutely required for proper folding in the yeast expression system. However, the pro-peptide may either stabilize rDer f 1 or facilitate its secretion. Strains expressing the unglycosylated version of pro-Der f 1 (HCY288) produced only 1.0–1.5 mg/L protein. Although we cannot rule out the possibility that these

strains contained different copy numbers of the expression cassette (15), these data suggest a role for oligosaccharides in expression of rDer f 1 in the yeast expression host. For example, these structures could affect secretion or the stability of the secreted protein (28).

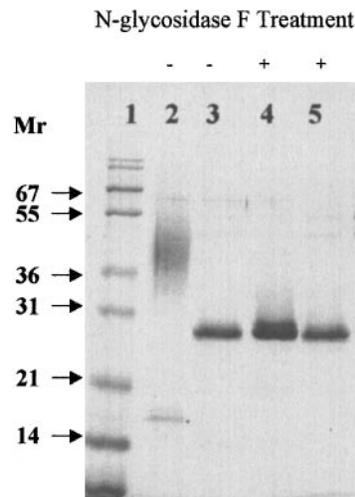
### Carbohydrate Estimation and Mass Spectrometry

Our nDer f 1 preparations migrated as a discrete protein band with an apparent molecular mass of ~25–30 kDa on SDS-PAGE. The predicted mass of Der f 1 is 25,185 Da (6). MALDI mass spectrometry of nDer f 1 indicated masses of 25,500 and 25,850 Da. The calculated mass for the unglycosylated version of rDer f 1 (HCY288) is 25,183 Da. MALDI-MS analysis of rDer f 1 isolated from strain HCY288 revealed three species of protein: two species (25,133 and 25,250 Da) are consistent with the correctly processed mature form (TSACR. . .), and one (25,402 Da) is consistent with a mature form that includes two residues of the pro-peptide (AETSACR. . .). Accurate mass determination of the glycosylated form of rDer f 1 (HCY215) was not successful due to its heterogeneity.

Mature Der f 1 contains a single, putative N-glycosylation site at residues 53–55 of the mature protein (6). Several investigators have reported that the group 1 HDM allergens are glycoproteins (12, 29, 30). One report identified galactose, mannose, *N*-acetylglucosamine, and *N*-acetylgalactosamine in fractions of purified nDer p 1 (29). Separate carbohydrate analyses of nDer f 1 and rDer f 1 (HCY215) confirmed the presence of carbohydrates on these molecules (data not shown). No carbohydrate was detected in the unglycosylated rDer f 1 (HCY288) (data not shown). Together with the mass spectroscopy data, these results suggest the presence of a mixture of short carbohydrate residues at the consensus N-linked glycosylation site in nDer f 1. Based on the observed masses (25,500 and 25,850 Da), it is tempting to speculate that the glycan portion of nDer f 1 may include one to three hexoses. However, we cannot rule out other amino acid modifications that could also affect the observed mass.

### Endoglycosidase Assays

Treatment of pro-rDer f 1 with *N*-glycosidase F reduces the mass of the protein from 35 to 50 kDa to ~25 kDa (Fig. 3, lanes 2 and 4). The molecular mass of endoglycosidase-treated pro-rDer f 1 is comparable to the mass of a recombinant form of Der f 1 that contains a mutation (N53Q) that alters the consensus N-linked glycosylation site (Fig. 3, lanes 3 and 4). Treatment of pro-Der f 1 (N53Q) with *N*-glycosidase F does not further reduce the apparent mass of the protein (Fig. 3, lane 5). These data strongly suggest that the diffuse mobility of *Pichia*-produced rDer f 1 on SDS-PAGE is



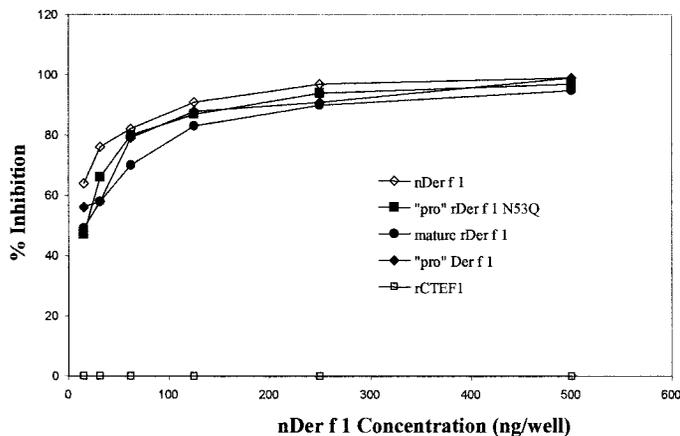
**FIG. 3.** Digestion of recombinant forms of Der f 1 with *N*-glycosidase F. Proteins were separated by SDS-PAGE and visualized by staining with Coomassie blue. Lane 1, molecular mass markers; lane 2, 3 µg pro-rDer f 1; lane 3, 1.5 µg pro-rDer f 1 (N53Q); lane 4, 3.0 µg *N*-glycosidase F-treated pro-rDer f 1; lane 5, 1.5 µg *N*-glycosidase F-treated pro-rDer f 1 (N53Q). The designation “pro” indicates that the gene construct was the pro-enzyme; the secreted product was shown to have the mature amino terminus.

due to heterogeneous N-linked glycosylation at a single site.

### Immunological Assays

Results from competition ELISA experiments suggest that recombinant versions of Der f 1 isolated from *P. pastoris* have tertiary conformations very similar to those of the native molecule. Competition ELISA experiments showed that the binding of nDer f 1 to serum IgE can be inhibited by recombinant forms of Der f 1 with efficiencies similar to those seen when nDer f 1 is allowed to inhibit its own binding to IgE (Fig. 4). A negative control consisting of a recombinant flea salivary antigen (rCte f 1) failed to inhibit Der f 1 binding, regardless of the amount of antigen used (Fig. 4). Our finding that glycosylated and unglycosylated versions of rDer f 1 compete equally well with nDer f 1 indicates that the oligosaccharides contributed by the *P. pastoris* expression system are not essential for IgE binding activity *in vitro*.

The amount of Der f 1-specific IgE antibodies in the sera of HMD allergic patients can be quantitated by comparing the ELISA signal obtained using a chimeric IgE antibody standard titrated against an allergen standard, with the ELISA signals obtained using either native or rDer f 1 (glycosylated pro-rDer f 1 isolated from HCY215) in an IgE capture ELISA (Fig. 5). These results show that, using sera from 42 HDM allergic patients, the amount of allergen specific IgE detected using rDer f 1 correlated well with the amount detected using native

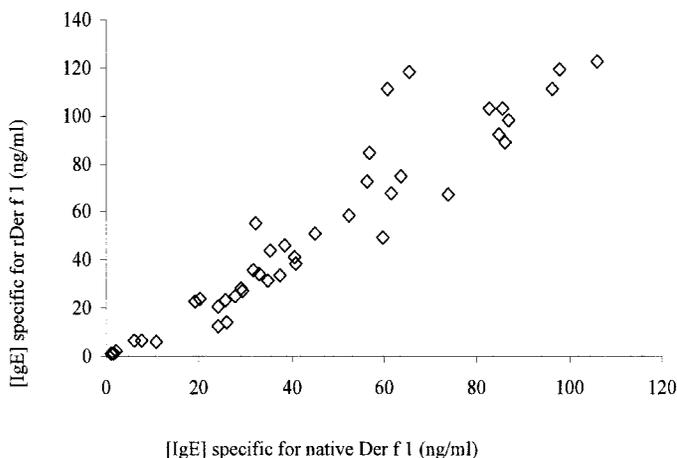


**FIG. 4.** Competition ELISA. Inhibition of nDer f 1 binding to IgE in sera of HDM allergic patients by recombinant Der f 1 (pro-rDer f 1, mature rDer f 1 unglycosylated (N53Q) pro-rDer f 1) and nDer f 1. Recombinant flea salivary antigen Cte f 1 was used as a negative control. The designation "pro" indicates that the gene construct was engineered to produce the pro-enzyme form; the protein recovered was shown to be the mature protein.

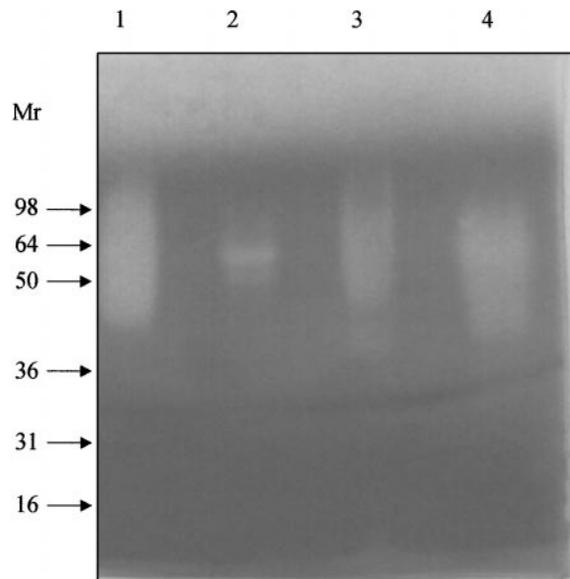
Der f 1 ( $r = 0.97$ ). Sera from seven nonallergic patients were used as negative controls. Sera from atopic patients had 1–125 ng/ml Der f 1-specific IgE and CAP scores of 0–6. Sera from seven nonatopic patients had  $\leq 1$  ng/ml Der f 1-specific IgE and CAP scores of 1.

#### Protease Activity

Hewitt and coworkers have demonstrated that native Der p 1, the archetypal group 1 HDM allergen isolated from *D. pteronyssinus*, has proteolytic activity against casein (4). Although Der p 1 is often described as a classical cysteine protease, Hewitt and coworkers demonstrated that nDer p 1 is inhibited by serine as



**FIG. 5.** Chimeric IgE ELISA. Quantitation of serum IgE binding to nDer f 1 or pro-rDer f 1 in HDM allergic sera. The protein designated "pro" rDer f 1 is a processed form containing the mature N-terminus.



**FIG. 6.** Zymogram gel of native and recombinant forms of Der f 1. Lane 1, 2  $\mu$ g protein from strain HCY215 (pro-rDer f 1 construct); lane 2, 1  $\mu$ g protein from strain HCY288 (unglycosylated (N53Q) pro-rDer f 1 construct); lane 3, 5  $\mu$ g Der f 1 from strain HCY286 (mature rDer f 1 construct); lane 4, 0.5  $\mu$ g Der f 1 isolated from dust mites.

well as cysteine protease inhibitors (4). This type of mixed mechanistic activity is unusual for a cysteine protease, and Hewitt and coworkers have speculated that the proteolytic activity of the group 1 proteins may contribute to their allergenicity (4, 5).

Results of zymogram gels (Fig. 6) indicate that nDer f 1 and recombinant versions of Der f 1 are capable of cleaving  $\beta$ -casein. Recombinant Der f 1 isolated from strain HCY215 was also proteolytically active against the fluorogenic substrate Z-Val-Leu-Arg-AMC (data not shown). This activity was inhibited by E64 (*L*-trans-epoxysuccinylleucylamide-(4-guanidino)-butane), a specific inhibitor of cysteine proteases. These data indicate that rDer f 1 isolated from *P. pastoris* has cysteine protease activity and are in agreement with Hewitt's finding about the enzymatic activity of the group 1 allergens (4).

The *in vitro* protease activity of rDer f 1, combined with our observation that *Pichia* strains engineered to produce pro-Der f 1 secrete mature Der f 1, suggests that the pro-peptide undergoes self-processing. We hope to obtain additional support for autoprocessing by constructing a version of Der f 1 with a residue change that alters residue C35, the predicted active site cysteine (6). The resulting rDer f 1 is expected to be secreted from *Pichia* as the pro-enzyme form if the processing event is self-catalyzed. The availability of a protease-defective version of Der f 1 could allow investigators to test Hewitt and coworkers' hypothesis that the proteolytic function of the protein "endows it with properties of a proallergic adjuvant" (5).

## ACKNOWLEDGMENTS

We are grateful to Stephanie Blehm for help with yeast molecular biology and expression work, Michael Ellis for DNA sequencing, and Eric Weber for his generous gift of a *D. farinae* cDNA library. We also thank Don Wassom, Dan Stinchcomb, and Carol Talkington Verser for helpful discussion and comments about the manuscript.

## REFERENCES

- Tovey, E. R., Chapman, M. D., and Platts-Mills, T. A. E. (1981) Mite faeces are a major source of house dust allergens. *Nature* **289**, 592–593.
- Meyer, C. H., Bond, J. F., Chen, M. S., and Kasaian, M. T. (1994) Comparison of the levels of the major allergens Der p I and Der p II in standardized extracts of the house dust mite, *Dermatophagoides pteronyssinus*. *Clin. Exp. Allergy* **24**, 1041–1048.
- Thomas, W. R. (1998) House-dust-mite allergens. *Allergy* **53**, 821–832.
- Hewitt, C. R. A., Brown, A. P., Hart, B. J., and Pritchard, D. I. (1995) A major house dust mite allergen disrupts the immunoglobulin E network by selectively cleaving CD23: Innate protection by antiproteases. *J. Exp. Med.* **182**, 1537–1544.
- Hewitt, C. R., Foster, S., Phillips, C., Horton, H., Jones, R. M., Brown, A. P., Hart, B. J., and Pritchard, D. I. (1998) Mite allergens: Significance of enzymatic activity. *Allergy* **53**, 60–63.
- Dilworth, R. J., Chua, K. Y., and Thomas, W. R. (1991) Sequence analysis of cDNA coding for a major house dust mite allergen, Der f I. *Clin. Exp. Allergy* **21**, 25–32.
- Chua, K. Y., Stewart, G. A., Thomas, W. R., Simpson, R. J., Dilworth, R. J., Plozza, T. M., and Turner, K. J. (1988) Sequence analysis of cDNA coding for a major house dust mite allergen, Der p I: Homology with cysteine proteases. *J. Exp. Med.* **167**, 175–182.
- Okumura, Y. (1996) Production of recombinant allergens and their application for immunotherapy. *Adv. Exp. Med.* **409**, 163–171.
- Greene, W. K., Cyster, J. G., Chua, K. Y., O'Brien, R. M., and Thomas, W. R. (1991) IgE and IgG binding peptides expressed from fragments of cDNA encoding the major house dust mite allergen Der p I. *J. Immunol.* **147**, 3768–3773.
- Chua, K. Y., Kehal, P. K., Thomas, W. R., Vaughan, P. R., and Macreadie, I. G. (1992) High-frequency binding of IgE to the Der p allergen expressed in yeast. *J. Allergy Clin. Immunol.* **89**, 95–102.
- Shoji, H., Hanawa, M., Shibuya, I., Hirai, M., Yasuhara, T., Okumura, Y., and Yamakawa, H. (1996) Production of recombinant mite allergen Der fI in insect cells and characterization of products—Removal of Pro-sequence is essential to IgE-binding activity. *Biosci. Biotech. Biochem.* **60**, 621–625.
- Shoji, H., Shibuya, I., Hirai, M., Horiuchi, H., and Takagi, M. (1997) Production of recombinant Der fI with the native IgE-binding activity using a baculovirus expression system. *Biosci. Biotech. Biochem.* **16**, 1668–1673.
- Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Horton, R. M., Cai, Z. L., Ho, S. N., and Pease, L. R. (1990) Gene splicing by overlap extension: Tailor-made genes using the polymerase chain reaction. *Biotechniques* **8**, 528–535.
- Higgins, D. R., Busser, K., Comiskey, J., Whittier, P. S., Purcell, T. J., and Hoeffler, J. P. (1998) Small vectors for expression based on dominant drug resistance with direct multicopy selection in "Methods in Molecular Biology," pp. 41–53, Humana Press, Totowa, NJ.
- Cregg, J. M., and Russell, K. A. (1998) Transformation in "Methods in Molecular Biology," pp. 27–39, Humana Press, Totowa, NJ.
- Dennis, J. J., and Zylstra, G. J. (1998) Improved antibiotic resistance cassettes through restriction site elimination using *Pfu* DNA polymerase PCR. *Biotechniques* **25**, 772–776.
- Lombardero, M., Heymann, P. W., Platts-Mills, T. A. E., Fox, J. W., and Chapman, M. D. (1990) Conformational stability of B cell epitopes on group I and group II Dermatophagoides spp. allergens. Effect of thermal and chemical denaturation on the binding of murine IgG and human IgE antibodies. *J. Immunol.* **144**, 1353–1360.
- Chapman, M. D., Heymann, P. W., and Platts-Mills, T. A. (1987) Epitope mapping of two major inhalant allergens, Der p I and Der f I, from mites of the genus *Dermatophagoides*. *J. Immunol.* **139**, 1479–1484.
- Laemmli, U. K. (1970) Cleavage of structural proteins during assembly of the head of T4 bacteriophage. *Nature* **227**, 680–681.
- Heymann, P. W., Chapman, M. D., and Platts-Mills, T. A. E. (1986) Antigen Der f I from the dust mite *Dermatophagoides farinae*: Structural comparison with Der p I from *Dermatophagoides pteronyssinus* and epitope specificity of murine IgG and human IgE antibodies. *J. Immunol.* **137**, 2841–2847.
- Ichikawa, K., Iwasaki, E., Baba, M., and Chapman, M. D. (1999) High prevalence of sensitization to cat allergen among Japanese children with asthma, living without cats. *Clin. Exp. Allergy* **29**, 754–761.
- Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89.
- Zhang, S., Zubay, G., and Goldman, E. (1991) Low-usage codons in *Escherichia coli*, yeast, fruit fly, and primates. *Gene* **105**, 61–72.
- Zhang, G., Goldman, E., and Zubay, G. (1994) Clustering of low usage codons and ribosome movement. *J. Theor. Biol.* **170**, 339–354.
- Illy, C., Araiishi, O., Wang, J., Purisima, E., Vernet, T., and Mort, J. S. (1997) Role of the occluding loop in cathepsin B activity. *J. Biol. Chem.* **272**, 1197–1202.
- Thomas, W. R., Smith, W., Hales, B. J., and Carter, M. D. (1997) Functional effects of polymorphisms of house dust mite allergens. *Int. Arch. Allergy Immunol.* **113**, 96–98.
- Zhu, A., Wang, Z.-K., and Beavis, R. (1998) Structural studies of a-N-acetylgalactosaminidase: Effect of glycosylation on the level of expression, secretion efficiency, and enzyme activity. *Arch. Biochem. Biophys.* **352**, 1–8.
- Chapman, M. D., and Platts-Mills, T. A. E. (1980) Purification and characterization of the major allergen from *Dermatophagoides pteronyssinus* antigen P1. *J. Immunol.* **125**, 587–592.
- Wolden, B., Paulsen, B. S., Wold, J. K., and Brimmer, O. (1982) Characterization of the carbohydrate moiety in a purified allergen preparation from the mite *Dermatophagoides farinae* and its importance for allergenic activity as tested by rast-inhibition method. *Int. Arch. Allergy Appl. Immunol.* **68**, 144–151.