Homology Modeling and Characterization of IgE Binding Epitopes of Mountain Cedar Allergen Jun a 3

Kizhake V. Soman,* Terumi Midoro-Horiuti,† Josephine C. Ferreon,* Randall M. Goldblum,† Edward G. Brooks,† Alexander Kurosky,* Werner Braun,* and Catherine H. Schein*

*Sealy Center for Structural Biology and Department of Human Biological Chemistry and Genetics, and †Department of Pediatrics, Child Health Research Center, University of Texas Medical Branch, Galveston, Texas 77555-1157 USA

ABSTRACT The Jun a 3 protein from mountain cedar (*Juniperus ashei*) pollen, a member of group 5 of the family of plant pathogenesis-related proteins (PR-proteins), reacts with serum IgE from patients with cedar hypersensitivity. We used the crystal structures of two other proteins of this group, thaumatin and an antifungal protein from tobacco, both ~50% identical in sequence to Jun a 3, as templates to build homology models for the allergen. The in-house programs EXDIS and FANTOM were used to extract distance and dihedral angle constraints from the Protein Data Bank files and determine energy-minimized structures. The mean backbone deviations for the energy-refined model structures from either of the templates is <1 Å, their conformational energies are low, and their stereochemical properties (determined with PROCHECK) are acceptable. The circular dichroism spectrum of Jun a 3 is consistent with the postulated β-sheet core. Tryptic fragments of Jun a 3 that reacted with IgE from allergic patients all mapped to one helical/loop surface of the models. The Jun a 3 models have features common to aerosol allergens from completely different protein families, suggesting that tertiary structural elements may mediate the triggering of an allergic response.

INTRODUCTION

Hypersensitivity to mountain cedar (*Juniperus ashei*, Cupressaceae) pollen is a frequent cause of severe, seasonal allergic disease (cedar pollinosis). Current treatment, which has not been particularly successful, is limited to symptomatic therapy and attempts at hyposensitization by injection of crude pollen extracts (Platts-Mills et al., 1998). Our overall goal is to evaluate the structural basis of the allergic immune response to mountain cedar pollen and to develop new immunotherapeutic agents based on defined IgE-binding epitopes. We isolated proteins from Texas mountain cedar pollen that react with the IgE in patient sera and cloned their related mRNAs. While one protein, Jun a 1 (Midoro-Horiuti et al., 1999a), was very similar to a protein previously identified as the major allergen in Japanese cedar pollen, we also discovered a second, novel allergen, Jun a 3 (Midoro-Horiuti et al., 2000). Jun a 3, a 30-kDa protein (199 residues), has high sequence identity with the PR-5 group of plant pathogenesis-related proteins (PR proteins) that are overexpressed when plants are subjected to stress conditions or infected with pathogens (Linthorst, 1981). More recently, an allergen (Pru a 2) believed to cause oral hypersensitivity to cherry (*Prunus avium*, Prunoideae) (Inschlag et al., 1998) was found to be a PR-5 group protein. The sequence of this protein is 45% identical to that of Jun a 3. Many members of the PR-5 family, including the PR-5D protein used here for modeling Jun a 3, have antifungal properties; some may also have antiviral activity, as they are produced in response to viral infection in plants (Linthorst, 1981). Structural knowledge of the epitopes responsible for allergenicity of these proteins is essential for designing therapeutic agents based on these proteins.

In this paper we describe experiments to define areas of Jun a 3 that bind IgE and potentially induce allergic reactions. A high-quality 3D structural model of the protein was prepared by homology modeling and energy minimization. The primarily β-sheet secondary structure predicted by the model was consistent with the circular dichroism (CD) spectrum of the protein isolated from pollen. Tryptic fractions of Jun a 3 were isolated by high-performance liquid chromatography (HPLC) and tested for reactivity with IgE in pooled patient sera, and positive fragments were identified by matrix-assisted laser desorption–mass spectrometry (MALDI-MS) and N-terminal sequence analysis. Of the four fragments detected in the HPLC fractions that reacted with patient IgE in dot blotting, three mapped to the same surface-exposed, helix/loop region in the model structure. The fourth, IgE binding to which was not confirmed, was located more internally in the same area. These data will be used in designing studies to further delineate the allergenic structures of this protein.

MATERIALS AND METHODS

Purification of Jun a 3 and isolation of tryptic peptides

Native Jun a 3, a 30-kDa protein (199 residues), was purified as described previously (Midoro-Horiuti et al., 2000). Briefly, defatted mountain cedar pollen was extracted in 0.125 M ammonium bicarbonate (pH 8.0) at 4°C for 48 h and precipitated with ammonium sulfate (40–80% of saturation

Received for publication 14 February 2000 and in final form 23 May 2000.

Address reprint requests to Dr. Catherine H. Schein, Sealy Center for Structural Biology, Route 1157, University of Texas Medical Branch, Galveston, TX 77555-1157. Tel.: 409-747-6810; Fax: 409-747-6850; E-mail: cathy@newton.utmb.edu.

0006-3495/00/09/1601/09 $2.00

fraction) (crude extract). Jun a 3 was isolated from the crude extract by 214TP510 (Vydac, Hesperia, CA) HPLC. The elution was performed with a 30–50% gradient of acetonitrile in 0.1% trifluoroacetic acid.

Tryptic fragments of native Jun a 3 were purified by reverse-phase HPLC (Midoro-Horiuti et al., 1999b). Briefly, 2 mg of Jun a 3 was reduced, alkylated, and repurified, using reverse-phase HPLC on the Vydac column. The Jun a 3 containing peak fraction 3 were lyophilized and redissolved in 0.1 M Tris-HCl (pH 8) containing 2 M urea and 0.01 M $CaCl_2$, *N*-tosyl-L-phenylalanine chloromethyl ketone (TPCK)-treated trypsin (enzyme/substrate ratio = 1:50; Promega, WI) was added, and the sample was incubated for 17 h at 37°C. The tryptic peptides were separated by HPLC on a 218TP52 (Vydac) column, using a 0–45% gradient of acetonitrile in 0.08% trifluoroacetic acid. Fractions were vacuum evaporated, resuspended in water, and used for dot-blot immunostaining, N-terminal sequencing, and mass spectrometry.

Dot blot assay

IgE binding of tryptic peptides was analyzed by dot-blot immunostaining. One microliter of each fraction from HPLC was dotted onto nitrocellulose. The membranes were blocked with 10% (w/v) fat-free milk overnight and incubated with a 1:4 dilution of pooled sera from cedar-hypersensitive patients or normal controls overnight. After the membranes were washed with Tween-TBS (0.05% Tween 20–Tris-buffered saline), they were incubated with 1 μg/ml of biotinylated anti-human IgE (Sigma, St. Louis, MO), followed by incubation with 1:20,000 dilution of horseradish peroxidase-streptavidin (Zymed, San Francisco, CA). The signal was detected using enhanced chemoluminescence Western blot detection reagents (Amersham, Piscataway, NJ). Human sera from allergic patients and purified Jun a 3 from pollen, dotted on the membrane served as positive controls.

Mass spectrometry, N-terminal amino acid sequence determination, and circular dichroism

Peptides reacting with IgE were analyzed by MALDI-MS (Perkin-Elmer-Applied Biosystems (PE-ABI) Voyager instrument) at the Mass Spectrometry Facility at the Louisiana State University Medical Center Core Laboratories in New Orleans. Twenty percent of each HPLC fraction was used. The N-terminal sequence of Jun a 3 and its tryptic peptides were determined using a PE-ABI Procise microsequencer. The CD spectrum was measured on an Aviv spectrophotometer (model 62DS).

Homology modeling and structure refinement

A Blast (Altschul et al., 1990) search of the Protein Data Bank (PDB) (Sussman et al., 1998) with the Jun a 3 sequence as probe yielded two entries: pathogenesis-related protein 5d from tobacco (PDB file *1aun*) and thaumatin from African berry (PDB file *1thv*). These proteins have, respectively, 51.5% and 46.5% sequence identity with Jun a 3 and crystal structures of resolution 1.80 Å and 1.75 Å, making them excellent reference templates for homology modeling. The homology models for Jun a 3 based on each of these templates, referred to as "aun" and "thv," were termed "Jun a 3_aun" and "Jun a 3_thv."

Homology modeling of Jun a 3

The sequence of Jun a 3 was aligned with that of the template protein, with the program CLUSTALW (Higgins et al., 1992) (Fig. 1). The program EXDIS developed in our group (http://www.scsb.utmb.edu/FANTOM/fm_home.html) was used to extract interatomic distance constraints and dihedral angle constraints from the structure of the template. During this process, short stretches corresponding to "gaps" or "loops" in the align-

```
juna3    1  VKFDIKNQCGYTVWAAGLP-GGGKRLDQGQTWTVNLAAGTASARFWGRTGCTFDASGKG  58
aun         GVFEVHNNCPYTVWAAATPVGGGRRLERGQSWWFWAPPGTKMARIWGRTNCNFDGAGRG
              *...*.* ******  * ***.**..**.*      **  ** ****.*.** .*.*

juna3   59 SCQTGDCGGQLSCTVSGAVPATLAEYTQ---SDQDYYDVSLVDGFNIPLAINPTN---AQ 112
aun         WCQTGDCGGVLECKGWGKPPNTLAEYALNQFSNLDFWDISVIDGFNIPMSFGPTKPGPGK
             ******** * *   *  * *****.    *  *. *.*..******.. .**     .

juna3  113 CTAPACKADINAVCPSELKVDGGCNSACNVFKTDQYCCRNAYVDNCPATNYSKIFKNQCP 172
aun         CHGIQCTANINGECPGSLRVPGGCNNPCTTFGGQQYCCTQG---PCGPTELSRWFKQRCP
            *    * * **  **  *.* ****  *. *   **** .     *  *  *. **..**

juna3  173 QAYSYAKDD-TATFACASG-TDYSIVFCP                                199
aun         DAYSYPQDDPTSTFTCTSWTTDYKVMFCP
             **** .** *.**.*.*  *** ..***
```

(a)

```
juna3    1 VKFDIKNQCGYTVWAAGLPG------GGKRLDQGQTWTVNLAAGTASARFWGRTGCTFDA  54
thv         ATFEIVNRCSYTVWAAASKGDAALDAGGRQLNSGESWTINVEPGTNGGKIWARTDCYFDD
              *.* *.* ******   *      **..*  *..**.*.  **   . * ** * **

juna3   55 SGKGSCQTGDCGGQLSCTVSGAVPATLAEYTQSDQ--DYYDVSLVDGFNIPLAINPTNAQ 112
thv         SGSGICKTGDCGGLLRCKRFGRPPTTLAEFSLNQYGKDYIDISNIKGFNVPMDFSPTTRG
            ** * *.****** * *   *  *.****..       ** *.* . ***.*.   **.

juna3  113 CTAPACKADINAVCPSELKVDGG-CNSACNVFKTDQYCCRNAYVDNCPATNYSKIFKNQC 171
thv         CRGVRCAADIVGQCPAKLKAPGGGCNDACTVFQTSEYCCTTG---KCGPTEYSRFFKRLC
            *    * ***   **. **  ** ** **.**.* .***  .     *  * **. **  *

juna3  172 PQAYSYAKDDTATFACASGTDYSIVFCP                                199
thv         PDAFSYVLDKPTTVTCPGSSNYRVTFCP
            * *.**  *  .* .*   . * .  ***
```

(b)

FIGURE 1 CLUSTAL W sequence alignments of Jun a 3 with the templates aun (*a*) and thv (*b*) used for homology modeling. Dashes indicate gaps: dots indicate conservative substitutions.

ment are left out, and constraints are extracted from the remaining "fragments" of the protein (see Table 1 for the fragments of Jun a 3 used). For a given atom, EXDIS selects a specified number of other atoms, chosen randomly, and calculates distances to them. For Jun a 3, specifying 10 constraints per atom, a total of 11,457 distances were extracted from the aun structure, and 11,542 from thv. Each distance was used as an upper and a lower bound for that atom pair, by adding a "tolerance" value of ± 0.1 Å. For dihedral angle constraints, EXDIS uses the following rule at each aligned position: if the amino acids in Jun a 3 and the template are identical, all dihedral angles are read from the template; if they differ, only the backbone dihedral angles are read. Values of unknown dihedral angles are assigned a starting value of 180°. These are converted to ranges by adding $\pm 10°$ to ω and $\pm 15°$ to the other torsion angles. For the two Jun a 3 models, 670 and 667 dihedral constraints were obtained from aun and thv, respectively.

The program FANTOM (Schaumann et al., 1990) was then used with the above constraints to minimize the conformational energy of the protein. FANTOM uses the ECEPP/2 all-atom force field (Abe et al., 1984). The total energy calculated is the sum of the conformational energy (electrostatic + hydrogen bond + Lennard-Jones + torsional energies) and the constraint energy (weighted penalties for violations of dihedral angles + upper + lower distance constraints). First, FANTOM constructs a starting structure of the protein, taking standard geometries from a library and dihedral angles from the template structure. The constraint energy is then calculated, by penalizing deviations from the distance and dihedral angle ranges that have already been set up. The total constraint energy is then minimized, to produce a crude model. This first stage is referred to as "regularization." Both template structures have eight disulfide bridges, and the corresponding residues are conserved in Jun a 3. The resulting disulfide bonds in Jun a 3 are 9–198, 50–60, 65–71, 113–187, 118–171, 126–136, 140–149, and 150–158 (Jun a 3 numbers). There are two X-*cis*-Pro peptide units in aun; the corresponding ones in Jun a 3 (Leu^{18}-Pro^{19} and Val^{77}-Pro^{78}) are modeled in *cis* configuration. In thv there is only one *cis* proline, Pro^{79} (Fig. 1 *b*). Hence in the thv-based model, Pro^{78} is modeled as *cis* and Pro^{19} as *trans*. The disulfide bridges and *cis*-Pro bonds were built at the start of FANTOM runs.

In the next stage, the full energy function was applied and minimized. A fourth-power energy function was used for distance constraints, which added $kT/2$ to the total energy for a violation by 0.2 Å in the regularization stage. This limit was raised in two steps to 1.0 Å by the end of the minimization stage. The distance constraints to the template were thus progressively relaxed. The dihedral angle constraint function added an energy of $10.0*kT/2$ for every 5° violation. The minimization was accomplished by the successive application of quasi-Newton and Newton- Raphson minimizers as implemented in FANTOM (Schaumann et al., 1990).

Continuum electrostatics calculations

The 15 Asp, two Glu, 10 Lys, and four Arg residues of Jun a 3 and the amino- (+1) and carboxy- (−1) terminal residues were considered charged. All of the other residues (including His) were treated as neutral. This charging scheme led to a total charge of −3 on the molecule. The protein was assigned a dielectric constant of 2.0 and the surrounding solvent, 80.0. For this system, the electrostatic potentials were calculated by solving the Poisson-Boltzmann equation by the method of Nicholls and Honig (1991), as implemented in the program MOLMOL (Koradi et al., 1996). MOLMOL then displays the electric potential on the protein's contact surface (Richards, 1977).

RESULTS

Structural and energetic evaluation of the models

Fig. 2 is a stereo view of the two models, showing their α-carbon backbones superimposed according to the sequence alignments given in Fig. 1. We used the program PROCHECK (Morris et al., 1992) to validate the models based on stereochemical and geometric considerations (Table 1). Only five residues were in the disallowed regions of

TABLE 1 Summary of Jun a 3 homology modeling

Property	aun-based model	thv-based model
Modeling details		
Resolution of template structures (Å)	1.80	1.75
Sequence identity	103 (51.5%)	93 (46.5%)
Fragments used by EXDIS*	1–15, 24–84, 89–106, 115–152, 158–199	1–18, 23–87, 92–152, 158–199
Disulfide bridges (in both models)	10–199, 51–61, 66–72, 114–188, 119–172, 127–137, 141–150, 151–159	
X-*cis*-Pro bonds in the model	L19–P20, V78–P79	V78–P79
Number of residues/atoms in the calculation	200/3229	199/3210
Dihedral angle constraints from template	670	667
Upper/lower distance constraints from template	11457	11542
Analysis of the models		
Backbone RMSD (Å) from template	0.633	0.954
Final van der Waals energy (kcal/mol)	−1116	−1067
Final constraint energy (kcal/mol)	389	933
Final conformational energy (kcal/mol)	−1055 (−5.3/residue)	−1010 (−5.1/residue)
Residues with disallowed backbone conformation†	5	2
Residues with ω deviation >20°	0	1
Number of dihedral constraints violated by ≥5°	139 (21%)	153 (23%)
Upper distance constraints violated by ≥1 Å	112 (<1%)	90 (<1%)
Lower distance constraints violated by ≥1 Å	84 (<1%)	56 (<1%)

*See text for definitions and details.
†From the Ramachandran map, as calculated by PROCHECK.

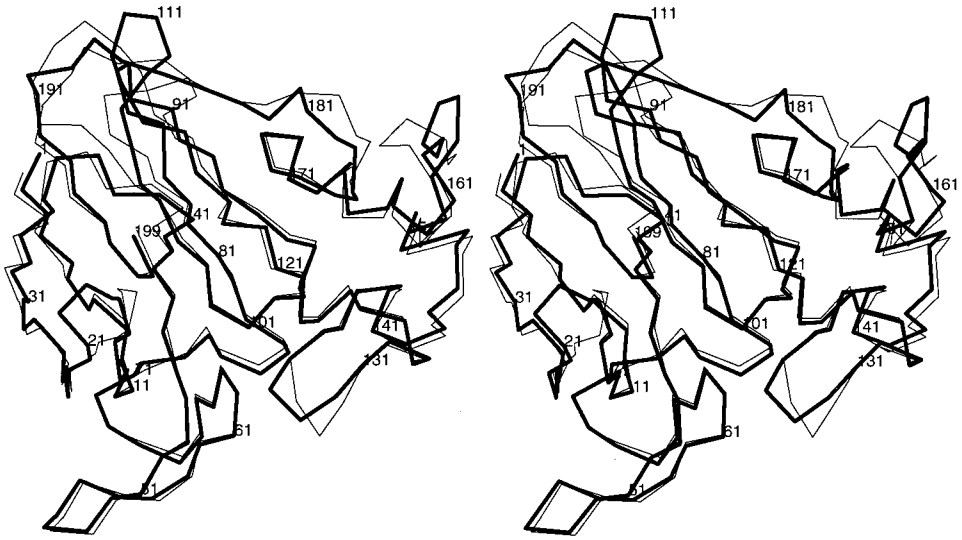

FIGURE 2 Stereo view of the aun-based and thv-based Jun a 3 models (*thick* and *thin lines*, respectively). Some residue numbers are marked. The orientation is the same as in Fig. 7.

the Ramachandran map for Jun a 3 aun, and two for Jun a 3 thv. There were no deviations in the peptide torsion angle ω above 20° for Jun a 3_aun, and only one residue deviated above that value in the thv-based model. Violations of the distance and dihedral angle constraints in the final models were within acceptable limits (Table 1). The conformational and van der Waals energies were negative for both models. These indicators show that both models were structurally and energetically acceptable. The backbone root mean square deviations (RMSDs) of the models from their respective templates were low (0.63 Å (Jun a 3_aun) and 0.95 Å (Jun a 3_thv)), indicating a high degree of structural similarity, as expected from their high degree of sequence identity. In addition, ~80% of the dihedral angles and >99% of the distance constraints extracted by EXDIS from the template structures were conserved in the model structures.

Comparison of the two models

The backbone RMSD of the central structure of the models, excluding the small loop regions where they differ because

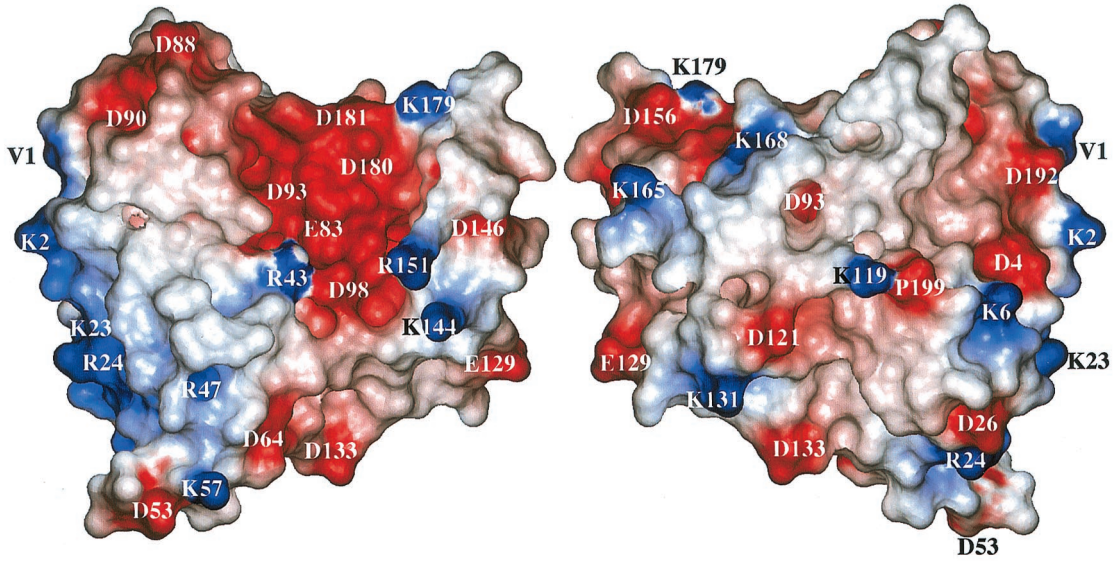

FIGURE 3 Electrostatic potentials at the Jun a 3 surface for the aun model. Blue represents positive potentials and red represents negative potentials. (*Left* and *right*) "Front" (corresponding to the orientation in Fig. 2) and "back" views of the molecule, respectively.

of gapping in the alignment (19–20, 86–90, 107–112, 132–134, 154–156, and 179–181; Fig. 2), is only 0.9 Å, showing that the two models for Jun a 3 are very similar. The segments where the gaps occur are in "loops" that connect secondary structures, but not parts of secondary structures themselves. The backbone RMSD between the two models is 1.9 Å over the whole protein.

Fig. 3 shows the electrostatic potentials at the surface of the Jun a 3 aun model; the labels indicate the approximate location of the charged residues on the surface. The potentials are (qualitatively) consistent with the location of the charged residues. Also note that the amino- and carboxy-terminal residues (Val^{1} and Pro^{200}) are charged in our calculations. A similar calculation based on the thv-based model produced a very similar diagram (results not shown).

Spectral evidence for the secondary structure of Jun a 3

The CD spectrum of Jun a 3 isolated from pollen (Fig. 4) is very similar to that of birch pollen allergen Bet v 1 (Ferreira et al., 1998), a predominantly β-sheet protein, and the Pru a 1 allergen from cherry (Scheurer et al., 1999). Analysis of this spectrum with the program CCA (Perczell et al., 1992; Balasubramanian et al., 1998) indicated that the protein was ~10% helical, ~28–32% β-sheet, and the rest random coil, with a 4% margin of error. This is consistent with our model, where 27/198 (13.6%) amino acids are in α-helices and 59/198 (29.8%) in a β-sheet conformation.

IgE-reactive tryptic peptides

The products of trypsin degradation of Jun a 3 were separated by HPLC (Fig. 5 *A*), and the fractions were tested for reactivity with pooled patient serum IgE by dot blotting (Fig. 5 *B*). The composition of fractions that showed high reactivity with IgE from cedar hypersensitive patients (50, 55, 56, 62, and 65) were analyzed by mass spectrometry (Fig. 6 and Table 2) for peptides specific for Jun a 3. Peptides 120–131 (fractions 62 and 65), 132–145 (fraction 56), and 152–165 (fractions 50 and 55) were determined to be IgE epitopes. A fourth peptide, 169–179, was also identified in the IgE-positive HPLC fraction 55, but not as the sole Jun a 3-derived component of any fraction. Thus IgE binding could not be conclusively attributed to this peptide.

These results were partially confirmed by a separate experiment. Half of the original tryptic digest was fractionated similarly on HPLC, the fractions were checked for reactivity with patient IgE, and positive fractions were analyzed by N-terminal amino acid sequencing (data not shown). While the amount of protein was limiting, peptide 152–165 was detected as the sole peptide in a fraction with which the IgE in patient sera reacted.

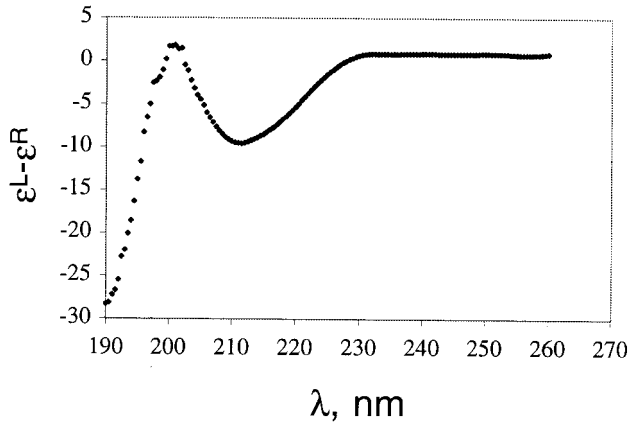

FIGURE 4 CD spectrum of Jun a 3 isolated from pollen. The spectrum was collected at room temperature; the protein concentration was ~0.5 mg/ml ($A_{280} \approx 0.5$) in phosphate-buffered saline.

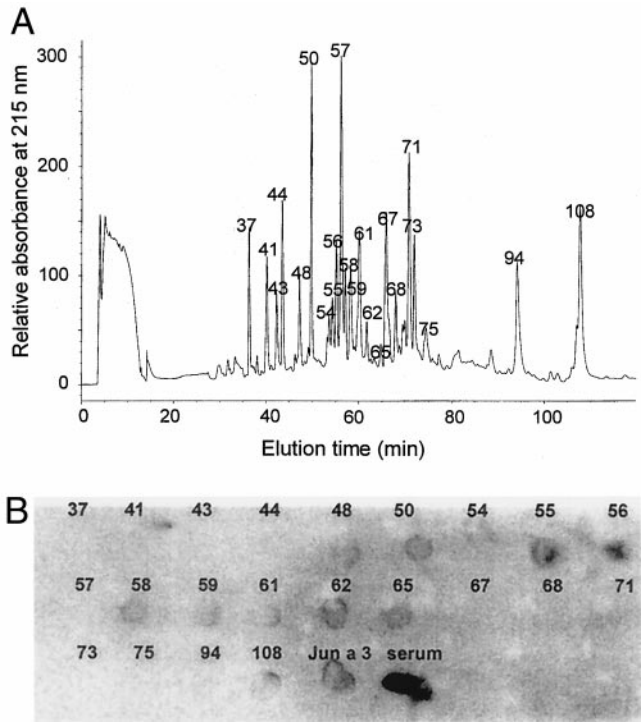

FIGURE 5 Tryptic fragments in fractions of Jun a 3 that react with patient IgE. A tryptic digest of Jun a 3 was fractionated by HPLC (*A*) as described in Materials and Methods, and the fractions were analyzed by immunodotspotting (*B*). The numbers in *B* correspond to the indicated fraction in *A*. The two positive controls are 1 μg of intact Jun a 3 from pollen and 1 μl of patient serum (1:100 dilution) for total IgE, applied with the samples before the membrane was blocked.

Location of epitopes on the protein surface

Fig. 7 is a ribbon and "neon" rendering of Jun a 3 depicting the three positively identified IgE epitopes (red, residues 120–131; gold, 132–145; and blue, 152–165). Peptide 169–

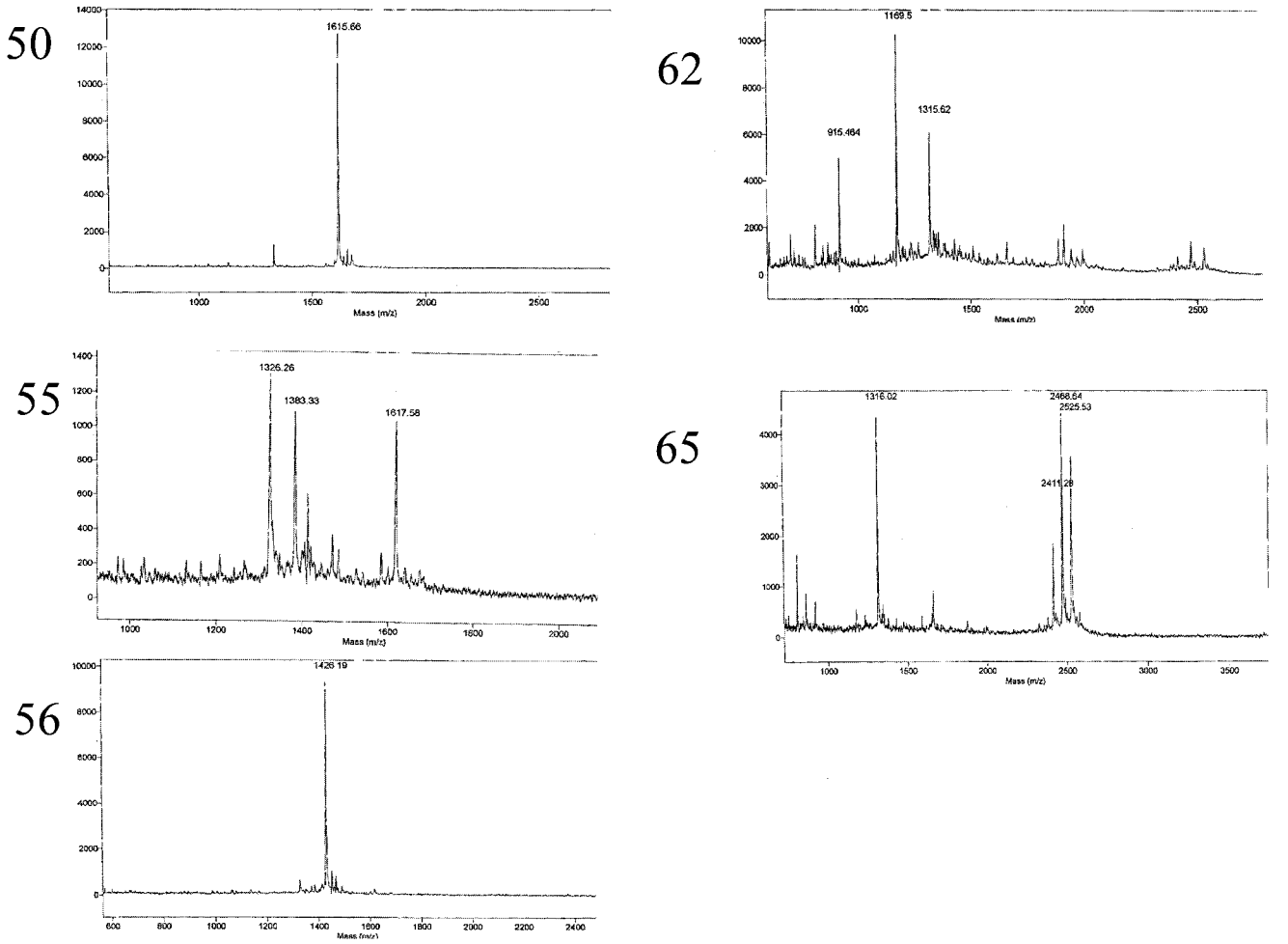

FIGURE 6 Mass spectrometry of HPLC fractions from the tryptic digest of Jun a 3 that bound IgE from pooled patient sera (Fig. 5). The molecular weights of the peptides from Jun a 3 detected in each fraction (fraction number in parentheses) were 1615.66 (50); 1326.26 and 1617.58 (55); 1426.19 (56); 1315.62 (62); and 1316.02 (65). The peaks at 1383.33 (55); 915.46 and 1169.5 (62); and 2411.29, 2468.64, and 2525.53 (65) probably resulted from trace contaminants.

179, which was also identified in fraction 55, is located behind these peptides in an area of lower solvent exposure. The orientation of the molecule is the same as in Figs. 2 and 3. Note that the epitopes are on one face of the protein, accessible for interaction with other macromolecules, such as immunoglobulins. This area maps to the front view of the electrostatic surface in Fig. 3, which indicates that the putative interaction surface is an extensive hydrophobic patch encircled by the charged side chains of Glu^{129}, Lys^{144}, Arg^{151}, Asp^{146}, Lys^{179}, Asp^{156}, and Lys^{165}. We propose that this large, solvent-exposed, hydrophobic area surrounded by charged residues contributes to both the binding affinity and specificity of the interaction with IgE.

DISCUSSION

Jun a 3 is among the first pollen allergens to be characterized as a PR protein, based on sequence identity with members of this family of plant proteins, the expression of which is induced by stress, osmotic shock drought, freezing temperature, infection, or ultraviolet B light. As there were high-resolution crystal structures for two proteins with >40% sequence identity with Jun a 3 (Fig. 1), we were able to prepare detailed model structures, using our in-house programs EXDIS and FANTOM (Fig. 2). As Table 1 shows, few of the ~12,000 angle and distance constraints extracted from either template were violated in the model structures, and the structures are stereochemically acceptable. These model structures are the first reported for this family of allergens and will be deposited in the PDB.

Similarities between the Jun a 3 models and the structures of other known allergens

Allergenic proteins identified to date can be grouped into discrete families based on sequence similarity (Stewart and

TABLE 2 Anticipated tryptic fragments of Jun a 3 in order of decreasing size, their calculated mass, and identification in HPLC fractions reacting positively with patient IgE in dot-blot assay

Sequence of the peptide	Residue numbers	Calculated molecular mass*	HPLC fractions reacting with patient IgE†
GSCQTGDCGGQLSCTVSGAVPATL EYTQSDQDYYDVSLVDGFNIPLAI PTNAQCTAPACK	58–119	6613.95	
DDTATFACASGTDYSIVFCP	180–199	2197.90	
LDQGQTWTVNLAAGTASAR	25–43	1995.99	
NQCGYTVWAAGLPGGGK	7–23	1735.82	
NAYVDNCPATNYSK	152–165	1616.70	50:1615.66; 55:1617.58
VDGGCNSACNVFK	132–144	1427.60	56:1426.19
NQCPQAYSYAK	169–179	1329.59	55:1326.26
ADINAVCPSELK	120–131	1316.65	62:1315.62; 65:1316.02
TGCTFDASGK	48–57	1043.45	
TDQYCCR	145–151	1002.37	
FWGR	44–47	565.29	
FDIK	3–6	522.29	

*Accounts for reduction and alkylation.
†See Fig. 6 for the mass spectra.

Thompson, 1996; Liebers et al., 1996). The new allergens of the PR-5 family, including Jun a 3, share no apparent sequence identity with birch pollen allergen Bet v 1 (Gajhede et al., 1996) or any of the other aerosol allergens for which a 3D structure is available in the PDB. However, as Fig. 8 illustrates, the model has features in common with the larger allergen proteins, the structures of which are available in the PDB (Rouvinen et al., 1999; Arruda et al., 1995). These proteins have a β-sheet core and flexible loop regions on the surface.

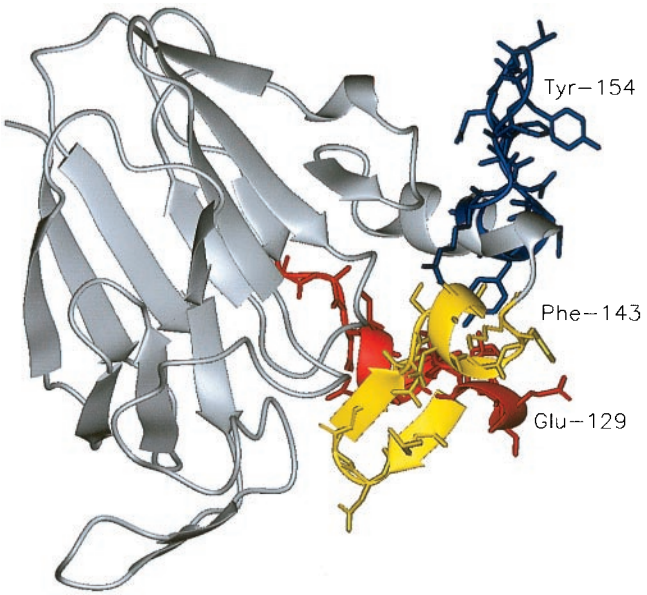

FIGURE 7 The IgE epitopes of Jun a 3 identified by trypsin hydrolysis and dot blotting are on one surface of the model. The backbone of the aun-based model is shown in gray; all nonhydrogen atoms are shown for the three epitopes in red (residues 120–131), gold (132–145), and blue (152–165). The orientation of the model is the same as in Fig. 2.

Although there is some helical character to these loop regions, several of the allergens contain no helix at all. According to SCOP, the structural classification for proteins (Murzin et al., 1995), Phl p 2, Der f 2, Der p 2, Bos d 2, Equ c 1, and mMUP belong to the structural class "all β," and Bet v 1 and Bet v 2 to "$\alpha+\beta$" proteins. Amb t 5 consists of a three-stranded antiparallel β-sheet with a short α-helix packed against it. Thus the overall features of the known structures are similar, despite their lack of sequence similarity.

The structural similarity among these allergens is further emphasized by the CD spectra of Jun a 3 (Fig. 4), which closely resembles that of recombinant Bet v 1 (Ferreira et al., 1998), and the Bet v 1 homologue from cherry, Pru a 1 (Scheurer et al., 1999).

Allergenic epitopes are in a helix-loop region on one face of Jun a 3

The tryptic fragments of Jun a 3 that reacted with patient IgE (Figs. 5 and 6) sufficiently to be detected in the dot-blot assay mapped to one surface-exposed helical/loop region of the model (Fig. 7). The location of these epitopes is consistent with that reported for other aeroallergens. The IgE-binding epitopes in birch pollen profilin (Fedorov et al., 1997a) are clustered at the N and C termini. Both termini are loop areas, proximal to the top and sides of the β-sheet core of the protein. Deletions of loop regions between either Cys^{21} and Cys^{27} or Cys^{73} and Cys^{78} in the mite allergen Der p 2 decreased the binding to IgE from certain patient sera by up to 1000 times (Hakkaart et al., 1998). Both of these areas map to surface loops peripheral to the central β-barrel core in the NMR structure of Der p 2 (Mueller et al., 1998). Mutating a serine residue in Bet v 1 (or its close relative Pru a 1, which should have a similar structure) at the end of a

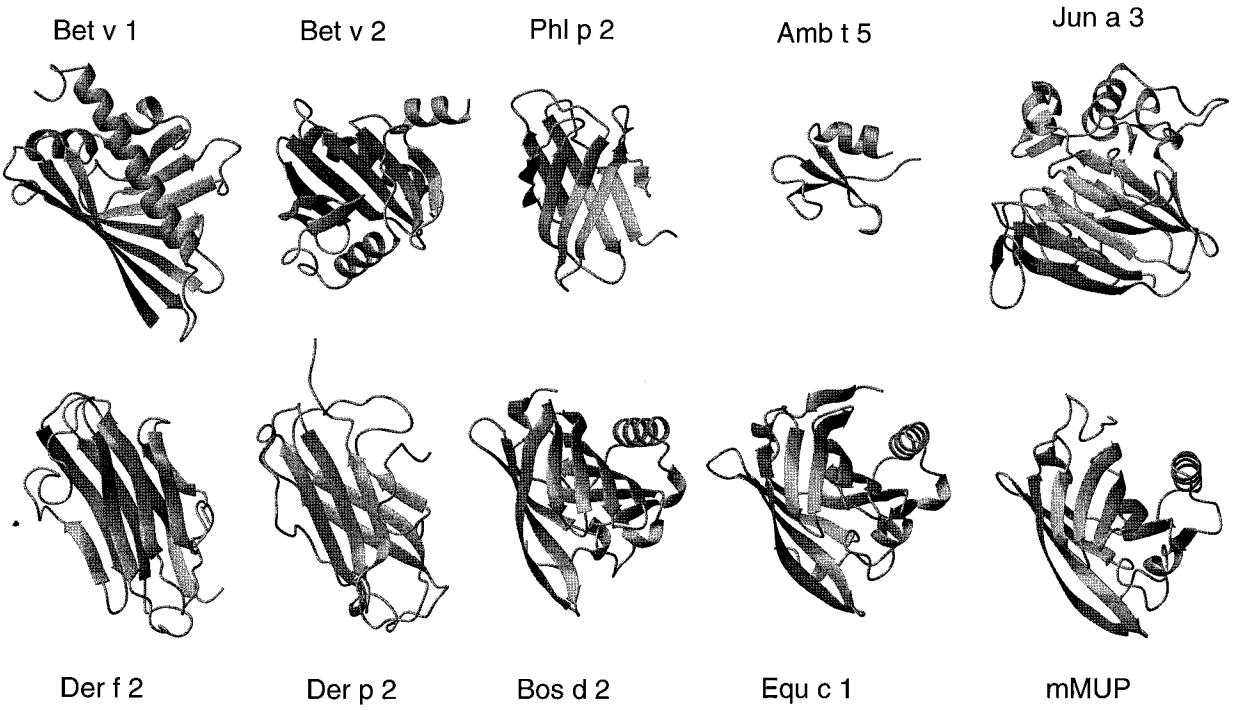

FIGURE 8 Comparison of the Jun a 3 model (this paper) with other allergen structures available in the PDB. The structures are (*in clockwise order from upper left corner*) birch pollen allergen Bet v 1 (crystal, 2-Å resolution, 1BV1) (Gajhede et al., 1996), birch pollen profilin Bet v 2 (crystal, 2.4-Å resolution, 1CQA) (Fedorov et al., 1997a), timothy grass pollen allergen Phl p 2 (crystal, 3.0-Å resolution, 1WHP) (Fedorov et al., 1997b), ragweed pollen allergen Amb t 5 (NMR structure, 1BBG), our Jun a 3 model based on 1AUN, mouse urinary protein Mus m 1 (crystal, 2.4 Å resolution, 1MUP) (Bocskei et al., 1992), horse allergen Equ c 1 (crystal, 2.3-Å resolution, 1EW3) (Lascombe et al., 2000), bovine lipocalin dander allergen Bos d 2 (crystal, 1.8-Å resolution, 1BJ7) (Rouvinen et al., 1999), mite allergen Der p 2 (NMR structure, 129 amino acids, 1A9V) (Mueller et al., 1998), and the mite fecal allergen Der f 2 (NMR structure, 129 amino acids, 1AHK) (Ichikawa et al., 1998). All structures are drawn to the same scale; where the proteins are known to form oligomers (e.g., Equ c 1), only the monomeric form is shown. The front helix in Bet v 1 is ~37 Å long.

loop immediately proximal to the β-sheet core, decreases IgE reactivity by a factor of 10–1000-fold (Scheurer et al., 1999). Mutations that reduced IgG binding were all in one loop of Amb t 5 (Rafner et al., 1998). Although two of the tryptic fragments we identified as IgE epitopes are long enough to assume some secondary structure after immobilization on the dot-blotting membrane surface, our analysis can identify only continuous epitopes. Judging from studies of other allergens (Collins et al., 1996; Mueller et al., 1998; Ichikawa et al., 1998; der Val et al., 1999; Engel et al., 1997), Jun a 3 may have areas of reactivity with IgE that are dependent on an intact 3D structure. For example, six amino acids, widely separated in the crystal structure of Bet v 1, were identified by comparative analysis of the aligned sequences of Bet v 1 and related proteins. Mutations at these positions reduced reactivity with patient IgE and skin-prick test responses (Ferreira et al., 1998). The effects of the mutations were, for the most part, cumulative.

In conclusion, our model structure of the novel allergen Jun a 3 shares many features with those of other allergens. The IgE epitopes mapped to one side of this structure, in a location similar to that seen for other aeroallergens. More detailed understanding of the similarities between the 3D-structures of allergens and their epitopes may provide an approach to the prediction of allergenicity in other proteins and design strategies for the therapeutic control of allergic responses (Valenta et al., 1998).

We thank Lucy Lee (UTMB) for recording the CD spectra of Jun a 3, S. V. Balasubramanian (SUNY, Buffalo, NY) for analysis of the CD spectra with the program CCA, and Steven Smith at the UTMB Protein Chemistry Laboratory for peptide sequence analysis.

This work was supported by the Sealy Center for Structural Biology, a UTMB President's Cabinet Award (to RMG, TM-H, and EGB), the James W. McLaughlin Fellowship Fund (to TM-H), the Child Health Research Center (RMG and TM-H), and grants to WB from the National Science Foundation (DBI-9632326 and DBI-9714937), the U.S. Department of Energy (DE-FG03-96ER62267), and the Texas Advanced Research Program (4952-0084-1999).

REFERENCES

Abe, H., W. Braun, T. Noguti, and N. Go. 1984. Rapid calculation of first and second derivatives of conformational energy with respect to dihedral angles for proteins. General recurrent equations. *Comput. Chem.* 8:239–247.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403–410.

Arruda, L. K., L. D. Vailes, M. L. Hayden, D. C. Benjamin, and M. D. Chapman. 1995. Cloning of cockroach allergen, Bla g 4, identifies ligand binding proteins (or calcynins) as a cause of IgE antibody responses. *J. Biol. Chem.* 270:31196–31201.

Balasubramanian, V., L. Nguyen, S. V. Balasubramanian, and M. Ramanathan. 1998. Interferon gamma inhibitory oligodeoxynucleotides alter the conformation of interferon. *Mol. Pharmacol.* 53:926–932.

Bocskei, Z., C. R. Groom, D. R. Flower, C. E. Wright, S. E. V. Phillips, A. Cavaggioni, J. B. Findlay, and A. C. North. 1992. Pheromone binding to two rodent urinary proteins revealed by x-ray crystallography. *Nature.* 360:186–188.

Collins, S. P., G. Ball, E. Vonarx, C. Hosking, M. Shelton, D. Hill, and M. E. H. Howden. 1996. Absence of continuous epitopes in the house dust mite major allergens *Der* p I from *Dermatophagoides pteronyssinus* and *Der* f I from *Dermatophagoides farinae. Clin. Exp. Allergy.* 26: 36–42.

der Val, G., B. C. Yee, R. M. Lozano, B. B. Buchanan, R. W. Ermel, Y. M. Lee, and O. L. Frick. 1999. Thioredoxin treatment increases digestibility and lowers allergenicity of milk. *J. Allergy Clin. Immunol.* 103: 690–697.

Engel, E., K. Richter, G. Obermeyer, P. Briza, J. Kungl, B. Simon, M. Auer, C. Ebner, H. Rheinberger, M. Breitenbach, and F. Ferreira. 1997. Immunological and biological properties of Bet v 4, a novel birch pollen allergen with two EF-hand calcium-binding domains. *J. Biol. Chem.* 272:28630–28637.

Fedorov, A. A., T. Ball, N. M. Mahoney, R. Valenta, and S. C. Almo. 1997a. The molecular basis for allergen cross-reactivity: crystal structure and IgE epitope mapping of birch pollen profilin. *Structure.* 5:33–45.

Fedorov, A. A., T. Ball, R. Valenta, and S. C. Almo. 1997b. X-ray crystal structures of birch pollen profilin and Phl p 2. *Int. Arch. Allergy Immunol.* 113:109–113.

Ferreira, F., C. Ebner, B. Kramer, G. Casari, P. Briza, R. Grimm, B. Jahn-Schmid, H. Breiteneder, D. Kraft, M. Breitenbach, H. Rheinberger, and O. Scheiner. 1998. Modulation of IgE reactivity of allergens by site-directed mutagenesis: potential use of hypoallergenic variants for immunotherapy. *FASEB J.* 12:231–242.

Gajhede, M., P. Osmark, F. M. Poulsen, H. Ipsen, J. N. Larsen, R. J. J. van Neerven, C. Schou, H. Lowenstein, and M. Spangfort. 1996. X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy. *Nature Struct. Biol.* 3:1040–1045.

Hakkaart, G. A. J., R. C. Aalberse, and R. van Ree. 1998. Epitope mapping of the house-dust-mite allergen Der p 2 by means of site-directed mutagenesis. *Allergy.* 53:165–172.

Higgins, D. G., A. J. Bleasy, and R. Fuchs. 1992. CLUSTAL W: improved software for multiple sequence alignment. *Comput. Appl. Biosci.* 8:189–191.

Ichikawa, S., H. Hatanaka, T. Yuuki, N. Iwamoto, S. Kojima, C. Nishiyama, K. Ogura, Y. Okumura, and F. Inagaki. 1998. Solution structure of Der f2, the major mite allergen for atopic diseases. *J. Biol. Chem.* 273:356–360.

Inschlag, C., K. Hoffmann-Sommergruber, G. O'Riordain, H. Ahorn, C. Ebner, O. Scheiner, and H. Breiteneder. 1998. Biochemical characterization of Pru a 2, a 23-kD thaumatin-like protein representing a potential major allergen in cherry (*Prunus avium*). *Int. Arch. Allergy Immunol.* 116:22–28.

Koradi, R., M. Billeter, and K. Wüthrich. 1996. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* 14: 51–55.

Lascombe, M.-B., C. Gregoire, P. T. Poncet, G. A., I. Rosinski-Chupin, J. Rabillon, H. Goubran-Botros, J.-C. Mazie, B. David, and P. M. Alzari. 2000. Crystal structure of the allergen Equ c 1: a dimeric lipocalin with restricted IgE-reactive epitopes. *J. Biol. Chem.* 275:21572–21577.

Liebers, V., I. Sanders, V. Van Kampen, M. Raulf-Heimsoth, P. Rozynek, and X. Baur. 1996. Overview of denominated allergens. *Clin. Exp. Allergy.* 26:494–516.

Linthorst, H. M. 1981. Pathogenesis-related proteins of plants. *Crit. Rev. Plant Sci.* 10:123–150.

Midoro-Horiuti, T., R. M. Goldblum, A. Kurosky, D. W. Goetz, and E. G. Brooks. 1999a. Isolation and characterization of mountain cedar (*Juniperus ashei*) pollen major allergen, Jun a 1. *J. Allergy Clin. Immunol.* 104:608–612.

Midoro-Horiuti, T., R. M. Goldblum, A. Kurosky, T. G. Wood, and E. G. Brooks. 2000. Identification and expression of an environmentally inducible allergen in mountain cedar (*Juniperus ashei*) pollen. *J. Immunol.* 164:2188–2192.

Midoro-Horiuti, T., R. M. Goldblum, A. Kurosky, T. G. Wood, C. H. Schein, and E. G. Brooks. 1999b. Molecular cloning of the mountain cedar (*Juniperus ashei*) pollen major allergen, Jun a 1. *J. Allergy Clin. Immunol.* 104:613–617.

Morris, A. L., M. W. MacArthur, E. G. Hutchinson, and J. M. Thornton. 1992. Stereochemical quality of protein structure coordinates. *Proteins Struct. Funct. Genet.* 12:345–364.

Mueller, G. A., D. C. Benjamin, and G. S. Rule. 1998. Tertiary structure of the major house dust mite allergen Der p 2: sequential and structural homologies. *Biochemistry.* 37:12707–12714.

Murzin, A., S. Brenner, T. Hubbard, and C. Chothia. 1995. SCOP: a structural classification of protein database for the investigation of sequences and structures. *J. Mol. Biol.* 247:536–540.

Nicholls, A., and B. H. Honig. 1991. A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson-Boltzmann equation. *J. Comput. Chem.* 12:435–445.

Perczell, A., K. Park, and G. D. Fasman. 1992. Analysis of the circular dichroism spectrum of proteins using convex constraint analysis algorithm: a practical guide. *Anal. Biochem.* 203:83–93.

Platts-Mills, T., G. Mueller, and L. Wheatley. 1998. Future directions for allergen immunotherapy. *J. Allergy Clin. Immunol.* 102:335–343.

Rafner, T., M. E. Brummet, D. Bassolino-Klimas, W. J. Metzler, and D. G. Marsh. 1998. Analysis of the three-dimensional antigenic structure of giant ragweed allergen, Amb t 5. *Mol. Immunol.* 35:459–467.

Richards, F. M. 1977. Areas, volumes, packing, and protein structure. *Annu. Rev. Biophys. Bioeng.* 6:151–176.

Rouvinen, J., J. Rautiainen, T. Virtanen, T. Zeiler, J. Kauppinen, A. Taivainen, and R. Mäntyjärvi. 1999. Probing the molecular basis of allergy. Three dimensional structure of the bovine lipocalin allergen Bos d 2. *J. Biol. Chem.* 274:2337–2343.

Schaumann, T., W. Braun, and K. Wuthrich. 1990. The program FANTOM for energy refinement of polypeptides and proteins using a Newton-Raphson minimizer in torsion angle space. *Biopolymers.* 29:679–694.

Scheurer, S., D. Y. Son, M. Boehm, F. Karamloo, S. Franke, A. Hoffmann, D. Haustein, and S. Vieths. 1999. Cross-reactivity and epitope analysis of Pru a 1, the major cherry allergen. *Mol. Immunol.* 36:155–167.

Stewart, G. A., and P. J. Thompson. 1996. The biochemistry of common allergens. *Clin. Exp. Allergy.* 26:1020–1044.

Sussman, J. L., D. Lin, J. Jiang, N. O. Manning, J. Prilusky, O. Ritter, and E. E. Abola. 1998. Protein Data Bank (PDB): database of three-dimensional structural information of biological macromolecules. *Acta Crystallogr.* D54:1078–1084.

Valenta, R., S. Almo, T. Ball, C. Dolecek, P. Steinberger, S. Laffer, P. Eibensteiner, S. Flicker, S. Vrtala, S. Spitzauer, P. Valent, S. Denepoux, D. Kraft, J. Banchereau, and S. Lebecque. 1998. The immunoglobulin E-allergen interaction: a target for therapy of type I allergic diseases. *Int. Arch. Allergy Immunol.* 116:167–176.