

Molecular and Crystal Properties of Bos d 2, an Allergenic Protein of the Lipocalin Family

J. Rautiainen,* S. Auriola,† J. Rouvinen,‡ J. Kauppinen,* T. Zeiler,* D. Novikov,§ T. Virtanen,* and R. A. Mäntyjärvi*,¹

*Department of Clinical Microbiology, University of Kuopio, POB 1627, 70211 Kuopio, Finland;

†Department of Pharmaceutical Chemistry, University of Kuopio, POB 1627, 70211 Kuopio, Finland;

‡Department of Chemistry, University of Joensuu, POB 111, 80101 Joensuu, Finland; and

§Department of Biochemistry and Biocenter Oulu, University of Oulu, 90570 Oulu, Finland

Received May 14, 1998

The relationship between the molecular structure of allergenic proteins and the allergenic determinants is one of the central issues in allergology. We report here that the natural preparation of Bos d 2, a mammalian lipocalin allergen, comprises three molecular variant proteins of 17,829, 17,781, and 17,800 Da. When cDNA of Bos d 2 (Genome Sequence Data Base No. L42867) was recloned and expressed in *Pichia pastoris*, two proteins were produced. One of the proteins (17,831 Da) and the proteins in the natural preparation had pyroglutamate as the N-terminal residue; in the other (17,849 Da) the N-terminal residue was glutamine. Recombinant Bos d 2 protein was crystallized and the native data set was collected at 1.8 Å resolution. © 1998

Academic Press

Lipocalins are a group of functionally and structurally related proteins (1). Although the sequence homology between the members of the family is generally low (10 to 20%), short stretches of conserved sequences can be identified in the molecules. Lipocalins have been roughly divided into core and outlier lipocalins, the former containing three regions of conserved sequences. Lipocalins are found in several species, and they function as carriers of pheromones, odorants and other small lipophilic molecules.

In recent years, it has become evident that a number of important mammalian allergens are members of the lipocalin family. These include Mus m 1 of mouse, Rat n 1 of rat, Equ c 1 of horse, and dog allergens Can f 1 and Can f 2 (2). We have previously shown that Bos d 2, the main allergen in cow dander, is also a lipocalin (3). It is produced by the apocrine sweat glands of skin

and it probably acts as a pheromone carrier (4). It is a potent allergen, up to 97% of patients presenting with cattle-associated asthma have IgE antibodies against natural Bos d 2 (5).

Binding of IgE antibodies is an important characteristic of an allergen. Antibodies, including IgE antibodies in case of allergic diseases, react with a protein antigen at sites defined by the three-dimensional structure of the molecule (6). These conformational epitopes comprise amino acid residues brought together on the surface of the molecule. Therefore, the information of the molecular properties of an allergen is important for the rational development of recombinant derivatives of the allergen suitable for new forms of immunotherapy.

In our previous papers we have described properties of natural Bos d 2 purified from cow dander (7) and cloned and sequenced the corresponding cDNA, later designated by the Allergen Nomenclature Committee of the International Union of Immunological Societies as Bos d 2 (8). In this paper we report a more detailed analysis of natural Bos d 2 which led us to produce recombinant Bos d 2 in *Pichia pastoris* yeast. The recombinant Bos d 2 was purified to the level allowing crystallization of the molecule.

MATERIALS AND METHODS

Affinity purification and anion exchange chromatography. Crude extract of bovine dander was prepared as described previously (9). Affinity chromatography with monoclonal anti-Bos d 2 antibodies was used to purify natural Bos d 2 from the extract (7).

For anion exchange chromatography, protein samples in 20 mM bis-Tris, pH 6.0 were applied to a Resource Q column (Pharmacia Biotech, Uppsala, Sweden) equilibrated with the same buffer. The bound proteins were eluted with a linear 40-ml gradient from 0 to 0.3 M NaCl in 20 mM bis-Tris, pH 6.0 and fractions of 1.0 ml were collected.

Immunoblotting. Protein fractions collected from anion exchange chromatography were immuno-blotted after separation in 15% so-

¹ To whom correspondence should be addressed. Fax +358-17-162705. E-mail: Rauno.Mantjarvi@uku.fi.

dium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) as described before (10). The separated components were transferred to a Hybond-ECL nitrocellulose membrane (0.45 μ m; Amersham, Little Chalfont, UK). Membrane strips were incubated overnight at room temperature with sera from cow-asthmatic persons (diluted 1:10) or with monoclonal antibodies (1:200 dilution of hybridoma culture supernatants). The strips were then washed in Tween-PBS four times for 5 min and incubated for 1 h with 1:1,000 diluted peroxidase labelled rabbit anti-human IgE (DAKO P295; Daco, Glostrup, Denmark), or with 1:1,000 diluted rabbit anti-mouse Ig (DAKO P161). After washing, the reactive bands were visualized with ECL detection reagents (Amersham).

HPLC-electrospray ionization mass spectrometry. Positive ion mass spectra were acquired with on-line high performance liquid chromatography-electrospray ionization mass spectrometry (HPLC-ESI-MS). The measurements were carried out using a LCQ quadrupole ion trap mass spectrometer equipped with an ESI ion source (Finnigan MAT, San Jose, CA), a Rheos 4000 HPLC pump (Flux instruments, Danderyd, Sweden) and a Rheodyne 7725 injector with 20 μ l loop (Cotati, CA). The spray was stabilized using nitrogen sheet gas flow (value 60 to 85) and the spray needle potential was set to 5 kV. The stainless steel inlet capillary was heated to 200°C.

The protein fractions from anion exchange chromatography were diluted to a concentration of 100 μ g/ml and 20 μ l were injected to reversed phase HPLC (Purospher C18e, 3 $\times$ 125 mm, 5 μ m, Merck, Darmstadt, Germany). The proteins were separated using a gradient from 5 to 70 % of acetonitrile (containing 50 mM formic acid) in 30 min. The full scan mass spectra from m/z 800 to m/z 2000 were measured using 500 ms for collection of the ions in the trap; 3 microscans were summed. Fraction B was further purified using trifluoroacetic acid as modifier. In this experiment, the flow from HPLC column was split between ion source and manual fraction collection.

Proteins were digested with trypsin as described before (7). The tryptic digests of the proteins (20 μ l) were separated using a Syncropak RP8 reversed phase column (2 $\times$ 100 mm, 5 μ m, 300 Å, Synchrom, Lafayette, IN). The full scan mass spectra of the peptides from m/z 300 to m/z 2000 were measured using 500 ms for collection of the ions in the trap; 1 to 3 microscans were summed. Collision induced dissociation (CID) MS-MS spectra of the peptides were obtained using the dependent scan mode of the instrument. Full scan MS data were collected until the base peak ion exceeded a present intensity threshold (2×10^4). This triggers acquisition of a CID MS-MS product ion spectrum of the ion. The collision energy was 35 to 40% and isolation width was 2 mass units. Scan range was dependent on the parent mass ion.

Production of recombinant Bos d 2 in *Pichia pastoris*. cDNA of Bos d 2 (Genome Sequence Data Base No. L42867) in a phage library (3) was used as a template for PCR. The primers were designed to include the leader sequence of Bos d 2 as well as EcoRI sites for cloning in the pHIL-D2 expression vector of *Pichia pastoris* (Invitrogen, Leek, the Netherlands). The sequences of the primers were 5'-TTA GAA TTC GCC ATG AAG GCC GTG TTC CTG-3' and 5'-TTA GAA TTC CTT ATG GAG GAC AAT TGT CTG-3'. The plasmid construction was performed with standard techniques (Invitrogen), and the correct insertion into the plasmid was verified by sequencing. The recombinant plasmid was transformed in *Pichia pastoris* strain GS115 (Invitrogen) by electroporation (Gene Pulser II, Bio-Rad, Hercules, CA).

Seed cultures of the transformed *Pichia* yeast were grown in a salt-based medium (GMG; Invitrogen) containing 0.5% (v/v) glycerol on a rotating incubator at 30°C. For the production of the recombinant allergen, yeast cells were collected by centrifugation and resuspended in one-fifth of the same medium base without glycerol, methanol was added to 0.5% (v/v) and the incubation continued for 6 days with daily adding of methanol. The supernatants of the cultures were harvested by centrifugation and the recombinant Bos d 2 protein was purified with a procedure comprising affinity chromatography, anion exchange chromatography and gel filtration as described above.

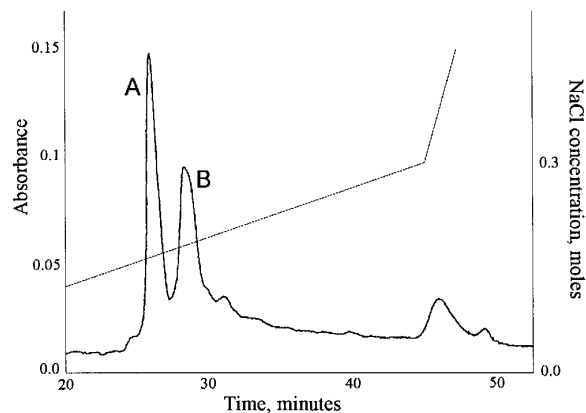


FIG. 1. Anion exchange chromatography profile of affinity-purified natural Bos d 2. Proteins in peaks A and B were collected separately for further analysis.

Crystallization of recombinant Bos d 2. The initial crystallization conditions were found by varying pH and PEG6000 as precipitant. The best crystals were obtained by using the hanging drop method at room temperature. The crystallization drop contained 2 μ l of protein solution (concentration 50 mg/ml) and 2 μ l of reservoir solution (10% (w/w) PEG6000, 0.1 M sodium acetate, pH 3.4-4.5). Crystals were mounted in a thin-walled glass capillary, and the X-ray data set was collected using R-Axis IIC area detector mounted on a Rigaku RU200HB X-ray source using CuK α radiation (50 kV, 180 mA) at 20°C.

RESULTS

Although a highly purified preparation of natural Bos d 2 was obtained by affinity chromatography followed by gel filtration, traces of other proteins were observed by two-dimensional immunoelectrophoresis (11). When the affinity-purified natural Bos d 2 was subjected to anion exchange chromatography, two major peaks were observed (Fig. 1). Immunological analysis of the corresponding proteins revealed that both reacted with two monoclonal antibodies recognizing different epitopes (11) and with IgE antibodies in the sera of two allergic patients (data not shown).

Proteins in peaks A and B obtained by anion exchange chromatography were next analysed by HPLC-ESI-MS. Peak A contained one major protein (Fig. 2). The ESI mass spectrum of this protein showed multiply protonated ions with 10 to 16 charges. The reconstructed molecular mass of the protein was 17,829 Da. On the other hand, two major proteins of 17,781 Da (B1) and 17,800 Da (B2) were discerned from the peak B. The separation of the peaks was enhanced by using trifluoroacetic acid which allowed baseline separation of B1 and B2 and collection of pure fractions for tryptic digestion (data not shown).

After tryptic digestion of the anion exchange peak A, ten peptides were separated and identified by using HPLC-ESI-MS and CID MS-MS fragmentation (Fig. 3). The main peak at retention time of 10.26 min in

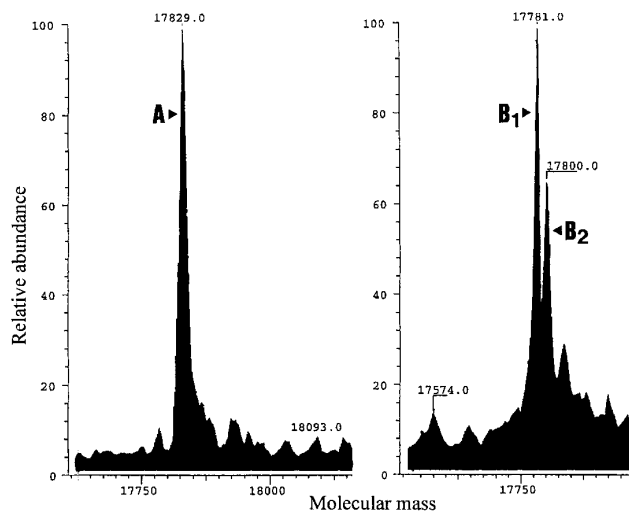


FIG. 2. Reconstructed electrospray ionization mass spectra of the two proteins detected in natural Bos d 2 by anion exchange chromatography. Left panel, peak A from the anion exchange chromatography; right panel, peak B from the anion exchange chromatography.

HPLC had a measured monoisotopic molecular mass of 1,935 Da and was identified with CID MS-MS as the N-terminal peptide of Bos d 2 with pyroglutamic acid (pyrrolidone carboxylic acid) modification (Fig. 4). The full scan CID mass spectra show mainly Y and B-type ions (12), which is typical to low energy collision mass spectra. The rest of the peptide sequences were also compatible with the sequence of cDNA of Bos d 2 covering 76% of the molecule. Furthermore, the theoretical molecular mass of Bos d 2 calculated from the published sequence with the pyroglutamate modification (17,832 Da or 17,830 Da with one or two sulfide bridges, respectively) is very close to that of peak A. Peptide sequencing of peaks B1 and B2 produced less clear-cut results. Peptides matching the sequence of Bos d 2 were present (amino acids 1-40 and 113-138) but the distribution of peptides was different from that of peak A.

The differences between Bos d 2 (peak A) and proteins B1 and B2 are apparently associated with amino acids 41-112 or 138-156.

Since one of our aims was to produce Bos d 2 in pure enough form for crystallization, we decided to reclone the cDNA into an expression vector of *Pichia pastoris*. By the procedure described in the Methods we were able to produce mg amounts of rBos d2 secreted in the growth medium by the yeast. Recombinant Bos d 2 was subjected to a purification procedure of affinity chromatography and anion exchange chromatography. A double peak was observed instead of a homogenous recombinant protein (Data not shown). Both peaks contained proteins which in immunoblotting reacted with patients' IgE antibodies. When the peaks were analyzed further by HPLC and mass spectrometry, the proteins from the peaks had molecular masses of 17,831 Da and 17,849 Da, respectively. HPLC-mass spectra of the tryptic digests showed that the larger protein had an N-terminal peptide beginning with a glutamin residue. The smaller protein had N-terminal sequence identical to that of the natural Bos d 2 containing pyroglutamate as the first residue. The proteins in the two peaks were collected separately for the crystallization experiments.

There was no difference between the two forms of recombinant Bos d 2 in crystallization, both produced well-shaped crystals within one week. The crystals resembled a tetragonal bipyramid with dimensions of approximately $0.15 \times 0.15 \times 0.15$ mm. A single crystal was used for determining crystal parameters and for collecting a native data set at 1.8 Å resolution. The reflections were collected within 45 frames of 2 degrees oscillation. The data was processed with the DENZO program and scaled with SCALEPACK (13). The first scaling was done in the space group P_1 and XPREP program (part of the SHELXTL program package) was used to determine the space group. The second scaling was done in the space group $P2_12_12_1$ as suggested by

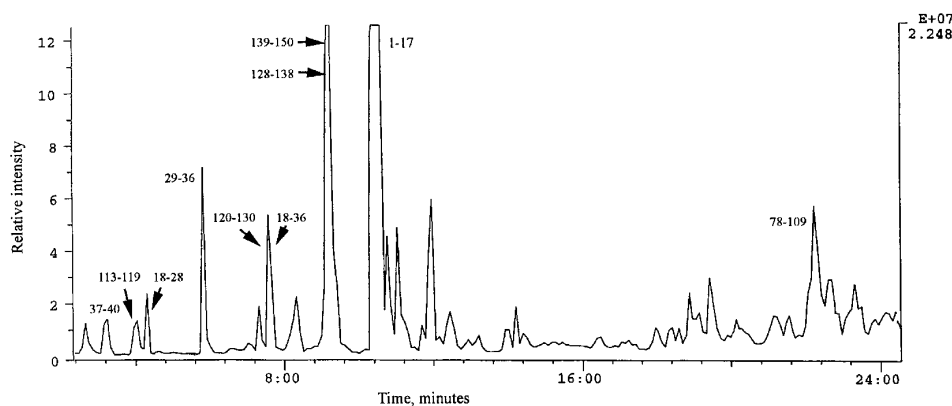


FIG. 3. HPLC-electrospray ionization mass spectroscopy base peak ion chromatogram of the tryptic digest of the protein from anion exchange chromatography peak A. Numbers next to the peaks indicate amino acids of the peptides which match with the sequence of Bos d 2.

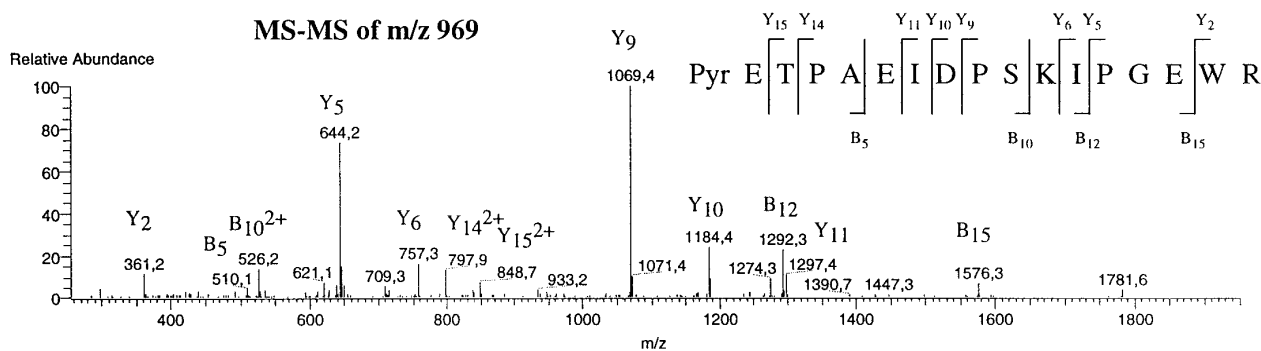


FIG. 4. Full scan electrospray ionization mass spectrum of the presumed N-terminal peptide of Bos d 2. The doubly protonated molecule ($M + 2H$)²⁺ at m/z 969 was selected for collision induced dissociation. The fragment ions are labelled according to Biemann (12). Pyr, pyroglutamic acid (pyrrolidone carboxylic acid) residue.

the XPREP program. The measured Bos d 2 crystal had the cell dimensions of $a = 37.02$ Å, $b = 55.62$ Å and $c = 77.19$ Å. Assuming one Bos d 2 molecule in an asymmetric unit, the specific volume of protein would be 2.23 Å³/Da. This value is close to the average for crystals in proteins observed by Matthews (14). The final merged data set consisted of 41,606 measurements of 14,230 unique reflections where $R_{\text{merge}} = 6.6\%$. The data set was 92.3 % complete at 1.8 Å resolution (completeness 60.3% in the shell 1.8-1.86 Å).

DISCUSSION

We have earlier cloned and published the sequence of the predominant bovine dander allergen (3) which was later designated Bos d 2 (8). A more detailed analysis, described in this report, of the corresponding allergen isolated from cow dander confirmed our earlier surmise that the natural allergen is in fact a group of related proteins. The current results showed convincingly that the natural allergen comprises at least two other proteins in addition to the Bos d 2 defined by the published sequence. The situation is similar to that of mouse allergens. A group of highly homologous lipocalins of mouse (major urinary proteins or MUPs or α_{2u} -globulins) is encoded by a family of 35 to 40 genes (15). Although called urinary proteins, MUPs are present in several types of secretions as proteins differing from each other by a variable number of amino acids (15). In the same way, several isoforms of rat α_{2u} -globulin, another rodent lipocalin allergen, have been identified (16). Bos d 2 was originally characterized as a dander allergen, but immunologically related proteins are also present in bovine urine (17). Bos d 2 is apparently a representative of a group of lipocalin proteins corresponding to the MUPs of mouse and rat.

In the allergen nomenclature, Mus m 1 is the designation allocated for allergens represented by MUPs (18). The realization that allergenicity is often shared by several closely similar molecules led to the definition

of isoallergen (18). Although the immunological characteristics of the proteins related to Bos d 2 has not yet been determined in detail, it is likely that they are in fact isoallergens of Bos d 2.

The crystal form obtained from recombinant Bos d 2 is suitable for determining the three-dimensional structure of Bos d 2 at a high resolution. Because of the low level of identity of Bos d 2 with other members of lipocalin family proteins of known three-dimensional structure (19), we are currently searching for heavy atom derivatives to solve the phase problem.

The detailed structural analysis of allergenic proteins is becoming more and more important because of the need for the development of new, more effective and safer reagents for immunotherapy. The recombinant protein technology offers a possibility to design and to produce modified allergen derivatives for that purpose. Examples of approaches include the use of selected peptides or reconstructed allergen molecules (20,21). Bos d 2 is a representative of the lipocalin family of proteins which includes several mammalian allergens of worldwide importance. The information on the molecular structure and immunological properties of Bos d 2 may be applicable to other lipocalin allergens when strategies for immunotherapies are being outlined.

ACKNOWLEDGMENTS

This study was supported by the Academy of Finland, the Finnish Allergy Research Foundation, the Finnish Cultural Foundation of Northern Savo, the Ida Montin Foundation, the Maud Kuistila Foundation, and Kuopio University Hospital (Project No. 5035). The skillful technical assistance of Riitta Korhonen, Virpi Fisk, and Reetta Kallio and the helpful advice of Kari Koivuranta, Ph.D., are gratefully acknowledged.

REFERENCES

1. Flower, D. R. (1996) *Biochem. J.* **318**, 1–14.
2. Stewart, G. A., and Thompson, P. J. (1996) *Clin. Exp. Allergy* **26**, 1020–1044.

3. Mäntyjärvi, R., Parkkinen, S., Rytönen, M., Pentikäinen, J., Pelkonen, J., Rautiainen, J., Zeiler, T., and Virtanen, T. (1996) *J. Allergy Clin. Immunol.* **97**, 1297–1303.
4. Rautiainen, J., Rytönen, M., Syrjänen, K., Pentikäinen, J., Zeiler, Th., Virtanen, T., and Mäntyjärvi, R. (1998) *J. Allergy Clin. Immunol.*, in press.
5. Ylönen, J., Mäntyjärvi, R., Taivainen, A., and Virtanen T. (1992) *Int. Arch. Allergy Immunol.* **99**, 112–117.
6. Berzofsky, J. A., and Berkower, I. J. (1993) in *Fundamental Immunology* (Paul, W. E., Ed.), 3rd ed., pp. 235–282, Raven Press, New York, NY.
7. Rautiainen, J., Pentikäinen, J., Rytönen, M., Linnala-Kankkunen, A., Pelkonen, J., Virtanen, T., and Mäntyjärvi, R. (1996) *Allergy* **51**, 378–382.
8. Rautiainen, J., Rytönen, M., Virtanen, T., Pentikäinen, J., Zeiler, T., and Mäntyjärvi, R. (1997) *J. Allergy Clin. Immunol.* **100**, 251–252.
9. Ylönen, J., Virtanen, T., Horsmanheimo, L., Parkkinen, S., Pelkonen, J., and Mäntyjärvi, R. (1994) *J. Allergy Clin. Immunol.* **93**, 851–858.
10. Virtanen, T., Zeiler, T., Rautiainen, J., Taivainen, A., Pentikäinen, J., Rytönen, M., Parkkinen, S., Pelkonen, J., and Mäntyjärvi, R. (1996) *Clin. Exp. Allergy* **26**, 188–196.
11. Ylönen, J., Virtanen, T., Rytönen, M., and Mäntyjärvi, R. (1994) *Allergy* **49**, 707–712.
12. Biemann, K. (1988) *Biomed. Environmental Mass Spectrometry* **16**, 99–111.
13. Otwinowski, Z. (1993) in *Data Collection and Processing* (Sawyer, L., Isaacs, N. W., and Bailey, S., Eds.), pp. 71–79, Daresbury Laboratory, Warrington, DC.
14. Matthews, B. W. (1968) *J. Mol. Biol.* **33**, 491–497.
15. Shahan, K., Gilmartin, M., and Derman, E. (1987) *Mol. Cell. Biol.* **7**, 1938–1946.
16. Bayard, C., Holmquist, L., and Vesterberg, O. (1996) *Biochim. Biophys. Acta* **1290**, 129–134.
17. Ylönen, J., Mäntyjärvi, R., Taivainen, A., and Virtanen, T. (1992) *Clin. Exp. Allergy* **22**, 83–90.
18. King, T. P., Hoffman, D., Löwenstein, H., Marsh, D. G., Platts-Mills, T. A. E., and Thomas, W. (1994) *Int. Arch. Allergy Immunol.* **105**, 224–233.
19. Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) *J. Mol. Biol.* **247**, 536–540.
20. Norman, P. S., Nicodemus, C. F., Creticos, P. S., Wood, R. A., Eggleston, P. A., Lichtenstein, L. M., Kagey-Sobotka, A., and Proud, D. (1997) *Int. Arch. Allergy Immunol.* **113**, 224–226.
21. Simons, F. E. R., Imada, M., Li, Y., Watson, W. T. A., and Hayglass, K. T. (1996) *Int. Immunol.* **8**, 1937–1945.