ELSEVIER

Journal of Chromatography B, 763 (2001) 91–98

JOURNAL OF CHROMATOGRAPHY B

www.elsevier.com/locate/chromb

Two new variants of the lipocalin allergen Bos d 2

Jaakko Rautiainen[a,*], Seppo Auriola[b], Anita Konttinen[b], Tuomas Virtanen[a], Marja Rytkönen-Nissinen[a], Thomas Zeiler[c], Rauno Mäntyjärvi[a]

[a]*The Department of Clinical Microbiology, University of Kuopio, P.O. Box 1627, 70211 Kuopio, Finland*
[b]*The Department of Pharmaceutical Chemistry, University of Kuopio, Kuopio, Finland*
[c]*Laboratory of Clinical Microbiology, Kuopio University Hospital, Kuopio, Finland*

Received 11 April 2001; received in revised form 24 July 2001; accepted 3 August 2001

Abstract

Allergens from various sources have been shown to comprise several isoforms. In the present study, a series of chromatographic steps was carried out to separate the lipocalin allergen Bos d 2 isoforms present in cow dander. Subsequent HPLC-MS–MS analyses revealed two new Bos d 2 variants. In one of the proteins, tyrosine (Y83) was substituted by aspartic acid, and in the other protein valine (V102) was replaced by alanine. We propose the three Bos d 2 variants be named as Bos d 2.0101 (previously sequenced Bos d 2), Bos d 2.0102 and Bos d 2.0103. Our results suggest that molecular polymorphism is a common property among lipocalin allergens. Since allergen isoforms may show variation in their IgE binding and/or T-cell reactivity, all of the many allergen forms should be taken into account when planning preparations for immunotherapy.

Keywords: Lipocalins; Allergen

1. Introduction

Recent studies applying recombinant protein technology have revealed the existence of extensive polymorphism in many allergens. Multiple forms of several allergens have been discovered by amino acid sequence analyses, most of which have been added to the official list of isoallergens [http://www.allergen.org/Isoall.htm, Apr 4, 2001] as isoallergens or variants as defined by the Allergen Nomenclature Subcommittee [1]. In particular, pollen allergens have been shown to exist in many isoforms [2–4]. Polymorphism of invertebrate allergens (mites and insects) is also well known [5–7]. The effects of the sequence polymorphism of allergens, expressed as differences in the immunological properties, have to be considered in the design of recombinant allergen preparations for diagnostic tests or for use as vaccines.

Lipocalins are a family of functionally and structurally related proteins. Typically they are transport proteins, able to bind small hydrophobic molecules such as retinol and pheromones. Recent studies have shown that a number of important mammalian allergens are members of this family. Aeroallergens from dog, horse, cow, mouse and rat (Can f 1, Can f 2, Equ c 1, Bos d 2, Mus m 1 and Rat n 1) as well as

*Corresponding author. Tel.: +358-17-162-712; fax: +358-17-162-705.
E-mail address: jaakko.rautiainen@uku.fi (J. Rautiainen).

0378-4347/01/$ – see front matter
PII: S0378-4347(01)00369-3

Bos d 5 or β-lactoglobulin of cow's milk are all lipocalins. In addition to the sequence data, the three-dimensional structures of Equ c 1, Mus m 1, Bos d 2 and Bos d 5 have been determined by X-ray crystallography [8–11]. On the other hand, immunological data of the molecules remains rather limited. T-cell epitope mapping results of Bos d 2 [12] and Bos d 5 [13] have been reported, but information of IgE epitopes is available only for Bos d 5 and Equ c 1 [8,14,15]. A significant decrease in IgE binding of Bos d 2 was seen with the mutation C154A which disrupts formation of an intracellular disulfide bridge [16]. This emphasizes the importance of the intact three-dimensional structure for IgE binding.

In contrast to pollen and mite allergens, there is limited information of isoallergens or variants in many of the source materials of lipocalin allergens. However, sequence analyses have revealed that there are multiple forms of rat and mouse major urinary proteins (MUPs) [http://www.expasy.ch/cgi-bin/nicesite.pl?PS00213, Apr 4, 2001]. Two of the rat MUPs have been designated as lipocalin allergens Rat n 1.01 and Rat n 1.02 [17]. Mouse MUP6 has been defined as corresponding to Mus m 1 [http://www.expasy.ch/cgi-bin/niceprot.pl?MUP6_MOUSE, Apr 4, 2001]. Because of the high level of sequence identity between the different mouse MUPs, they would be predicted also to be allergenic and to represent isoforms or variants of Mus m 1. Two allergen variants of Equ c 2, designated Equ c 2.0101 and Equ c 2.0102 have been found by partial sequencing of the corresponding natural proteins [18]. Sequence variants of Bos d 5 have also been described but not registered as allergen isoforms [http://www.expasy.ch/cgi-bin/niceprot.pl?P02754, Apr 4, 2001].

We have previously cloned the cDNA of a bovine dander allergen, subsequently designated as Bos d 2 [19,20]. When the natural Bos d 2 was purified by a combination of techniques including affinity chromatography, two additional proteins co-purified with Bos d 2 [21]. These proteins were assumed to be isoallergic forms of Bos d 2. In this study, we have determined the differences of the three proteins at the amino acid sequence level with HPLC–MS–MS analysis. Based on these results, we propose that the sequences represent three lipocalin allergen variants, Bos d 2.0101 (previously cloned and sequenced Bos d 2), Bos d 2.0102 and Bos d 2.0103.

2. Experimental

2.1. HPLC–electrospray-mass spectrometry

Mass spectra were recorded using a LCQ quadrupole ion trap mass spectrometer (Finnigan, San Jose, CA). The instrument was tuned with standard 10 μg/ml peptide solution in 30% acetonitrile and 50 m*M* HCOOH (MRFA-peptide, Finnigan). The spray needle was set to 4.5–5.5 kV in the positive ion mode and the spray was stabilized by a nitrogen sheath flow, the value was set to 95–100 (arbitrary units). The inlet capillary temperature was 200–225°C. The high-performance liquid chromatography (HPLC) system consisted of a Rheos 4000 pump (Flux instruments, Danderyd, Sweden) and a LaChrom L-7200 autosampler (Merck–Hitachi, Tokyo, Japan).

2.2. Protein and peptide analysis

The proteins were purified with immunoaffinity chromatography, anion-exchange chromatography and reversed-phase chromatography as described by Rautiainen et al. [21]. However, the reversed-phase chromatography was slightly modified to enhance separation: the column used in this study was a Sephasil Protein C_4, 5 μm ST4.6/250, 300A, Pharmacia Biotech, Uppsala, Sweden), the gradient was 0–100% acetonitrile in 40 min, 0.1% trifluoroacetic acid was added both in water and acetonitrile.

The analytical procedure is described in Fig. 1. The procedure with bovine dander extract is started with affinity purification, anion-exchange chromatography, and reversed-phase HPLC separation of proteins. The molecular weights of the undigested proteins were determined with high-performance liquid chromatography–electrospray ionization mass spectrometry (HPLC–ESI-MS). The purified proteins were reduced with dithiothreitol (Boehringer-Mannheim, Mannheim, Germany) alkylated with iodoacetamide (Sigma, St. Louis, MO) and digested with either trypsin (TPCK-treated, Sigma) or with endoproteinase Glu-C (Boehringer-Mannheim, Mannheim, Germany) as instructed by the product information sheets. In brief, the purified proteins were dissolved in 30–50 μl of 0.1 *M* ammonium bicarbonate (pH 8), reduced with 6 μl of 0.1 *M* dithiothreitol for 15 min at 56°C, and alkylated with 4 μl

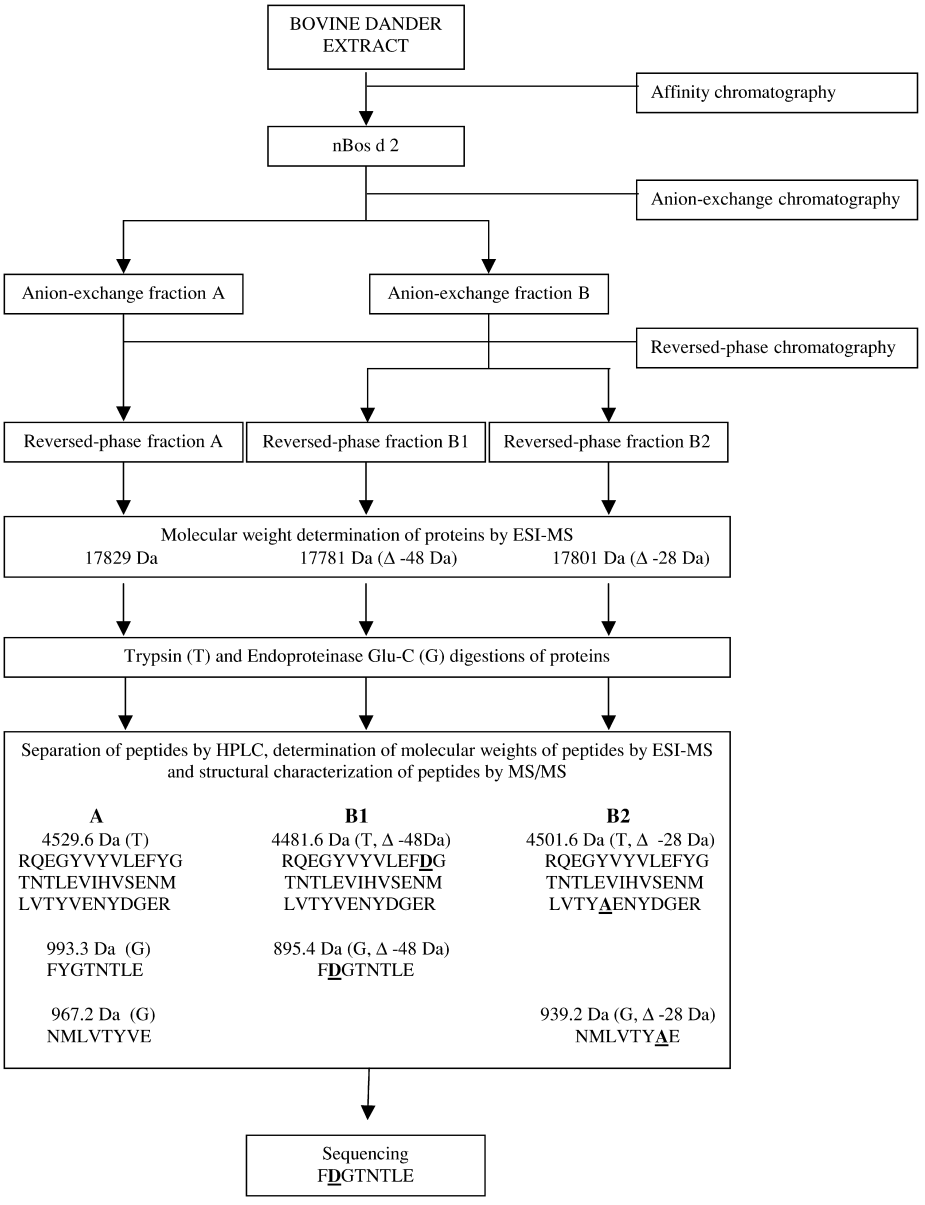

Fig. 1. Process scheme for the purification and characterization of Bos d 2 variants. The peptides obtained by trypsin digestion are marked by (T), and those obtained by Glu-C digestion are marked by (G).

of 0.5 *M* iodoacetamide for 15 at 23°C. The proteins were digested either with 2.6 μg trypsin, or with 2.6 μg Glu-C for 16 h at 37°C. Peptides in the protein digests were identified using HPLC–tandem mass spectrometry (HPLC–MS–MS) and with a protein sequencer (Applied Biosystems 477A).

3. Results

We have reported earlier [21] that the anion-exchange chromatography of affinity-purified natural Bos d 2 yielded two protein fractions (A and B) both reacting with IgE antibodies of cattle-allergic pa-

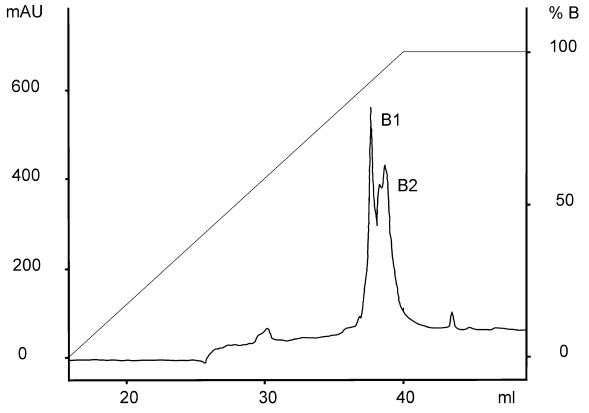

Fig. 2. Reversed-phase HPLC separation of anion exchange fraction B.

tients. Fraction B could be further separated with reversed-phase HPLC into two fractions B1 and B2 (Fig. 2). The molecular mass determination of intact proteins showed that fraction A was Bos d 2 with two cystein bridges (average molecular mass 17 829±1 Da). The molecular mass of the protein in fraction B1 was 48 Da lower than Bos d 2, and the protein B2 was 28 Da smaller than Bos d 2.

In order to find the relationship between Bos d 2 and the proteins in B1 and B2, we applied HPLC–MS–MS analysis after digestion of the proteins with either trypsin or Glu-C. The digestion with the conditions used resulted in both fully and partially cleaved peptides. The mass spectra of the tryptic peptides showed that the loss of 48 Da in protein B1 and loss of 28 Da in protein B2 was due to modifications in the tryptic peptide RQEGYVYVLEFYGTNTLEVIHVSENMLVTYVENYDGER (Figs. 1 and 3). The MS–MS spectrum of the unmodified peptide (M_w 4529.6) showed carboxy terminal y_7^+ ion at m/z 882.4 and y_8^+ ion at m/z 981.3 (difference between these ions is 99 Da, which is the residual mass of valine). The MS–MS spectrum of the modified peptide from protein B2 (M_w 4501.6) showed ion y_7^+ at m/z 882.4 (same as original) and y_8^+ ion at m/z 953.3 (28 Da less than the unmodified peptide, indicating a change of valine to alanine). The amino terminal b ions can be used for confirmation of this result, for the original peptide the b_{31}^{+3}, can be found at m/z 1217.0; and b_{28}^{+2} at m/z 1643.0, which corresponds to the sequence TYV. For the modified peptide in protein B2, b_{31}^{+3} was found at m/z 1207.4; and b_{28}^{+2} at m/z 1642.9, which corresponds to the amino acids TYA. The MS–MS spectrum of the tryptic peptide from B1 (M_w 4481.6) was less informative, but the site of variation was located between positions 73 and 97.

The HPLC–MS–MS analysis of peptides obtained with Glu-C digestion was also used to detect the positions of the modifications. In the case of the Glu-C-digested peptides, the carboxy terminal amino acid is glutamic acid. The presence of an acidic amino acid in the C-terminus does not favor the formation of y-ions which are typical of tryptic peptides containing basic arginine or lysine in the C-terminus. The MS–MS spectra (Fig. 4) of Glu-C digested peptide NMLVTYVE (M_w 967.2) from fraction A and of the modified peptide from fraction B2 (M_w 939.2) showed a series of structurally indicative b, b-NH_3, and b-H_2O ions. The b-ion

```
Peak A              1   QETPAEIDPS KIPGEWRIIY AAADNKDKIV EGGPLRNYYR RIECINDCES
Peak B1             1   QETPAEIDPS KIPGEWRIIY AAADNKDKIV EGGPLRNYYR RIECINDCES
Peak B2             1   QETPAEIDPS KIPGEWRIIY AAADNKDKIV EGGPLRNYYR RIECINDCES

Peak A             51  LSITFYLKDQ GTCLLLTEVA KRQEGYVYVL EFYGTNTLEV IHVSENMLVT
Peak B1            51  LSITFYLKDQ GTCLLLTEVA KRQEGYVYVL EFDGTNTLEV IHVSENMLVT
Peak B2            51  LSITFYLKDQ GTCLLLTEVA KRQEGYVYVL EFYGTNTLEV IHVSENMLVT

Peak A            101 YVENYDGERI TKMTEGLAKG TSFTPEELEK YQQLNSERGV PNENIENLIK
Peak B1           101 YVENYDGERI TKMTEGLAKG TSFTPEELEK YQQLNSERGV PNENIENLIK
Peak B2           101 YAENYDGERI TKMTEGLAKG TSFTPEELEK YQQLNSERGV PNENIENLIK

Peak A            151 TDNCPP
Peak B1           151 TDNCPP
Peak B2           151 TDNCPP
```

Fig. 3. Amino acid sequence alignment of Bos d 2 variants. Sequences were identified by HPLC–ESI-MS–MS. The amino acid sequence of peak A protein corresponds to the one previously reported for Bos d 2 [19]. The variant forms of Bos d 2 from the peaks B_1 and B_2 are Bos d 2.0102 and Bos d 2.0103 (Protein Identification Resource, PIR, accession numbers A59225 and B59225).

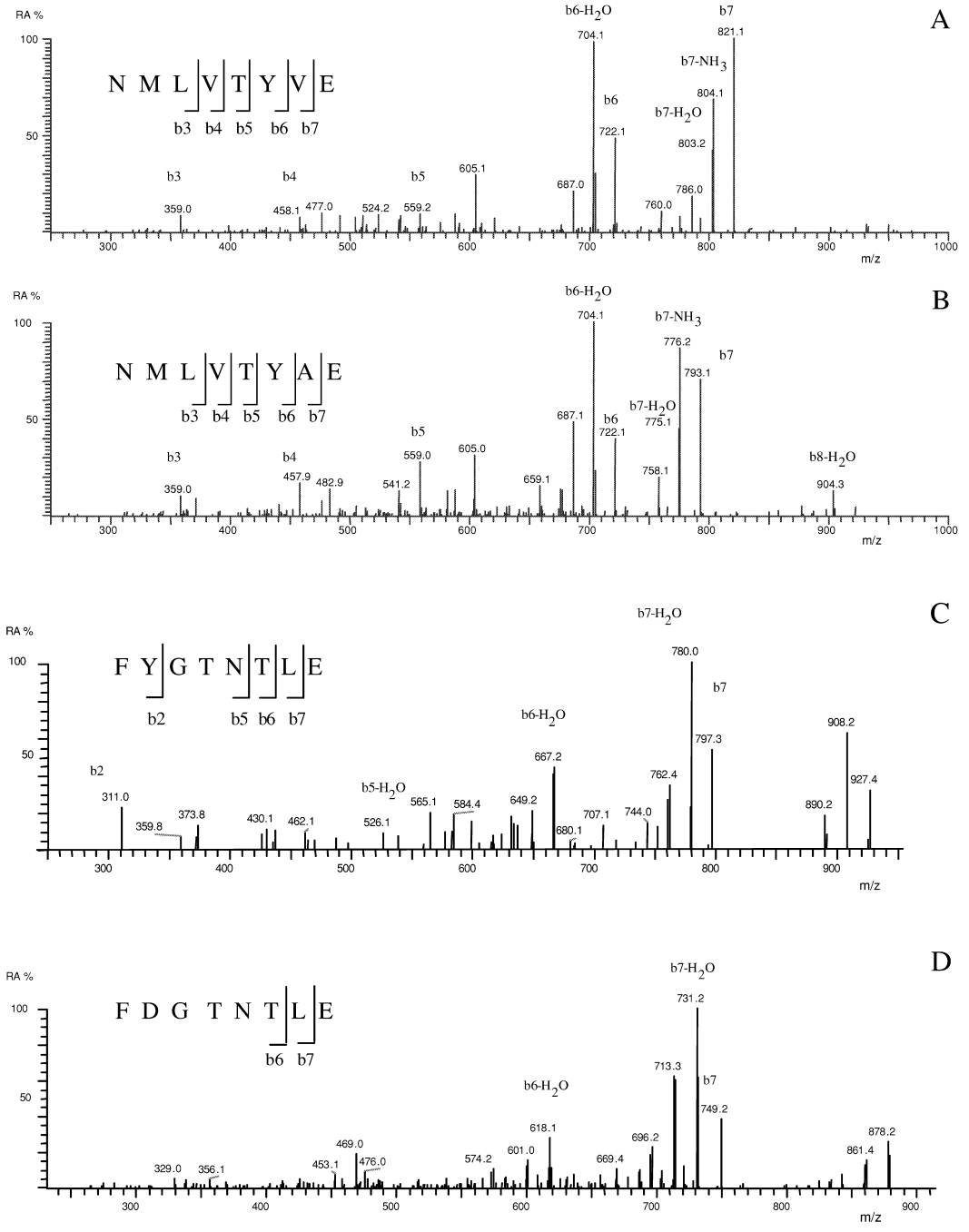

Fig. 4. Electrospray MS–MS spectra of two peptides obtained by Glu-C digestion. (A) Peptide NMLVTYVE originates from protein fraction (A), and (B) peptide NMLVTYAE is from protein fraction B2. (C) The peptide FYGTNTLE originated from protein fraction A and peptide FDGTNTLE (D) originated from B1. The protonated molecules MH^+ were selected for collisional-induced dissociation, and fragment ions were labeled according to Biemann [27].

series verify that the modified amino acid is valine at position 101 (V>A, loss of 28 Da).

The MS–MS spectrum of Glu-C digested peptide FYGTNTLE (M_w 993.3 Da) showed b_7 at m/z 797.3, b_7-NH_3 at m/z 780.0, b_7-H_2O at m/z 779 and b_6-H_2O at m/z 666. The MS–MS spectrum of modified peptide from fraction B1 (M_w 895.4), showed b_7 at m/z 749.2, b_7-H_2O at m/z 732, b_7-H_2O at m/z 731.2, and b_6-H_2O at m/z 618.1. The site of modification in protein B1 is thus located between positions 82 and 87 and the mass difference of 48 Da is attributable to a change from phenylalanine to valine or from tyrosine to aspartic acid. The structure of Glu-C peptide FDGTNTLE was finally verified with a protein sequencer. Both modifications (from tyrosine to aspartic acid in B1, and from valine to alanine in B2) can be explained by a modification of only one base pair in DNA.

4. Discussion

The aim of the present study was to identify two proteins copurifying with Bos d 2 through a multistep procedure including affinity chromatography with a monoclonal antibody against Bos d 2 [21]. The two approaches commonly used to characterize allergens at the sequence level are cDNA cloning and the standard sequencing of highly purified natural proteins. In the present study, we have applied an alternative method, HPLC–MS–MS analysis, to determine the amino acid sequences of the three proteins, designated as peaks A, B1 and B2 in the HPLC chromatogram and which were present in our preparation of natural Bos d 2 [21]. Bos d 2 has been defined by the sequence of the cDNA cloned from cow epithelium [19]. When analyzing the HPLC–MS–MS results, A, B1 and B2 proved identical with Bos d 2 except for amino acid substitutions in B1 and B2, one substitution in each protein. In protein B1, tyrosine (Y83) was substituted by aspartic acid, and in protein B2, valine (V102) was replaced by alanine. Therefore, the proteins in B1 and B2 represent two variants of Bos d 2 and they are proposed to be designated as Bos d 2.0102 and Bos d 2.0103.

Affinity chromatography with a single monoclonal antibody was used as the first step in the purification of the preparation containing the two new variants of Bos d 2. It is also possible that additional isoforms or variants of Bos d 2, not recognized by this monoclonal antibody, could be present in bovine tissues or body fluids. For example, the allergen found in bovine urine could well be a lipocalin allergen related to Bos d 2 [22].

The practical importance of the different sequence variations in allergens depends on the effects they have on the immunological properties of the molecule. Although the variants of Bos d 2 copurified in the monoclonal antibody-based affinity chromatography, differences in IgE epitopes cannot be excluded. A small alteration in an epitope may have a significant effect on IgE binding of an allergen [23,24]. When studying the localization of the amino acid substitutions of the Bos d 2 variants in the three-dimensional model [10], it is noteworthy that the Y83D substitution in protein B1 is located in the loop between two β-strands. Since this loop extends to the surface of the molecule (Fig. 5), it is possible that the substitution interferes with the IgE binding

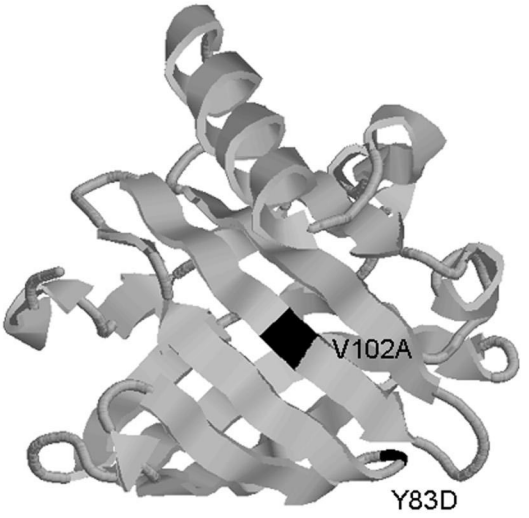

Fig. 5. Localization of the Y83D and V102A substitutions in the three-dimensional model of Bos d 2 (Bos d 2 coordinates: Brookhaven PDB code #1BJ7, figure by the RasMol 2.6 program).

ability. The V102A substitution in protein B2 is located in a β-strand which is one of the strands forming the barrel typical of all lipocalins. This substitution is especially interesting because the corresponding β-strand in another lipocalin allergen, Bos d 5 (β-lactoglobulin), has been shown to contain an IgE epitope [14,15].

The effects of the substitutions on the T-cell epitopes are equally important. We have earlier determined the T-cell epitopes of Bos d 2 [12]. The amino acid substitution Y83D will not change any of its known T-epitopes, whereas the substitution V102A in protein B2 is located in epitope E [12]. However, the role of B2 variant in T-cell responsiveness has not yet been studied. A T-cell response against Bos d 5 was observed in seven out of 12 newborns who had atopic parents [13]. An epitope mapping with CNBr peptides revealed that epitopes were scattered along the molecule, but that most of the reactivity (11 of the total of 28 positive responses) was directed against the C-terminal fragment p145–161. The corresponding region in Bos d 2 is adjacent to epitope G [12].

In general, the clinical significance of the variants of lipocalin allergens is largely unknown. The numerous sequence variants of mouse and rat urinary proteins or Mus m 1 and Rat n 1, respectively, have not been compared for their IgE-binding capacity or T cell reactivity. The same is true for variants of Bos d 5 and Equ c 2. More data are available for pollen allergens. Ferreira et al. [25] have demonstrated the wide variation between Bet v 1 isoforms in terms of IgE binding capacity. On the other hand, isoforms of Bet v 1 which had lost most of their capacity for IgE-binding still were capable of eliciting an intense T-cell response.

Recombinant DNA technology has opened new possibilities for designing and producing recombinant allergen vaccines [26]. On the other hand, most, if not all, allergens seem to comprise numerous isoforms, and this represents a challenge for the vaccine development [4]. Dealing with the isoforms is only one of the factors to be taken into account. Considering all the options available for designing recombinant vaccines, e.g., the combination with the immunostimulatory DNA sequences, the main problem in the future will not be the production of vaccine candidates but the testing of their clinical efficacy.

Acknowledgements

This study was supported by Kuopio University Hospital project no. 5035.

References

[1] T.P. King, D. Hoffman, H. Løwenstein, D.G. Marsh, T.A. Platts Mills, W. Thomas, Int. Arch. Allergy Immunol. 105 (1994) 224.
[2] T. Rafnar, I.J. Griffith, M.C. Kuo, J.F. Bond, B.L. Rogers, D.G. Klapper, J. Biol. Chem. 266 (1991) 1229.
[3] I. Swoboda, A. Jilek, F. Ferreira, E. Engel, K. Hoffmann-Sommergruber, O. Scheiner, D. Kraft, H. Breiteneder, E. Pittenauer, E. Schmid et al., J. Biol. Chem. 270 (1995) 2607.
[4] H. Løwenstein, S.H. Sparholt, S.S. Klysner, H. Ipsen, J.N. Larsen, Int. Arch. Allergy Immunol. 107 (1995) 285.
[5] K.Y. Chua, P.K. Kehal, W.R. Thomas, Int. Arch. Allergy Immunol. 101 (1993) 364.
[6] W.A. Smith, W.R. Thomas, Clin. Exp. Allergy 26 (1996) 571.
[7] C.H. Wu, M.F. Lee, N.M. Wang, S.F. Luo, Mol. Immunol. 34 (1997) 1.
[8] M.B. Lascombe, C. Gregoire, P. Poncet, G.A. Tavares, I. Rosinski-Chupin, J. Rabillon, H. Goubran-Botros, J.C. Mazie, B. David, P.M. Alzari, J. Biol. Chem. 275 (2000) 21572.
[9] Z. Böcskei, C.R. Groom, D.R. Flower, C.E. Wright, S.E. Phillips, A. Cavaggioni, J.B. Findlay, A.C. North, Nature 360 (1992) 186.
[10] J. Rouvinen, J. Rautiainen, T. Virtanen, T. Zeiler, J. Kauppinen, A. Taivainen, R. Mäntyjärvi, J. Biol. Chem. 274 (1999) 2337.
[11] S. Brownlow, J.H. Morais Cabral, R. Cooper, D.R. Flower, S.J. Yewdall, I. Polikarpov, A.C. North, L. Sawyer, Structure 5 (1997) 481.
[12] T. Zeiler, R. Mäntyjärvi, J. Rautiainen, M. Rytkönen-Nissinen, P. Vilja, A. Taivainen, J. Kauppinen, T. Virtanen, J. Immunol. 162 (1999) 1415.
[13] M. Piastra, A. Stabile, G. Fioravanti, M. Castagnola, G. Pani, F. Ria, Int. Arch. Allergy Immunol. 104 (1994) 358.
[14] G. Ball, M.J. Shelton, B.J. Walsh, D.J. Hill, C.S. Hosking, M.E. Howden, Clin. Exp. Allergy 24 (1994) 758.
[15] I. Selo, G. Clement, H. Bernard, J. Chatel, C. Creminon, G. Peltre, J. Wal, Clin. Exp. Allergy 29 (1999) 1055.
[16] J. Kauppinen, T. Zeiler, J. Rautiainen, M. Rytkönen-Nissinen, A. Taivainen, R. Mäntyjärvi, T. Virtanen, Clin. Exp. Allergy 29 (1999) 989.

[17] G. Stewart, P. Thompson, Clin. Exp. Allergy 26 (1996) 1020.
[18] V. Bulone, T. Krogstad Johnsen, B. Smestad Paulsen, Eur. J. Biochem. 253 (1998) 202.
[19] R. Mäntyjärvi, S. Parkkinen, M. Rytkönen, J. Pentikäinen, J. Pelkonen, J. Rautiainen, T. Zeiler, T. Virtanen, J. Allergy Clin. Immunol. 97 (1996) 1297.
[20] J. Rautiainen, M. Rytkönen, T. Virtanen, J. Pentikainen, T. Zeiler, R. Mäntyjärvi, J. Allergy Clin. Immunol. 100 (1997) 251.
[21] J. Rautiainen, S. Auriola, J. Rouvinen, J. Kauppinen, T. Zeiler, D. Novikov, T. Virtanen, R.A. Mäntyjärvi, Biochem. Biophys. Res. Commun. 247 (1998) 746.
[22] J. Ylönen, R. Mäntyjärvi, A. Taivainen, T. Virtanen, Clin. Exp. Allergy 22 (1992) 83.
[23] H. Breiteneder, F. Ferreira, K. Hoffmann-Sommergruber, C. Ebner, M. Breitenbach, H. Rumpold, D. Kraft, O. Scheiner, Eur. J. Biochem. 212 (1993) 355.
[24] G.A. Hakkaart, R.C. Aalberse, R. van Ree, Adv. Exp. Med. Biol. 409 (1996) 395.
[25] F. Ferreira, K. Hirtenlehner, A. Jilek, J. Godnik-Cvar, H. Breiteneder, R. Grimm, K. Hoffmann-Sommergruber, O. Scheiner, D. Kraft, M. Breitenbach, H.J. Rheinberger, C. Ebner, J. Exp. Med. 183 (1996) 599.
[26] R. Valenta, S. Vrtala, M. Focke-Tejkl, S. Bugajska, T. Ball, A. Twardosz, S. Spitzauer, H. Grönlund, D. Kraft, Biol. Chem. 380 (1999) 815.
[27] K. Biemann, Biomed. Environ. Mass Spectrom. 16 (1988) 99.