

Structural aspects of dog allergies: The crystal structure of a dog dander allergen Can f 4



Merja H. Niemi^a, Marja Rytkönen-Nissinen^{b,c}, Janne Jänis^a, Tuomas Virtanen^b, Juha Rouvinen^{a,*}

^a Department of Chemistry and Biocenter Kuopio, University of Eastern Finland, PO Box 111, 80101 Joensuu, Finland

^b Department of Clinical Microbiology, Institute of Clinical Medicine and Biocenter Kuopio, University of Eastern Finland, PO Box 1627, 70211 Kuopio, Finland

^c Institute of Dentistry, School of Medicine, University of Eastern Finland, PO Box 1627, 70211 Kuopio, Finland

ARTICLE INFO

Article history:

Received 5 March 2014

Received in revised form 10 April 2014

Accepted 14 April 2014

Keywords:

Allergen

Dog dander

Lipocalin

Crystal structure

Native mass spectrometry

Dimerization

ABSTRACT

Four out of six officially recognized dog allergens are members of the lipocalin protein family. So far, a three-dimensional structure has been determined for only one dog allergen, Can f 2, which is a lipocalin protein. We present here the crystal structure of a second lipocalin allergen from dog, a variant of Can f 4. Moreover, we have compared and analyzed the structures of these two weakly homologous (amino acid identity 21%) dog allergens. The size and the amino acid composition of the ligand-binding pocket indicate that Can f 4 is capable of binding only relatively small hydrophobic molecules which are different from those that Can f 2 is able to bind. The crystal structure of Can f 4 contained both monomeric and dimeric forms of the allergen, suggesting that Can f 4 is able to form transient (weak) dimers. The existence of transient dimers in solution was confirmed by use of native mass spectrometry. The dimeric structure of Can f 4 is formed when the ends of four β -strands are packed against the same strands from the second monomer. The residues in the interface are mainly hydrophobic and the formation of the dimer is similar to the major horse allergen Equ c 1. Interestingly, the crystal structure of dog Can f 2 has been reported to show a different type of dimer formation. The capability of these allergens to form dimers may be important for the development of immediate allergic reaction (mast cell activation) because oligomeric allergens can effectively present multivalent epitopes.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Among pets, dogs and cats are the major sources of human respiratory allergens. It has been estimated that 5–10% of the population worldwide is allergic to the domestic dog (*Canis familiaris*). Dog dander, urine and saliva contain several allergens that can induce IgE-mediated allergic reactions, such as allergic rhinitis or asthma. As allergens from dog dander and secretions can easily disperse in the environment, they are not only found in dog-owning households, but also in homes without pets as well as in public places where dogs are not usually present (Arbes et al., 2004; Morris, 2010). The development of a dog allergy is emotionally difficult for dog owners, since dogs are kept as family members. Thus, many people are not willing to give up their dogs, despite the symptoms

they cause. Although some dog breeds have been suggested to be hypoallergenic and therefore suitable for allergic individuals, investigations have proven that they are as allergenic as other breeds (Nicholas et al., 2011; Vredegoor et al., 2012).

So far, six respiratory allergens have been reported for domestic dogs, namely Can f 1 to Can f 6. Can f 3 is a serum albumin (69 kDa) (Spitzauer et al., 1994). Approximately 35% of dog-allergic individuals have IgE against it and thus it is classified as a minor allergen of dog. A major allergen, mainly present in dog urine, was identified as prostatic (urine) kallikrein or arginine esterase (28 kDa). It was named Can f 5 (Mattsson et al., 2009). Up to 70% of patients with dog allergy were found to have specific IgE against Can f 5 and almost one-third was sensitized solely to Can f 5. Therefore, Can f 5 was recognized as one of the major dog allergens.

Interestingly, no less than four out of the six reported dog allergens (Can f 1, Can f 2, Can f 4, and Can f 6) are members of the lipocalin protein family. Can f 1 and Can f 2 were the first two dog

* Corresponding author. Tel.: +358 40 705 5105; fax: +358 29 445 7319.

E-mail address: juha.rouvinen@uef.fi (J. Rouvinen).

allergens purified from dog hair and dander in 1991 (De Groot et al., 1991; Schou et al., 1991). They were later produced as recombinant proteins (Konieczny et al., 1997). Can f 1 is expressed in tongue epithelial tissue, parotid and mandibular glands, and skin. It is a major allergen, as it causes IgE antibody response in 50–70% of dog-allergic individuals. Another salivary lipocalin allergen of dog, Can f 2, sensitizes approximately one-third of dog-allergic patients and is thus a minor allergen. The IgE response of dog-sensitized individuals is often directed towards Can f 1 and Can f 2 (Konieczny et al., 1997; Saarelainen et al., 2004). A recent study indicated that a minority of patients (5%) may have Can f 2-specific IgE in the absence of Can f 1-specific IgE. Can f 2 has been suggested to be IgE cross-reactive with the cat lipocalin allergen Fel d 4. Furthermore, it was the first dog allergen for which the three-dimensional structure was solved by X-ray crystallography (Madhurantakam et al., 2010). Can f 6, which is expressed in the submaxillary gland, skin and bladder tissues of dog, was shown to sensitize 61% of dog-allergic individuals. Although Can f 6 displays low sequence identity with other dog lipocalin allergens (19–23%), its sequence similarity with the major lipocalin allergens from cat (Fel d 4, 67%) and horse (Equ c 1, 57%) is high. ELISA inhibition analyses using the sera from patients sensitized to both cat and dog indicated that Can f 6 and Fel d 4 are IgE cross-reactive (Hilger et al., 2012).

Soon after the allergenic lipocalins Can f 1 and Can f 2 were characterized, it was noticed that they were not the only dog proteins to induce IgE response. Konieczny et al. observed that three out of the 20 plasma samples examined contained IgE that bound to a dog skin test extract but did not recognize either Can f 1 or Can f 2 (Konieczny et al., 1997). Furthermore, Saarelainen et al. observed that the highest IgE reactivity in immunoblotting with a dog epithelial extract was directed against an unidentified, 18 kDa lipocalin protein presented in a dog epithelial extract (Saarelainen et al., 2004). This allergenic component was later named Can f 4. Can f 4 is a 158 amino-acid lipocalin protein expressed by tongue epithelial tissue (Mattsson et al., 2010). The allergen is principally found in saliva and dander. Can f 4 was initially purified from dog dander by Mattsson et al. and then produced as a recombinant protein in *Escherichia coli* for immunological characterization (Mattsson et al., 2010). An ImmunoCAP analysis, carried out using the sera from 37 dog-allergic patients, showed that 35% of the individuals displayed IgE against Can f 4. Moreover, one of the dog allergic patients had a strong IgE reaction to rCan f 4 without a response to other dog allergens. The results suggested that Can f 4 is a potentially significant allergen and may also act as an independent sensitizer that leads to dog allergy. Can f 4 was detected to be IgE cross-reactive with a 23 kDa odorant-binding protein isolated from cow dander, thus offering a possible linkage between dog and bovine dander allergies (Mattsson et al., 2010).

Here, we present the three-dimensional structure of the recombinant dog lipocalin allergen Can f 4, determined at a 2.6 Å diffraction resolution. The Can f 4 used in this study represents a variant in which four amino acids are different from what was previously reported (Ile20Val, Asp30Glu, Met38Leu, and Ser51Leu; Mattsson et al. (2010)). Apart from the X-ray diffraction analysis, rCan f 4 was also characterized with high-resolution mass spectrometry to verify the correctness of the expressed protein structure and to identify possible modifications. Furthermore, native mass spectrometry was used to detect the oligomeric state of the allergen in native-like solution conditions. Before this study, only one allergen structure from dog, namely the structure of the minor dog allergen Can f 2, had been determined (Madhurantakam et al., 2010). Thus, this is the first publication to describe the structure of another dog allergen, Can f 4. As the allergen is distantly homologous to Can f 2, the setting gives a possibility for structural comparisons and insights into the origin of our allergic reactivity to dogs.

2. Materials and methods

2.1. Isolation of the Can f 4 variant cDNA and the production of recombinant Can f 4

To generate the DNA fragment that encodes the Can f 4 sequence for cloning, Can f 4-specific primers were designed based on information on the canine genomic DNA sequences and the nucleotide sequence of Can f 4.

Total RNA was isolated from the mixture of tongue, mandibular and parotid gland tissue samples of a single dog (Beagle, male) by the RNeasy[®] total RNA isolation system (Promega, Madison, WI, USA) following the manufacturer's instructions. The total RNA obtained was utilized for cDNA synthesis using oligo(dT)18 primers and AccuScript reverse transcriptase (Stratagene, La Jolla, CA, USA) according to the supplier's instructions. The oligo(dT)18 primed cDNAs were used with Can f 4 gene-specific primers (the 5' primer includes the leader sequence of Can f 4) to amplify the fragments of the allergen by using a Phusion DNA polymerase (Finnzymes Oy, Espoo, Finland). To subclone and sequence the amplicons, products were gel-extracted by using an Ultrafree-DA Centrifugal Filter Device (Millipore, Bedford, MA, USA). Then, a Biotools DNA polymerase (Biotools, Madrid, Spain) was used to add a 3' A overhang to purified PCR fragments for UA cloning by a Qiagen PCR cloning plus kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. The nucleotide sequence of the generated construction was confirmed by DNA sequencing by using a Thermo Sequenase CY5 Dye Terminator Kit and the A.L.F. Express DNA Sequencer (GE Healthcare Bioscience AB, Uppsala, Sweden). The GenBank accession number for the sequence of the full-length cDNA of the Can f 4 variant reported in this paper is KF192077.

The achieved construct was cloned into the pPIC9 expression vector. *Pichia pastoris* yeast was transformed with the construct pPIC9rCan f 4 and the allergen expression was performed according to the manufacturer's instructions (Invitrogen, San Diego, CA, USA). The culture supernatant was clarified by centrifugation and filtration. The HiTrap NHS column (GE Healthcare) was coupled with a Can f 4 allergen-specific monoclonal antibody (48F) (see below) and the recombinant allergen was purified from the extract with the use of affinity chromatography, according to the instructions of the manufacturer (GE Healthcare). Then, the rCan f 4 preparation was subjected to gel filtration on the Superdex 75 preparative grade column XK16/70 (GE Healthcare Bioscience AB) with the GE Healthcare Äktapurifier HPLC system with PBS as the eluant. The calibration curve for the column was obtained by using a molecular-weight standard kit, MWGF70 (Sigma-Aldrich, St. Louis, USA). The rCan f 4-containing gel filtration fractions were pooled and then the buffer was exchanged for the anion-exchange chromatography with a HiTrap desalting column (GE Healthcare). Anion-exchange chromatography was performed with a Resource Q column (GE Healthcare Bioscience AB) and the bound rCan f 4 was eluted with a linear 20 CV gradient from 0 to 0.5 M NaCl in 20 mM Tris-HCl buffer at pH 7.5. The purity and IgE reactivity of the preparation was verified by sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting (see below). Protein concentrations were determined by absorbance spectroscopy which used the molar absorption coefficients (<http://us.expasy.org/tools/protparam.html>). Sterile-filtered preparations were stored at –70 °C.

2.2. Generation of the monoclonal antibody 48F specific to the Can f 4 allergen

The concentrated skin prick test extract of dog epithelium (ALK Abelló, Hørsholm, Denmark) was subjected to gel filtration, as described above. BALB/c mice were immunized intraperitoneally

with the fractions that contained the Can f 4 allergen combined with the Freund's complete adjuvant, basically as described previously (Ylönen et al., 1994). Three subsequent immunizations were done at 4-week intervals. The mice were sacrificed four days after the last immunization and the spleen cells were fused with the Ag-8 mouse myeloma cell line by using a hybridoma fusion and cloning supplement HFCS (Roche, Penzberg, Germany), according to the manufacturer's instructions. The hybridomas secreting specific antibody to Can f 4 was cloned two to three times by a limiting dilution method and were then further propagated. The monoclonal antibody 48F was purified by affinity chromatography on a protein-G column (GE Healthcare, USA). The isotype of the antibody was tested by a Clonacell InstantCHEK One-minute isotype Kit (SteamCell, Vancouver, Canada). Protein concentrations were determined by the Bio-Rad Protein Assay Kit by using bovine gamma globulin as a standard. Experiments were performed in agreement with the European Convention for the Protection of Vertebrate Animals used for Experimental and Other Scientific Purposes (Strasbourg 18 March 1986, adopted in Finland 31 May 1990). The study was approved by the Animal Care and Use Committee of the University of Kuopio (now University of Eastern Finland) with the permit number 04-37/2004.

2.3. Immunoblotting

The IgE reactivity of rCan f 4 was verified with a representative serum from a clinically diagnosed dog-allergic patient (dog dander UniCAP IgE measurement (Pharmacia, Uppsala, Sweden) was 0.79 kU/l) by using immunoblotting, as described previously (Saarelainen et al., 2004). Using the same technique, the IgE reactivity of the serum against the commercial dog epithelial extract from ALK Abelló (Hørsholm, Denmark) was also verified. In brief, after SDS-PAGE and protein transfer, the membrane strips with the allergen or the epithelial extract were incubated overnight with the patient serum (1:10), followed by incubation for 2 h with monoclonal horseradish peroxidase-labelled mouse anti-human IgE (Southern Biotechnology Associates, Inc., Birmingham, AL, USA) at 1:2000. The reactive bands were visualized with the ECL detection reagents (Amersham International, Buckinghamshire, England).

2.4. Crystallization and X-ray data collection

The recombinant Can f 4 allergen was crystallized by the hanging-drop vapour-diffusion method at 293 K. Crystallization drops were prepared by mixing 1 μ l of rCan f 4 protein solution (1.6 mg ml⁻¹ in 50 mM ammonium acetate buffer, pH 6.7) and 1 μ l of precipitant solution that contained 0.2 M potassium phosphate monobasic, 10% (w/w) PEG8000 at pH 4.6. The obtained crystals were often clusters of several individual crystals tightly joined together. Techniques used to improve the crystal quality, such as seeding and the careful optimization of initial crystallization conditions, were not successful in yielding crystals for data collection. However, along with the crystal clusters, single crystals occasionally appeared in crystallization drops.

For the data collection, crystals were cryoprotected by soaking them quickly into a cryo-protective solution (0.2 M potassium phosphate, 20% PEG8000, 15% glycerol) and then freezing them in liquid nitrogen. The X-ray diffraction data were collected from a single crystal that diffracted to a resolution of 2.6 Å on beamline ID29 at the ESRF (Grenoble, France). In order to collect a complete data set, the crystal was rotated 90° with a 0.1° oscillation range and the reflections were recorded with a Pilatus 6M-F pixel detector (Dectris Ltd). The diffraction data set was processed with a XDS program and scaled and merged with XSCALE (Kabsch, 1993).

Table 1

Data collection and refinement statistics. Values in parentheses are for the highest resolution shell.

<i>Data collection</i>	
Space group	P2 ₁ 2 ₁ 2 ₁
Resolution (Å)	44.6–2.6 (2.7–2.6)
Unit cell parameters	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	39.6, 74.4, 167.4
No. of observed reflections	67,789 (6943)
No. of unique reflections	15,857 (1665)
Redundancy	4.3 (4.2)
Completeness (%)	99.4 (99.1)
Average <i>I</i> / σ	23.4 (4.4)
<i>R</i> _{obs} (%)	4.2 (33.1)
<i>R</i> _{meas} (%)	4.8 (38.0)
<i>Refinement</i>	
<i>R</i> _{work} / <i>R</i> _{free} (%)	23.7/28.1
No. of non-hydrogen atoms	3527
Average B-factors (Å ²)	59.7
Main-chain atoms	57.0
Side-chain atoms	62.4
Rms deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.245
PDB code	4ODD

*R*_{free} test set comprises 5% of the randomly selected reflections omitted from a model refinement.

2.5. Structure determination and refinement

The rCan f 4 was crystallized in a space group P2₁2₁2₁. The crystal structure was solved by the molecular replacement method, carried out by a PHASER program (McCoy et al., 2007) from a CCP4 package (Winn et al., 2011). The odorant-binding protein from nasal mucosa of porcine (PDB code 1A3Y (Spinelli et al., 1998)), having a 37% sequence identity with Can f 4, was selected as a search model. The molecular replacement gave a solution for three rCan f 4 molecules in an asymmetric unit. The structure was refined with a Phenix.refine program (Adams et al., 2010) and the protein model was constructed manually with a Coot program (Emsley et al., 2010). Non-crystallographic symmetry (NCS) restraints were used in refinement in order to avoid the over-fitting of the protein model. Briefly, at the beginning of the refinement all the monomers of the asymmetric unit were assumed to be similar. Later, the NCS restraints were removed from the loop regions and the regions where the electron density showed differences between the protein molecules. The protein model was refined to *R*_{work} and *R*_{free} values of 23.7 and 28.1%, respectively. Solvent molecules were not added to the protein model due to the medium resolution of the data set. The protein model was validated with a Rampage program (Lovell et al., 2003). The Ramachandran plot for the rCan f 4 model indicated that 99.5% of the residues are in the favoured or allowed regions and 0.5% in outlier regions. The data collection and refinement statistics are given in Table 1. Dimeric interface areas were calculated with the Arealmol (Winn et al., 2011). The amino acid residues of the ligand-binding sites and the volumes of the central cavities were determined with a CASTp program (Dundas et al., 2006). The atomic coordinates and the structure factors have been deposited in the RCSB Protein Data Bank with the accession code 4ODD.

2.6. Mass spectrometry

Mass spectrometric analyses were performed with a 12 T APEX-Qe Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer (Bruker Daltonics, Billerica, MA, USA), equipped with an ESI source (Apollo-II). The accurate molecular mass of rCan f 4 was determined in the denaturing solution conditions. Briefly,

the protein sample in 50 mM ammonium bicarbonate buffer was diluted to 2 μ M with MeCN:H₂O:HOAc (49.5:49.5:1.0, v/v) and directly infused at a flow rate of 1.5 μ L/min. Dry nitrogen was used as the drying (250 °C, 6 mbar) and nebulizing gas. ESI-generated ions were accumulated in the hexapole ion trap for 1.0 s and transferred to the ICR cell for trapping, excitation and detection. For a mass spectrum, a total of 256 co-added (1 Mword) time-domain transients were recorded and subjected to the fast Fourier transform and magnitude calculation. To study the oligomeric state of rCan f 4 in native-like solution conditions, the protein sample was desalted with a PD-10 column (Amersham-Biosciences, Billingham, USA) using 10 mM ammonium acetate buffer (pH 6.9) as an eluent. In order to maintain weak non-covalent interactions in the gas-phase, the instrumental parameters were carefully optimized to avoid any unintentional oligomer dissociation upon ionization/desolvation. Moreover, prior to the determination of native mass spectra of rCan f 4, bovine beta-lactoglobulin (Bos d 5, ordered from Sigma–Aldrich) was measured as a reference protein to observe the dimer formation, as previously reported (Rouvinen et al., 2010). The data sets (500 co-added, 256 kword time-domain transients) were fast Fourier-transformed before the magnitude calculation. Mass calibration was done externally with respect to the ions of an ES Tuning Mix (Agilent Technologies). All mass data were processed with a Bruker XMASS software (version 7.0.8). The equilibrium dissociation constant for the protein dimer (K_d) was determined from the equation $K_d = [M]_{\text{total}} \times 1/(R(2R+1))$, where $[M]_{\text{total}}$ is the total protein (monomer) concentration and R is a dimer-monomer concentration ratio, as described previously (Liu and Konermann, 2011). A mass spectrum directly provides R through peak intensities with the assumption that the response factors (ion transmission and ionization efficiencies) are the same for the monomer and the dimer. In fact, Liu and Konermann have recently shown that the response factors for the monomers and dimers of Bos d 5 are surprisingly similar (Liu and Konermann, 2011). Thus, given the similar size and fold between Bos d 5 and Can f 4, it is reasonable to assume very similar response factors for the monomer and dimer of Can f 4 as well. All intensities were summed over all observed charge states for the calculation of R in order to avoid any bias due to slightly different charge state distributions for the monomer and the dimer.

3. Results and discussion

3.1. Isolation, expression and immunological characterization of the Can f 4 variant

The variant of rCan f 4 used in this study was cloned from the total RNA obtained from a mixture of tongue, mandibular and parotid gland tissues of a single dog, a male Beagle, as described in the Methods section. Comparison of the deduced amino acid sequence of this variant with the previously published one (Mattsson et al., 2010) revealed four amino acid substitutions: Ile20Val, Asp30Glu, Met38Leu, and Ser51Leu.

Recombinant Can f 4 was expressed as a soluble protein in *Pichia pastoris*. It was purified to a high extent, as shown in SDS-PAGE (see Fig. 1, Lane 1). The yield of the allergen after all purification steps, was around 1.6 mg per litre of yeast culture. The IgE binding capacity of the allergen was verified by immunoblotting (Fig. 1).

3.2. Mass spectrometry

To analyze the rCan f 4 allergen preparation in detail, the accurate molecular mass of the allergen was measured by high-resolution ESI FT-ICR mass spectrometry in denaturing solution conditions (Fig. 2). The experimentally determined

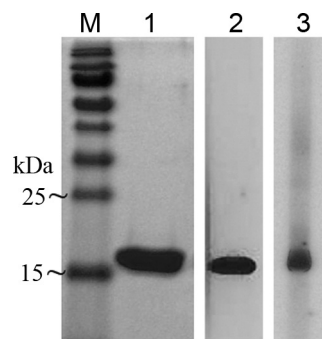


Fig. 1. SDS-PAGE pattern of the purified rCan f 4 stained with Coomassie Brilliant blue (Lane 1) plus a representative example of IgE reactivity of one clinically verified dog-allergic subject to rCan f 4 (Lane 2) and commercial dog epithelial extract from ALK Abelló (Lane 3). The membrane strips with the allergen preparations (Lanes 2 and 3) were incubated overnight with a patient serum (1:10), followed by incubation for 2 h with HRP-labelled mouse anti-human IgE at 1:2000. The reactive bands were visualized with the ECL detection reagents. Lane M, molecular weight markers (PAGERULER™ Prestained protein ladder, Fermentas, #SM0671).

molecular mass (averaged over the charge state distribution) was $17,627.11 \pm 0.01$ Da, which agrees precisely with the mass calculated from the full amino acid sequence of Can f 4, therefore two cysteines (Cys62 and Cys154) form a disulfide bridge (17,627.08 Da). This is in agreement with the crystal structure. Apart from the main form, the mass spectrum revealed the presence of two modifications in the protein. In a part of the protein molecules, the N-terminal glutamine had undergone a cyclization into a pyrrolidone carboxylic acid (PCA) form (17,610.05 Da, mass difference –17 Da). In addition, the protein was partly oxidized (most probably a single methionine residue) and displayed a +16 Da difference to the main form (17,643.07 Da). The determined crystal structure shows that one of the total of the three methionine residues of rCan f 4 (Met92) is located on the surface of the protein and is thus vulnerable to oxidation. The mass-spectrometric characterization of wild-type Can f 4 will be published elsewhere (Rytönen-Nissinen et al., submitted for publication).

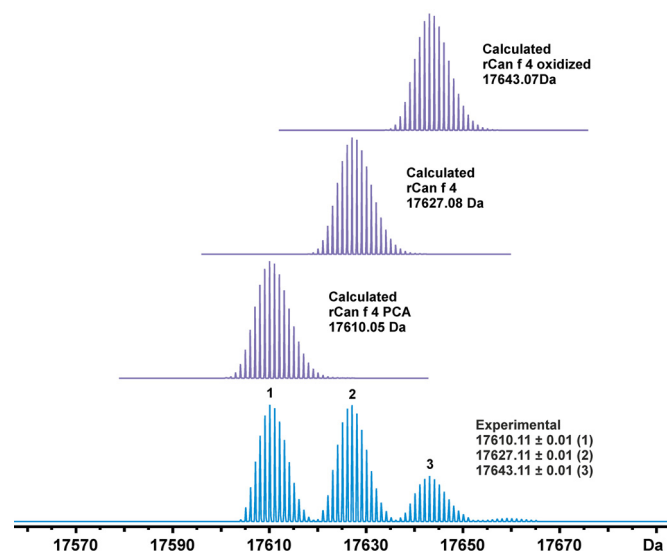
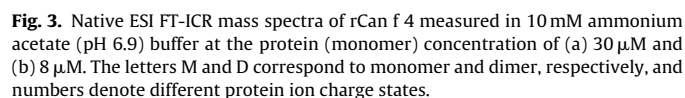
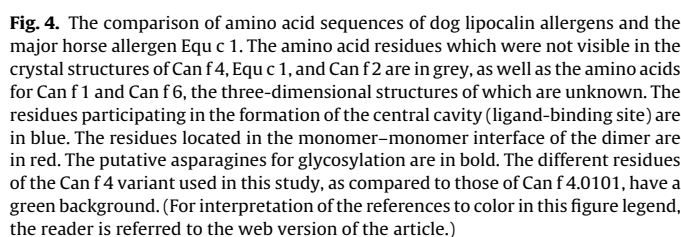


Fig. 2. Deconvoluted ESI FT-ICR mass spectrum of rCan f 4 measured in MeCN:H₂O:HOAc (49.5:49.5:1.0, v/v). The blue trace shows the experimental isotopic pattern and the magenta traces the theoretical isotopic patterns for the three protein forms observed with the most abundant isotopic masses indicated. For the undeconvoluted charge-state spectrum, see Supplemental Data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)



Native mass spectrometry (native MS) is a novel method for the analysis of proteins in their native state, without destroying weak non-covalent interactions in the gas-phase (Heck, 2008; Van den Heuvel and Heck, 2004). Compared to the more traditional methods for studying protein oligomerization, such as size-exclusion chromatography or analytical ultracentrifugation, native MS possesses several benefits, including speed of analysis, high sensitivity, selectivity, low sample volume requirements and direct stoichiometric information. Native mass spectrometry is a particularly suitable method for studying transient protein complexes in which different oligomeric states are presented simultaneously and their relative amounts vary, depending on the solution conditions (Rouvinen et al., 2010). In this study, native MS was used to detect the oligomeric state of rCan f 4 in solution and to calculate the relative dissociation constant (K_d) of the dimer observed in the crystal structure. A native mass spectrum of rCan f 4 at a protein (monomer) concentration of 30 μ M shows a mixture of monomers and dimers (Fig. 3a). At a lower protein concentration of 8 μ M, less dimers were observed as compared to the monomeric form of rCan f 4 (Fig. 3b). The relative K_d -value for the rCan f 4 dimer was determined by measuring the monomer-dimer ratios at two different protein concentrations, 8 and 30 μ M, and averaged over two repeated measurements. The determined K_d value of 27 ± 8 μ M (see, Materials and Methods for details) proves that Can f 4 is, indeed, a weak transient dimer, similar to many other allergens (Rouvinen et al., 2010).

The structure of the recombinant Can f 4 allergen was determined at a diffraction resolution of 2.6 Å. The protein was crystallized in the orthorhombic space group P2₁2₁2₁ with three rCan f 4 monomers (A, B, and C) in an asymmetric unit. Due to the medium resolution diffraction data, partial non-crystallographic symmetry (NCS) restraints were used for atomic coordinates and individual B factors to keep the three monomers similar during the refinement. The crystallographic *R*-factor was 23.7% and the free *R*-factor was 28.1%. The stereochemical quality of the model is good, and the clear electron density maps enabled a reliable interpretation of the structure. The first eight residues at the N-terminus (Gln-Leu-Pro-Leu-Pro-Asn-Val-Leu) and the last residue at the C-terminus (Glu) were not included in monomers A and C, whereas



the first four residues at the N-terminus and the last two residues (His-Glu) were not included in monomer B (Fig. 4). Because mass spectrum for the protein showed the presence of all residues, the lack of electron density is probably due to the flexible nature of both termini. In addition, the loop, which connects the H strand of the β -barrel to the C-terminal α -helix, has high B-factors, thus indicating remarkable flexibility. The average B-factor for rCan f 4 in the final model is relatively high ($59.7 \text{ \AA}^2$), but close to that calculated from the Wilson plot ($52.7 \text{ \AA}^2$). The temperature factor analysis shows that the average B-factor for the residues located inside the binding cavity is lower in comparison to the overall B-factor ($41.7 \text{ \AA}^2$).

The solved crystal structure shows that rCan f4 adopts a highly conserved fold that is typical for the members of the lipocalin protein family (Flower, 1996). The eight-stranded anti-parallel β -barrel forms the core of the protein. The β -strands of the barrel, labelled from A to H, enclose the central ligand-binding site. The structure also contains a four-turn C-terminal α -helix packed against the β -barrel, and a short 9th β -strand (I) at the C-terminus (Fig. 5a). A single disulfide bond formed between the residues of Cys62 and Cys154 stabilizes the protein by connecting the C-terminal part of the protein to the central β -barrel. The open end of the β -barrel, containing the entrance to the ligand-binding cavity, is comprised of three short β -hairpin loops (L3, L5, L7) and a longer L1 loop. The other side of the β -barrel, the closed end, contains one β -hairpin loop (L4) and two other short loops (L2 and L6) and the N-terminal extension of the polypeptide chain, which closes access into the β -barrel (Fig. 5a).

The Can f4 variant used in this study contains four mutations, as compared to the previously reported Can f 4.0101 (Mattsson et al., 2010). All mutations are in the N-terminal part of the polypeptide chain (Fig. 4). The ligand-binding site contains the substitutions I20V and M38L, which may alter the ligand-binding preference. The substitution D30E is located in the loop L1 and S51L in β -strand A on the protein surface. Because the IgE epitopes of Can f 4 are not known, it is not possible to estimate their influence on IgE binding. However, it could be noted that allergen variants with only a few

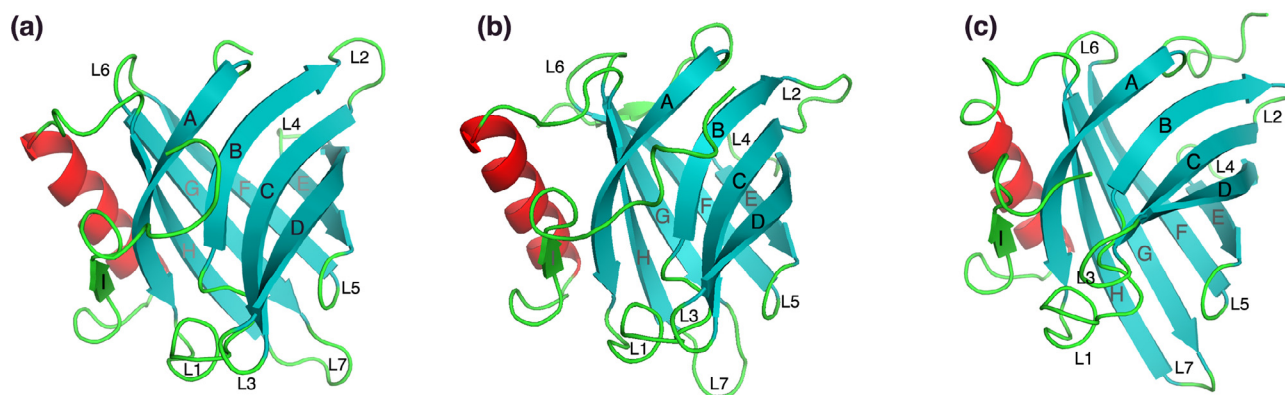


Fig. 5. Structures of three lipocalin allergens: (a) Can f 4 (4ODD), Equ c 1 (1EW3), and Can f 2 (3L4R). The eight antiparallel β -strands which form a central barrel are in cyan and labelled from A to H. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

substitutions may have considerably altered allergenicity, as is the case with the major birch pollen allergen Bet v 1 (Wagner et al., 2006).

An analysis based on the data extracted from the AllFam database pointed out that allergens could only be found in a limited number of protein families (Radauer et al., 2008). In practice, a great majority of significant respiratory allergens of mammalian origin belong to the lipocalin protein family (Radauer et al., 2008; Virtanen et al., 2012). The dog lipocalin allergen Can f 4 shows higher amino acid identity with the non-allergenic odorant-binding proteins from swine, bovine and rat (31–37%) than with the other mammalian lipocalin allergens (16–31%). The amino acid identity less than 50% often results in no immunological cross-reactivity (Radauer et al., 2008). Although several lipocalin allergens are potentially glycosylated (Fig. 4), the sequence of Can f 4 does not contain putative N-glycosylation sites.

3.4. Structural comparison with the Can f 2 allergen

The crystal structure of the recombinant Can f 2, determined at a 1.45 Å diffraction resolution, was the first lipocalin allergen of dog characterized at the three-dimensional structure level (Madhurantakam et al., 2010). Although lipocalins share a highly similar overall structure, they commonly show low sequential identity. Likewise, the sequence similarity is low between the allergenic dog dander lipocalins Can f 2 and Can f 4: only 21% of the amino acids are identical.

Superimposition of the rCan f 2 and rCan f 4 structures, performed with the programme PyMol (Delano, W. L. The PyMol Molecular Graphics System, www.pymol.org), indicated that the r.m.s. deviation for C α atoms of 148 structurally equivalent residues is 1.1 Å. Differences between the structures are mainly found in loop regions. Loops L2, L3, L6, and L7 show slight changes, whereas the longer L1 loop occupies a highly similar conformation in both structures (Fig. 5a and c). This is consistent with the other crystal structures of lipocalins, where the conformation of L1 loop is highly conserved. Differences were also detected in the loop located between the H-strand of the β -barrel and the α -helix. Because this loop is highly flexible in the structure of rCan f 4, no reliable comparison of loop conformations can be made. The superimposition with the Equ c 1 allergen (Fig. 5b) from horse showed a larger difference of 1.63 Å deviation, although it has a similar dimeric structure as Can f 4.

3.5. The ligand-binding cavity

The electron density observed inside the central cavity of rCan f 2 was interpreted as a glycerol, used to cryoprotect the protein

crystal before the data collection at 100 K (Madhurantakam et al., 2010). Similarly, in the case of rCan f 4, the electron density maps showed residual electron density inside the highly hydrophobic ligand-binding cavity. Due to the limited resolution of the diffraction data, the identity of the bound small molecule remained unclear. Although the protein maintains its functional conformation in the native mass spectrometry and also allows for the analysis of weakly binding ligands, native mass spectra measured for rCan f 4 did not show any evidence of small molecule binding (see above). However, the unknown ligand detected in the binding site in the crystal structure is probably not biologically significant since the protein was produced in a recombinant form. The bound molecule most probably originates from production and purification steps or from the crystallization solution.

Lipocalins are known for their capability to bind and transport small hydrophobic molecules, such as fatty acids or pheromones (Flower, 1996). The ligand-binding specificity and thus the biological role of Can f 4 is currently unknown. The solved crystal structure revealed the putative ligand-binding site of rCan f 4, located inside the central β -barrel typical for lipocalins. The surface area of the cavity is approximately of 390 Å² and the volume 470 Å³, calculated by CASTp (Dundas et al., 2006). A large number of hydrophobic side-chains participate in the formation of the cavity. These predominantly include amino acid residues from seven of eight β -strands (strands B–H), which form the central antiparallel β -barrel (Fig. 5). In total, 18 residues define the shape of the cavity (Fig. 6a). The flat face of the aromatic rings of residues Phe52, Phe54, Tyr81, and Phe87 covers a large part of the cavity surface. In addition, aliphatic side-chains of Ile36, Leu38, Val67, Val113, Leu115, and Met117 have an important role in shaping the cavity. The cavity contains only two polar residues, Asn85 (from the β -strand F) and Asp101 (from the β -strand G), the side-chains of which are close to each other. The N δ 2 atom of Asn85 is very close (2.6 Å) to the carbonyl main chain oxygen of Tyr81, which suggests a hydrogen bond between them. This leaves the O δ 1 of Asn85 very close (2.6 Å) to the O δ 1 of Asp101. In the latter case, there is a hydrogen bond only if the O δ 1 is protonated, which would mean that Asp101 is probably in a neutral form in the structure. Asn85 and Asp101 may form hydrogen bonds with a ligand bound to the cavity.

We have compared in detail the central cavity of Can f 4 with that of the Equ c 1 allergen of horse that has a similar dimeric structure to Can f 4, as well as with Can f 2, another lipocalin allergen from dog (Fig. 6). The amino acid sequence identity is low between Can f 4 and Equ c 1 as well as between Can f 4 and Can f 2, 28 and 22%, respectively. The volume of the cavity is smaller in Equ c 1 (340 Å³) and slightly larger in Can f 2 (510 Å³), compared with that of Can f 4 (470 Å³). The shapes of the cavities are clearly different in all three lipocalin allergens (Fig. 6), which would suggest that

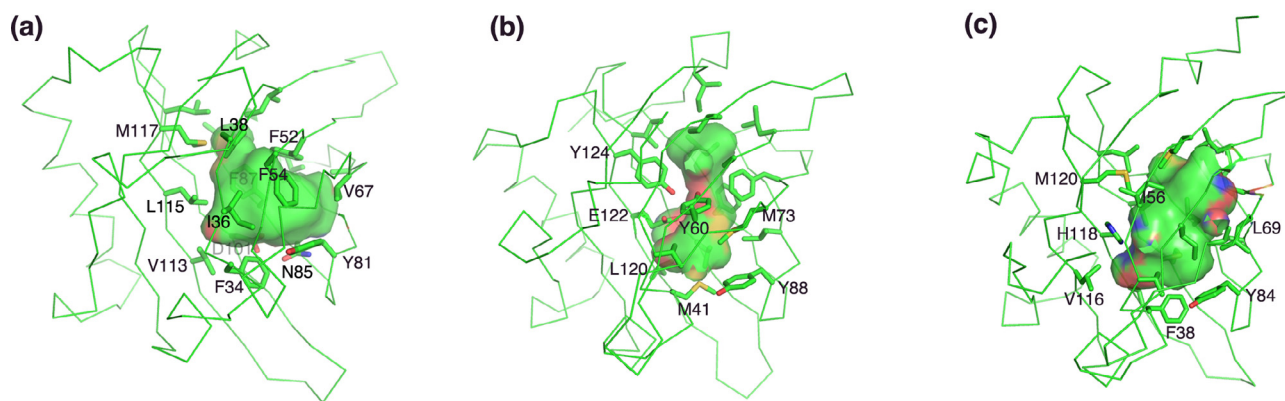


Fig. 6. The ligand-binding cavity of three lipocalin allergens: (a) Can f 4, (b) Equ c 1, and (c) Can f 2. The major residues that participate in definition of the cavity are shown as stick models.

these proteins would preferentially bind structurally clearly different ligands. However, the amino acid residues that participate in the formation of the cavities are clearly predominantly hydrophobic in all three lipocalin allergens. There are only one or two hydrophilic amino acid side-chains in the cavities, all located at the C-terminal β -strands F, G, or H. Can f 4 contains Asn85 in β -strand F and Asp101 in β -strand G. Equ c 1 contains Glu122 in β -strand H, and Can f 2 contains His118 in β -strand H. The comparison of amino acid sequences of the three lipocalin allergens Can f 4, Equ c 1 and Can f 2 shows that the residues which participate in the formation of the cavities are very well conserved, to be more precise, the position in the sequence (and in the structure) and the hydrophobic character of the amino acids are conserved (Fig. 4). The entrance to the central cavity is closed by Tyr81. It has an equivalent residue in Equ c 1 (Tyr88) and in Can f 2 (Tyr84). It is possible that the conformational change in the side-chain of this tyrosine may allow access to the cavity. Can f 4 has a low amino acid sequence identity with other dog lipocalin allergens (the three-dimensional structures of which are not known) Can f 1 and Can f 6, namely 17 and 26%, respectively. It can be assumed that the amino acids which participate in the formation of the ligand-binding cavity in Can f 1 and Can f 6 are located in similar positions in the amino acid sequence as in other lipocalin allergens. The sequence comparison suggests that Can f 1 and Can f 6 would have a different ligand-binding site. However, Can f 6 has surprisingly similar residues to those of Equ c 1 in these positions (Fig. 4).

The crystal structures of the following mammalian lipocalin allergens have previously been determined: horse Equ c 1 (Lascombe et al., 2000), bovine Bos d 2 (Rouvinen et al., 1999), Bos d 5, a bovine milk allergen (Brownlow et al., 1997), mouse Mus m 1 (Bocskei et al., 1992), rat Rat n 1 (Chaudhuri et al., 1992), and dog Can f 2 (Madhurantakam et al., 2010). Structure comparisons reveal that mammalian lipocalin allergens share relatively similar binding pockets with a high content of aromatic or aliphatic amino acid residues. The structures of lipocalin allergens from insects, Bla g 4 and Per a 4 from cockroach (Tan et al., 2008) and Arg r 1 from pigeon tick (Hilger et al., 2005), differ significantly from mammalian lipocalin allergens.

3.6. The dimeric nature of rCan f 4

An allergic reaction is triggered by cross-linking of Fc ϵ RI-bound IgE antibodies on mast cells and basophil leukocytes. During the process, the dimerization of an allergen is probably a favourable structural feature because cross-linking is then possible with two identical IgE antibodies. If an allergen is monomeric (asymmetric), two different IgE antibodies are required for cross-linking. rCan f 4 was crystallized with three protein molecules (named A, B,

and C) in the asymmetric unit. The packing of Can f 4 monomers in the crystal was studied by the use of a PISA server (Krissinel and Henrick, 2007). The largest interface area was found between molecules A and C (640 Å²). In addition, there is a two-fold symmetry axis between these two molecules. There are no other two-fold symmetry axes between the molecules in the crystal, and therefore we suppose that the dimer formed by A and C molecules in the asymmetric unit corresponds most probably with the dimeric structure of Can f 4. The moderate size of the monomer–monomer interface suggests that Can f 4 is able to form weak or transient dimers (Rouvinen et al., 2010). Consequently, there would be a concentration dependent equilibrium of monomers and dimers in solution. This would explain why the crystal structure contained both monomers and dimers of Can f 4.

The dimer interface of Can f 4 comprises regions from four adjacent strands of the central β -barrel (E, F, G, and H). The α -helices are located on the same side of the dimer and because of the two-fold symmetry they run in opposite directions (Fig. 7a). Hydrophobic residues located on the β -strands, Leu76 (strand E), Gly84, Tyr86 (strand F), Val102, Val104 (strand G), and Ile110 (strand H), form the core of the dimeric interface (Figs. 4 and 7d). Asp74 and Lys107 may form a salt bridge on the outer region of the interface. Although the ligand-binding cavity is close to the dimeric interface, it is not buried upon dimerization and is thus available for ligand binding.

When other crystal structures of lipocalin allergens were compared it was found that the Can f 4 dimeric structure resembles only the symmetric dimer reported for the major horse allergen Equ c 1 (Lascombe et al., 2000). The monomer–monomer interface in Equ c 1 is also formed by the same β -strands (E, F, G, and H) (Fig. 7b and e). The amino acid residues on the interface are hydrophobic and different but located in similar positions in the sequence (Fig. 4). The buried area of rCan f 4 monomers in a dimer covers 640 Å² of the exposed molecular surface, which is 8% of the total molecular surface of the monomer (7650 Å²). The monomer–monomer interface area of the Equ c 1 dimer is wider in comparison with that of rCan f 4. This is caused by the more extensive packing of the Equ c 1 monomers to a dimer mediated by additional hydrophobic residues (Fig. 7e), compared to Can f 4. The interface is 1025 Å² (13% of the molecular surface of a monomer (8135 Å²), which would suggest a more stable dimer than that of Can f 4.

Lipocalin allergens widely exist as dimers or higher oligomers. For example, Equ c 1, Bos d 5, Per a 4, Bla g 4, and Mus m 1 have been described to be dimers, whereas Rat n 1 forms a tetramer in the crystal. The comparison of known dimeric lipocalin structures shows that the monomer–monomer interfaces differ considerably among the structures and that it is not possible to distinguish any universal manner for lipocalin dimerization. In the crystal structures of dimeric lipocalin allergens, the monomer–monomer

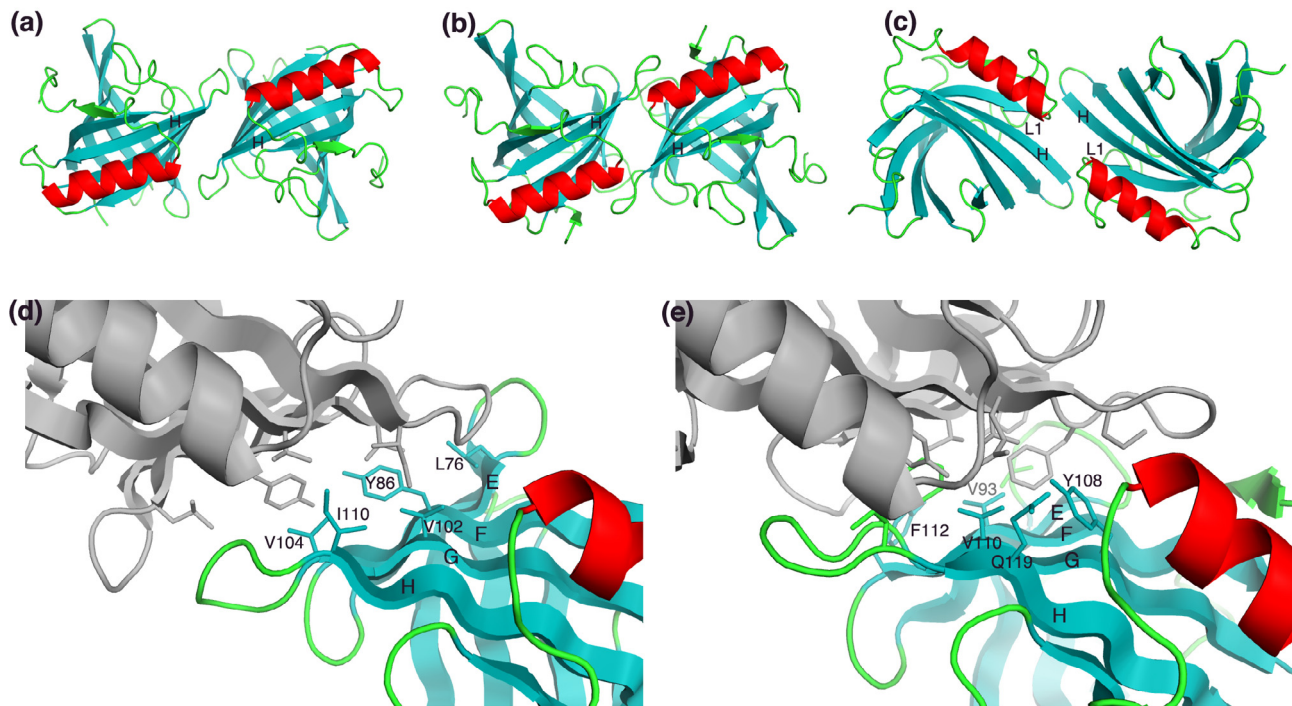


Fig. 7. Dimerization of lipocalin allergens: (a) Can f 4, (b) Equ c 1, and (c) Can f 2. The view is through a 2-fold symmetry axis. Dimerization of Can f 4 is similar to Equ c 1 but clearly different from that of the Can f 2 allergen. The residues that form the monomer–monomer interface are different in Can f 4 (d) and in Equ c 1 (e). This interface is also more extensive in Equ c 1.

interface areas vary from 390 to 1025 Å² (Rouvinen et al., 2010). As an example of different oligomerization, the rCan f 2 allergen from dog was crystallized in the orthorhombic space group C222₁ which in this case had one protein molecule in an asymmetric unit. Can f 2 was reported to be a monomeric allergen, at least at nearly neutral pH 6.0. The authors did not exclude the possibility that rCan f 2 can form a dimer at a lower pH, since the monomer–dimer ratio can also change as a function of pH. The PISA server at EBI-MSD (Krissinel and Henrick, 2007), which can be used to find macromolecular assemblies in crystal structures, predicts that Can f 2 exists as a monomer. Another option for searching dimers in crystal structures is to use the rotational symmetry axis between the monomers as a search criterion (Rouvinen et al., 2010). Using this method, one possible dimer was detected in the crystal structure of rCan f 2. Its structure is clearly different from the dimer structures of Can f 4 and Equ c 1 (Fig. 7c). The rCan f 2 dimer especially forms through loop L1 and β-strand H. Interactions are predominantly based on hydrogen bonds and also on main-chain atoms. Interestingly, the crystal structure of rCan f 2 shows also a homotetrameric structure with C2 symmetry which would indicate that at higher protein concentrations even larger oligomeric structures are possible for Can f 2. The three-dimensional structures as well as possibly quaternary structures of Can f 1 and Can f 6 allergens are not known. However, when the amino acid sequences are compared (Fig. 4), it can be noted that the hydrophobic residues which contribute to the formation of the monomer–monomer interface in Can f 4 and Equ c 1 are not hydrophobic in Can f 1 and Can f 6. This suggests that Can f 1 and Can f 6 are not able to form similar dimers, as Can f 4 and Equ c 1 do.

Dimers can be classified as weak transient dimers if the interface area covers less than 1000 Å² of a monomer surface (Nooren and Thornton, 2003). Changes in the local environmental conditions, like solution pH or temperature, or an increase in the protein concentration, may affect the monomer–dimer equilibrium. For example, a protein that mainly exists as a monomer at a low protein concentration can dimerize transiently when the protein

concentration increases. The local increment in the protein concentration would be reached, obviously, in crystallization but also in colocalization within the cells (Kuriyan and Eisenberg, 2007). Both the determined crystal structure and native mass spectrometric analysis suggest that rCan f 4 is a transient dimer. Because the dimerization mode of Can f 4 is similar to that of another major allergen, Equ c 1, this suggests that this kind of dimerization of a lipocalin may contribute to IgE binding and subsequent mast cell and basophil activation. The significance of allergen dimerization would be smaller if IgE response is polyclonal.

Acknowledgments

We acknowledge the European Synchrotron Radiation Facility for provision of synchrotron radiation facilities and we would like to thank the staff of beamline ID-29 for their assistance in data-collection. We are grateful to Ritva Romppanen and Mirja Saarelainen for their skilful technical assistance and to Dr. Jukka Randell for the acquisition of dog-allergic patients' sera. This work was supported by the Academy of Finland (project 137988).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.molimm.2014.04.003>.

References

- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst. D66*, 213–221.
- Arbes, S.J., Cohn, R.D., Yin, M., Muilenberg, M.L., Friedman, W., Zeldin, D.C., 2004. Dog allergen (Can f 1) and cat allergen (Fel d 1) in US homes: results from the national survey of lead and allergens in housing. *J. Allergy Clin. Immunol.* 114, 111–117.

- Bocskei, Z., Groom, C.R., Flower, D.R., Wright, C.E., Phillips, S.E., Cavaggioni, A., Findlay, J.B., North, A.C., 1992. Pheromone binding to two rodent urinary protein revealed by X-ray crystallography. *Nature* 360, 186–188.
- Brownlow, S., Morais Cabral, J.H., Cooper, R., Flower, D.R., Yewdall, S.J., Polikarpov, I., North, A.C., Sawyer, L., 1997. Bovine beta-lactoglobulin at 1.8 Å resolution – still an enigmatic lipocalin. *Structure* 5, 481–495.
- Chaudhuri, B.N., Kleywegt, G.J., Bjorkman, J., Lehman-McKeeman, L.D., Oliver, J.D., Jones, T.A., 1992. The structures of alpha 2u-globulin and its complex with hyaline droplet inducer. *Acta Cryst. D55*, 753–762.
- De Groot, H., Goei, K.G.H., Van Swieten, P., Aalberse, R.C., 1991. Affinity purification of a major and minor allergen from dog extract: serologic activity of affinity-purified Can f 1-depleted extract. *J. Allergy Clin. Immunol.* 87, 1056–1065.
- Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., Liang, J., 2006. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucl. Acids Res.* 34, W116–W118.
- Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. *Acta Cryst. D66*, 486–501.
- Flower, D.R., 1996. The lipocalin protein family: structure and function. *Biochem. J.* 318, 1–14.
- Heck, A.J.R., 2008. Native mass spectrometry: a bridge between interactomics and structural biology. *Nat. Methods* 5, 927–933.
- Hilger, C., Bessot, J.C., Hutt, N., Grigioni, F., De Blay, F., Pauli, G., Hentges, F., 2005. IgE-mediated anaphylaxis caused by bites of the pigeon tick *Argas reflexus*: cloning and expression of the major allergen Arg r 1. *J. Allergy Clin. Immunol.* 115, 617–622.
- Hilger, C., Swiontek, K., Arumugam, K., Lehnert, C., Hentges, F., 2012. Identification of a new major dog allergen highly cross-reactive with Feld 4 in a population of cat- and dog-sensitized patients. *J. Allergy Clin. Immunol.* 129, 1149–1152.
- Kabsch, W.J., 1993. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constant. *J. Appl. Cryst.* 26, 795–800.
- Konieczny, A., Morgenstern, J.P., Bizinkauskas, C.B., Lilley, C.H., Brauer, A.W., Bond, J.F., Aalberse, R.C., Wallner, B.P., Kasaian, M.T., 1997. The major dog allergens, Can f 1 and Can f 2, are salivary lipocalin proteins: cloning and immunological characterization of the recombinant forms. *Immunology* 92, 577–586.
- Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774–797.
- Kuriyan, J., Eisenberg, D., 2007. The origin of protein interactions and allostery in colocalization. *Nature* 450, 983–990.
- Lascombe, M.-B., Grégoire, C., Poncet, P., Tavares, G.A., Rosinski-Chupin, I., Rabillon, J., Goubran-Botros, H., Mazie, J.-C., David, B., Alzari, P.M., 2000. Crystal structure of the allergen Equ c 1. A dimeric lipocalin with restricted IgE-reactive epitopes. *J. Biol. Chem.* 275, 21572–21577.
- Liu, J., Konermann, L., 2011. Protein–protein binding affinities in solution determined by electrospray mass spectrometry. *J. Am. Soc. Mass Spectrom.* 22, 408–417.
- Lovell, S.C., Davis, I.W., Arendal III, W.B., de Bakker, P.I.W., Word, J.M., Prisant, M.G., Richardson, J.S., Richardson, D.C., 2003. Structure validation by Cα geometry: phi, psi and Cβeta deviation. *Proteins* 50, 437–450.
- Madhurantakam, C., Nilsson, O.B., Uchtenhagen, H., Konradsen, J., Saare, T., Högbom, E., Sandalova, T., Grönlund, H., Achour, A., 2010. Crystal structure of the dog lipocalin allergen Can f 2: implications for cross-reactivity to the allergen Fel d 4. *J. Mol. Biol.* 401, 68–83.
- Mattsson, L., Lundgren, T., Everberg, H., Larsson, H., Lidholm, J., 2009. Prostatic kallikrein: a new major dog allergen. *J. Allergy Clin. Immunol.* 123, 362–368.
- Mattsson, L., Lundgren, T., Olsson, P., Sundberg, M., Lidholm, J., 2010. Molecular and immunological characterization of Can f 4: dog dander allergen cross-reactive with a 23 kDa odorant-binding protein in cow dander. *Clin. Exp. Allergy* 40, 1276–1287.
- McCoy, A.J., Grosse-Kuntleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. *J. Appl. Cryst.* 40, 658–674.
- Morris, D.O., 2010. Human allergy to environmental pet danders: a public health perspective. *Vet. Dermatol.* 21, 441–449.
- Nicholas, C.E., Wegienka, G.R., Havstad, S.L., Zoratti, E.M., Ownby, D.R., Johnson, C.C., 2011. Dog allergen levels in homes with hypoallergenic compared with nonhypoallergenic dogs. *Am. J. Rhinol. Allergy* 25, 252–256.
- Nooren, I.M.A., Thornton, J.M., 2003. Structural characterization and functional significance of transient protein–protein interactions. *J. Mol. Biol.* 325, 991–1018.
- Radauer, C., Bublin, M., Wagner, S., Mari, A., Breiteneder, H., 2008. Allergens are distributed into few protein families and possess a restricted number of biochemical functions. *J. Allergy Clin. Immunol.* 121, 847–852.
- Rouvinen, J., Rautiainen, J., Virtanen, T., Zeiler, T., Kauppinen, J., Taivainen, A., Mäntytjärvi, R., 1999. Probing the molecular basis of allergy. Three-dimensional structure of the bovine lipocalin allergen Bos d 2. *J. Biol. Chem.* 274, 2337–2343.
- Rouvinen, J., Jänis, J., Laukkanen, M.-L., Jylhä, S., Niemi, M., Päivinen, T., Mäkinen-Kiljunen, S., Hahtela, T., Söderlund, H., Takkinen, K., 2010. Transient dimers of allergens. *PLoS ONE* 5, e9037.
- Rytkönen-Nissinen, M., Saarelainen, S., Randell, J., Häyrynen, J., Kalkkinen, N., Virtanen, T., submitted for publication.
- Saarelainen, S., Taivainen, A., Rytkönen-Nissinen, M., Auriola, S., Immonen, A., Mäntytjärvi, R., Rautiainen, J., Kinnunen, T., Virtanen, T., 2004. Assessment of recombinant dog allergens Can f 1 and Can f 2 for the diagnosis of dog allergy. *Clin. Exp. Allergy* 34, 1576–1582.
- Schou, C., Svendsen, U.G., Lowenstein, H., 1991. Purification and characterization of the major dog allergen Can f 1. *Clin. Exp. Allergy* 21, 321–328.
- Spinelli, S., Ramoni, R., Grolli, S., Bonicel, J., Cambillau, C., Tegoni, M., 1998. The structure of the monomeric porcine odorant binding protein sheds light on the domain swapping mechanism. *Biochemistry* 37, 7913–7918.
- Spitzauer, S., Schweiger, C., Sperr, W.R., Pandjaitan, B., Valent, P., Muhl, S., Ebner, C., Scheiner, O., Kraft, D., Rumpold, H., Valenta, R., 1994. Molecular characterization of dog albumin as a cross-reactive allergen. *J. Allergy Clin. Immunol.* 93, 614–627.
- Tan, Y.W., Chan, S.L., Ong, T.C., Yit, L.Y., Tiong, Y.S., Chew, F.T., Sivaraman, J., Mok, Y.K., 2008. Structures of two major allergens, Bla g 4 and Per a 4, from cockroaches and their IgE binding epitopes. *J. Biol. Chem.* 284, 3148–3157.
- Van den Heuvel, R.H.H., Heck, A.J.R., 2004. Native protein mass spectrometry: from intact oligomers to functional machineries. *Curr. Opin. Chem. Biol.* 8, 519–526.
- Virtanen, T., Kinnunen, T., Rytkönen-Nissinen, M., 2012. Mammalian lipocalin allergens – insight into their enigmatic allergenicity. *Clin. Exp. Allergy* 42, 494–504.
- Vredegoor, D.W., Willemse, T., Chapman, M.D., Heederik, D.J.J., Krop, E.J.M., 2012. Can f 1 levels in hair and homes of different dog breeds: lack of evidence to describe any dog breed as hypoallergenic. *J. Allergy Clin. Immunol.* 130, 904–909.
- Wagner, S., Radauer, C., Bublin, M., Hoffmann-Sommergruber, K., Kopp, T., Greisenegger, E.K., Vogel, L., Vieths, S., Scheiner, O., Breiteneder, H., 2006. Naturally occurring hypoallergenic Bet v 1 isoforms fail to induce IgE responses in individuals with birch pollen allergy. *J. Allergy Clin. Immunol.* 121, 246–252.
- Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Pottorff, E.A., Powell, H.R., Read, R.J., Vagin, A., Wilson, K.S., 2011. Overview on the CCP4 suite and current developments. *Acta Cryst. D67*, 235–242.
- Ylönen, J., Virtanen, T., Rytkönen, M., Mäntytjärvi, R., 1994. Quantification of a major bovine allergen by a two-site immunometric assay based on monoclonal antibodies. *Allergy* 49, 707–712.

