PLOS ONE

RESEARCH ARTICLE

Structural characteristics of lipocalin allergens: Crystal structure of the immunogenic dog allergen Can f 6

Gina M. Clayton[1,2,3], **Janice White**[1], **Schuyler Lee**[1], **John W. Kappler**[1,2,4], **Sanny K. Chan**[1,5,6]*

1 Department of Biomedical Research, National Jewish Health, Denver, Colorado, United States of America, **2** Program in Structural Biology and Biochemistry, University of Colorado Denver School of Medicine, Aurora, Colorado, United States of America, **3** Department of Biochemistry and Molecular Genetics, University of Colorado Denver School of Medicine, Aurora, Colorado, United States of America, **4** Department of Immunology and Microbiology, University of Colorado Denver School of Medicine, Aurora, Colorado, United States of America, **5** Department of Pediatrics, University of Colorado Denver School of Medicine, Aurora, Colorado, United States of America, **6** Division of Pediatric Allergy-Immunology, National Jewish Health, Denver, Colorado, United States of America

* chans@njhealth.org

Check for updates

OPEN ACCESS

Citation: Clayton GM, White J, Lee S, Kappler JW, Chan SK (2019) Structural characteristics of lipocalin allergens: Crystal structure of the immunogenic dog allergen Can f 6. PLoS ONE 14 (9): e0213052. https://doi.org/10.1371/journal.pone.0213052

Editor: Alexander Wlodawer, NCI at Frederick, UNITED STATES

Received: February 18, 2019

Accepted: August 13, 2019

Published: September 16, 2019

Copyright: © 2019 Clayton et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Lipocalins represent the most important protein family of the mammalian respiratory allergens. Four of the seven named dog allergens are lipocalins: Can f 1, Can f 2, Can f 4, and Can f 6. We present the structure of Can f 6 along with data on the biophysical and biological activity of this protein in comparison with other animal lipocalins. The Can f 6 structure displays the classic lipocalin calyx-shaped ligand binding cavity within a central β-barrel similar to other lipocalins. Despite low sequence identity between the different dog lipocalin proteins, there is a high degree of structural similarity. On the other hand, Can f 6 has a similar primary sequence to cat, horse, mouse lipocalins as well as a structure that may underlie their cross reactivity. Interestingly, the entrance to the ligand binding pocket is capped by a His instead of the usually seen Tyr that may help select its natural ligand binding partner. Our highly pure recombinant Can f 6 is able to bind to human IgE (hIgE) demonstrating biological antigenicity.

Introduction

Analyses over the last 2 decades have revealed that most allergens can be grouped into a few families of proteins[1]. One such family is made up of lipocalins. These are amongst the most important inhalant animal allergens. As far as their functions in their host of origin are concerned, these, depending on the lipocalin involved, include a number of properties such as pheromone transport, prostaglandin synthesis, retinoid binding, odorant binding and cancer cell interactions[1]. Lipocalins are small proteins, usually composed of less than 200 amino acids. In spite of their shared ability to act as allergens they have limited sequence homology with each other, even as low as 20%. Yet for the lipocalins whose structures have been

determined, they have a common tertiary structure consisting of an 8 stranded antiparallel beta barrel with a hydrophobic ligand binding pocket[2]. Human sensitization has been identified against lipocalins from cows, guinea pigs, horses, cats, mice, rats[3] and several arthropods[4]. Human sensitization to dogs and cats is a known major risk factor for the development of asthma and allergic rhinitis. Because the prevalence of allergic disease to furry animals unfortunately continues to increase worldwide[5–7], it will be useful to expand our knowledge of the structure and function of the lipocalins that are most often involved in such human diseases.

There are 7 sequentially named dog component allergens, Can f 1 through Can f 7[5] and 4 of them are lipocalins: Can f 1, Can f 2, Can f 4, and Can f 6. The ligands and biological functions of these lipocalins in vivo are poorly defined. One study of individuals sensitized to unfractionated dog dander showed subjects had IgE that bound different dog component allergens, for example, 64% were allergic to Can f 1 while only 27% were allergic to Can f 3[8]. Symptoms from patients when exposed to dogs range from those of asthma to no symptoms at all, with responses depending to some extent on the dog allergens to which the sensitized individuals were exposed[9].

In the studies described here, we chose to focus on one of the dog lipocalin allergens, Can f 6. Recent reports suggests that Can f 6 is a key lipocalin responsible for allergic symptoms[10]. Responses to Can f 6 are present in between 38%[11] and 56.3%[12] of dog sensitized individuals and Can f 6 allergic individuals are more likely to suffer from allergic rhinitis symptoms at a rate of 86% vs non-sensitized at 53% and those with asthma symptoms are found to be sensitized at 64% vs 38% in non-sensitized[10]. Therefore, Can f 6 is a major allergen driving disease amongst those people who are allergic to dogs. Moreover, in some individuals, antibodies to the cat lipocalin, Fel d 4, and/or the horse lipocalin, Equ c 1, cross react with Can f 6 suggesting some structural and sequence similarities between the 3 lipocalins[11,13]. Since mouse Mus m 1 cross reacts with Equ c 1[14], it is likely that there is also cross reactivity with Can f 6. Although the structure for Fel d 4 is unknown, the ones for Equ c 1[15] and Mus m 1[16] are, so a comparison of the structures of Equ c 1 and Mus m 1 with Can f 6 might reveal the reasons for their immunological cross reactivity.

Here the structure of the clinically relevant Can f 6 allergen was determined at 2 Å (Table 1) resolution along with human antibody binding data that shows its biologic activity. Although structurally similar to other dog lipocalins, Can f 6's primary sequence is more similar to other mammalian lipocalins. Unique regions in the capping region and shape of the ligand binding pocket may play a role in ligand recognition, binding, and ultimate function in vivo. We also describe both unique and common features of Can f 6 and similar lipocalins that underlie their allergenicity.

Material and methods

Cloning, production, purification, and biochemical characterization

The cDNA sequence based on *Can f 6*[11] was synthetically created from KEGG cfa: 481674 encompassing the entire protein but without the signal sequence and cloned with a C-terminal His_6-tag. An additional valine was added before the start methionine codon and 2 amino acids, leucine and aspartic acid, was added to keep the fragment in frame when using the restriction sites. Can f 6 was then expressed in the periplasmic space of the Rosetta strain of *E. coli*, as previously described, with minor modifications[17]. Can f 6 was purified using a His-Pur Ni-NTA Superflow Agarose nitrilotriacetic acid (NTA) column, followed by size exclusion in PBS on a Superdex 10/300 column (GE), and anion exchange chromatography (GE) in 10mM Tris pH 8 with a 10mM Tris/10 mM Tris and1M NaCl gradient. The purified protein,

Table 1. Crystal data and model refinement statistics for the structure of Can f 6.

Crystal	Overall	High res bin	Highest resolution shell
PBD Accession Code: 5X7Y			
Mean I/Sigma (I)	40.58	3.98	
Wavelength (Å)	0.9791		
Low resolution limit (Å)	58.16	-	58.16
High resolution limit (Å)	2.06	2.16–2.05	2.05
Rpim	0.012	0.194	
R/R Free			24.17/25.46
Ramachandran			99.37% Favored
			0.63% Allowed
			0.0 outlier,
Total reflections	206632	4293	
Total unique	12122	872	
CC(1/2)	0.997	0.93	
CC1/2 in the highest shell		2.06	
Completeness overall (%)	99.4	93	
Multiplicity	17	4.9	
Average B			62
RMSD bonds (Å)			0.0048
RMSD angles (°)			1.112

Cell (Å, °) a = 67.15 b = 67.22 c = 72.64, α = 90 β = 90 γ = 120
Space group P3121

https://doi.org/10.1371/journal.pone.0213052.t001

Can f 6 with the C-terminal His_6-tag, was utilized in all subsequent experiments. It was subjected to SDS-PAGE followed by excision of the band, tryptic digestion, peptide ionization and its identification was analyzed on a Q Exactive quadrupole orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) coupled to an Easy nLC 1000 UHPLC (Thermo Fisher Scientific) through a nanoelectrospray ion source. The MS/MS spectra were extracted from raw data files and converted into mgf files using Proteome Discoverer 2.2. These mgf files were then independently searched against the sequence of CF6 protein using an in-house Mascot™ server (Version 2.6, Matrix Science) and identified 96% of the full Can f 6 sequence.

Measurements of the thermal stability of Can f 6 were made using an Agilent Stratagene MX3005P qPCR thermal shift assay as previously described[18]. Briefly, the fluorescence of Sypro Orange was monitored as temperatures were increased, leading to unfolding of the Can f 6 protein.

Crystallization

Purified recombinant Can f 6 was buffer exchanged into 10mM Tris pH 8 and 100mM NaCl and concentrated to 15mg/ml. Initial crystallization trials were undertaken at room temperature using the mosquito LCP Nanoliter Protein crystallization robot. Crystals were grown and used to seed larger crystals using the same solutions but by the hanging drop vapor diffusion method containing a reservoir solution of 6% v/v 2-propanol, 0.1M sodium acetate trihydrate pH 4.5, and 26% v/v polyethylene glycol monomethyl ether 550.

Data collection, structure determination and refinement

Can f 6 crystals were harvested and cryoprotected in a reservoir solution containing 35% PEG 500 MME followed by storage in liquid N_2. Data were collected on a single crystal at the NE-CAT 24-ID-C beam line Advanced Photon Source, Argonne National Laboratory. Data were processed in X-ray Detector Software (XDS)[19], the scaling program AIMLESS[20], and further analyzed in Phenix Xtriage[21].

Structural determination

The structure of Can f 6 was determined by molecular replacement in Phaser[22] using Can f 4 (PDB ID: 4ODD)[23] as an initial search model. The resulting solution was further modelled and refined in the CCP4[24] suite of programs using Coot and Refmac. Phenix Validation and Refine[21] were used to confirm the final structure. The crystal packing of Can f 6 was analyzed in PISA[25] and its secondary structure was analyzed in STRIDE[26]. See Table 1 for data and refinement statistics. Pymol was used for analysis of the structure and to generate figures[27]. The atomic coordinates and the structure factors have been deposited in the RCSB Protein Data Bank with accession code 6NRE.

The amino acid residues of the ligand-binding sites and the volumes of the central cavities were determined using CASTp[28] and Voidoo[29] used to generate a cavity volume map and visualized as Connolly surfaces[30]. Post-translational modifications were analyzed with NetNGlyc 1.0 server[31] for potential N-glycosylated sites, NetOGlyc 4.0 server was used to predict GalNAc O-glycosylation sites[32], NetPhos 3.1 Server[33] was used to predicted phosphorylation sites, and Sulfinator[34] was used to predict tyrosine sulfation.

Sequence alignments for Can f 6, Can f 2, Can f 4, Fel d 4, Equ c 1, and Mus m 1 was performed in Clustal Omega[35]. Superpositions of Can f 6, Can f 2 (PDB ID: 3L4R), Can f 4 (PDB ID: 4ODD), Equ c 1 (PDB ID: 1EW3), and Mus m 1 (PDB ID: 1QY0) were performed using Coot and Pymol.

Detection of human antibody binding by ELISA

Patients' sample data were downloaded from the National Jewish Health Research Database, Data Set Identifier: 473-06-20-2018. Human serum samples banked by National Jewish Health under National Jewish Health Institutional Review Board (NJH IRB, OHRP #IRB00000024) approved protocol were retrieved after NJH IRB approval. Sera from 5 patients who had dog IgE >5 kU/L (ImmunoCAP; Phadia, Uppsala Sweden) and skin prick testing >3x3mm to AP dog extract were selected at random for analysis. Sera from 5 age matched controls who had sIgE<0.35 kU/L and negative skin testing were then selected from the biobank. For detection of biological activity by ELISA, plates were coated with recombinant Can f 6 at 50 μg/ml overnight at 4˚C and then incubated with serially diluted sera. Bound IgE was detected using alkaline phosphatase-conjugated goat anti-human secondary antibodies (A18796, Invitrogen) followed by incubation with p-nitrophenyl phosphate and measurement by spectrophotometry. The average baseline background OD405nm was subtracted from all samples and readings from the 1:10 dilution was used for analysis of significance using an unpaired t-test.

Results and discussion

Overall structure of Can f 6

The recombinant Can f 6 crystallized in space group $P3_121$ with one monomer in the asymmetric unit (ASU) and was refined to a resolution of 2.06 Å. Final refinement and model statistics are in Table 1. The overall architecture of Can f 6 contains a highly conserved lipocalin

fold of an 8 stranded anti-parallel β-barrel enclosing a central ligand binding cavity core, a four-turn C-terminal α-helix that packs against the β-barrel, and a short 9th β-strand (β_I) at the C-terminus (Fig 1A). The top open end of the β-barrel, which marks the entrance to the presumptive ligand-binding cavity, is comprised of four β-hairpin turns. The other side of the β-barrel is closed by one β-hairpin turn, two smaller β-hairpin turns, and an N-terminal coil (Fig 1A). There is a disulfide bond between residues Cys69 and Cys162 that connects the C-terminal part of the protein to the central β-barrel. This overall structure is similar to that previously reported earlier this year PDB ID: 5X7Y[36] which has a resolution of 2.35 Å resolution compared to ours that is at 2.06 Å. The sequence of main secondary structures is the same in both structures although the span of the β strands, in particular the $\beta_{B\text{-}E,\ H\ \text{and}\ I}$ vary slightly in the formation of the calyx. Their $P4_1$ space group suggested a tetragonal shape while ours is trigonal/rhombohedral and likely due to the alternative locations of β sheet interactions. The c-terminal tail is disorganized in our structure and is not modeled, which is similar to their report. Additionally, an electron density of map of the unmodeled ligand (S1 Fig) is surrounded by hydrophobic amino acids.

In this structure, three lipocalin Structurally Conserved Regions (SCRs) are present (Fig 1A and 1B) that contain conserved residues. The SCRs hold together the N- and C-terminus of the overall fold (Fig 1A) under the ligand cavity calyx.

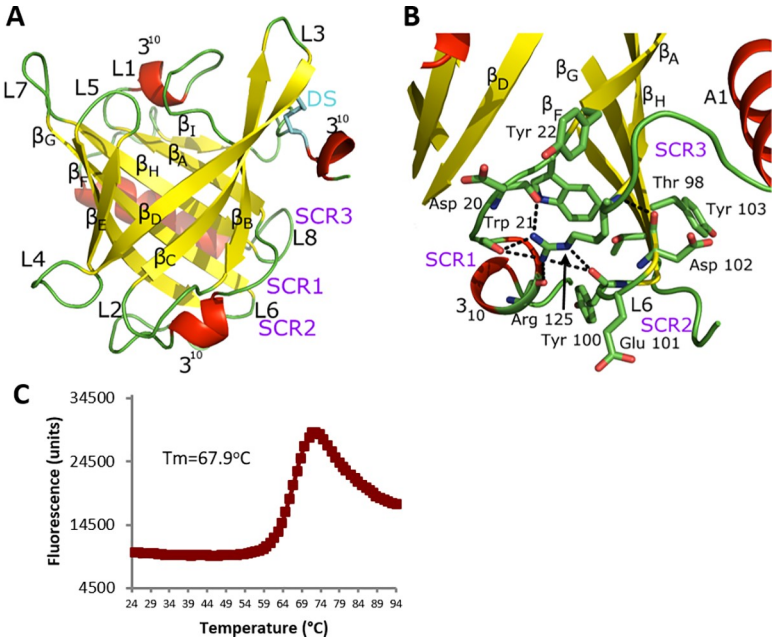


Fig 1. The overall architecture of Can f 6 displays the classic lipocalin structure. (A) Eight antiparallel β-strands (yellow) form a central ligand binding cavity ($\beta_{A\text{-}I}$) flanked by C-terminal α-helix (posteriorly located in the figure) and three 3_{10}-helices (red), with the connecting loops (L, in green) numbered. A disulfide bond (**DS**, cyan) clamps the C-terminal loop to the beta core via β_D. The cavity opening is covered by L1 and a loop-3_{10} helix-loop that links together β_A and β_D. Three core lipocalin Structurally Conserved Regions (**SCR, purple**) that hold the overall fold together are labeled: **SCR1** connects the N-terminal loop to the C-terminal **SCR3** which also interacts with **SCR2**. (**B**) Figure A is rotated by 90° out of the plane and shows the main residues of the core SCR. In **SCR3**, the side chain and distal guanidium group of a conserved Arg125 are stabilized by hydrogen bonds (dashed lines) to the backbone carbonyls of other SCR residues and a Cation-Pi interaction with a conserved **SCR1** aromatic residue, here Trp21. Other residues in the SCR also form numerous non-covalent interactions. **(C)** Purified Can f 6 was assayed by Thermofluor shift to demonstrate the stability of our recombinant Can f 6. The average Tm is 67.9˚C with standard deviation of 1.7˚C.

https://doi.org/10.1371/journal.pone.0213052.g001

By comparison, the Can f 6 cavity cap and calyx are more distinct in sequence and conformation when considered across all lipocalins. For example, in the structures of the lipocalins apo, holo Retinol-binding protein (RBP) [37] and apolioprotein M [38], the ligand cap conformations are different and have a more open architecture than the ligand cavity cap described here.

Can f 6 crystalized as a monomer reflecting the method of purification, on size exclusion chromatography at pH 8 where the protein of correct size was identified and further purified. The protein is quite stable to heat (Fig 1C) which is perhaps helped by the disulfide bond and the interactions of the SCR. PISA analysis of the crystal lattice packing around the Can f 6 monomer reveals 3 adjacent lattice packing molecules. The largest packing surface, at symmetry operator position Y, X, -Z, forms a dimer buried surface area of 736 $Å^2$ (S1 Table). The dimer interface is the same as one of the dimers observed in the other deposited structure of Can f 6 (PDB ID: 5X7Y) which forms a tetramer in the ASU.

Interface surfaces of less than 1000 $Å^2$ between monomers can be classified as having weak transient interactions[39]. The PISA analysis of the Can f 6 crystal did not reveal any specific interactions that might result in the formation of a stable quaternary structure in solution[25]. The dimeric interface of the Can f 6 structure is considered below that for the complexation criteria. However, the dimeric crystal packing interface consists of 27 residues and 14 hydrogen bonds from each monomer (S1 Table) and as such would be susceptible to changes in pH conditions during purification and crystallization.

To date, lipocalin structures reveal buried interface surfaces that vary considerably from 390–23475 $Å^2$[39] as dimers, trimers, or tetramers without a consistent quaternary pattern. The crystal lattice packing of Can f 6 does not resemble those observed in the structures of Can f 4[23] nor Equ c 1[15]. In the Can f 4 structure there is a dimer interface (largest interface of 640 $Å^2$)[23]. Similarly, in the Equ c 1 structure there is also a dimer interface (largest interface of 1070 $Å^2$)[15]. In the structure of Mus m 1 the dimer interface is much smaller (interface of 389 $Å^2$)[39].

The packing interfaces in the lipocalin structures may be physiologically relevant or may simply be a consequence of crystallization conditions. Differences in the monomeric and transient multimeric states of lipocalins are theorized to contribute to a protein's allergenicity and are likely unique to each allergen[4]. For example, the dimerization of the lipocalin allergen Bet v 1 has been demonstrated to increase allergenicity in skin tests of allergic mice compared to monomers[40]. However, the role of multimerization in allergenic lipocalins is inconsistent. It is not known if Can f 6 forms multimers in solution at different pH conditions as reported for other lipocalins[39].

Comparison of the structure and sequence of Can f 6 with those of other mammalian lipocalins

Some members of the lipocalin family have limited sequence homology with each other. For example, of the known dog lipocalins, Can f 2 is only 21% identical to Can f 4 and likewise Can f 6 is 25% identical to Can f 2 and 29% identical to Can f 4 (Figs 2A and 3). On the other hand, Can f 6 is 66%; 57%; and 47% identical to the cat lipocalin, Fel d 4; the major horse allergen, Equ c 1; and the mouse lipocalin Mus m 1 respectively (Figs 2B and 3). Cross reactivity with IgE antibody binding has been show between Can f 6, Fel d 4 and Equ c 1[11] as well as between Equ c 1 and Mus m 1 [14] which would also suggest that Can f 6 and Mus m 1 would also cross react.

The crystal structures of Can f 2 (PDB ID: 3L4R)[41], Can f 4 (PDB ID: 4ODD)[23], Equ c 1 (PDB ID: 1EW3)[15], and Mus m 1 (PDB ID: 1QY0)[16] have been determined previously.

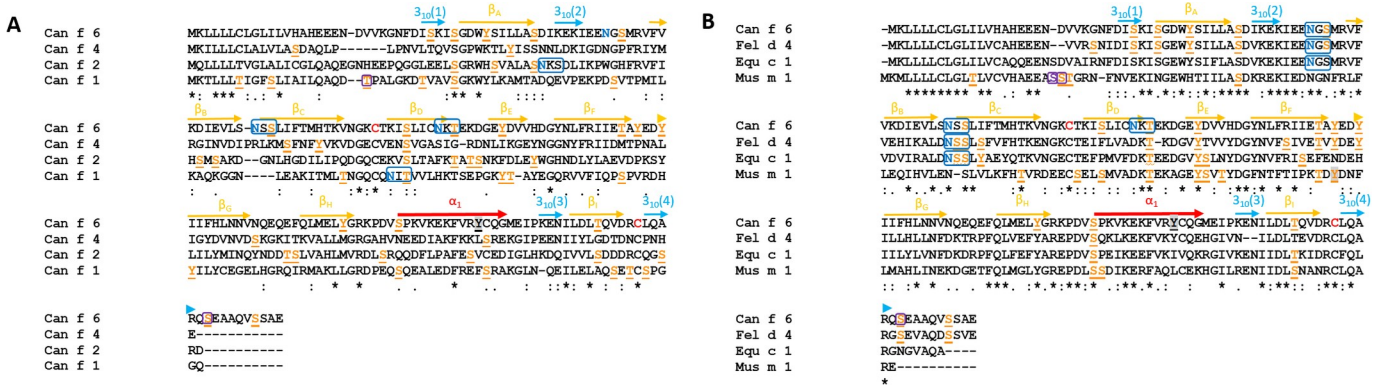


Fig 2. Primary structure alignment of lipocalins. (A) The secondary structure for Can f 6 is drawn above the sequences. The colors correspond to the model of Can f 6 depicted in Fig 1 (A). Despite their conserved lipocalin architecture there is low sequence similarity between the dog lipocalins Can f 1, 2, 4, and Can f 6. The only disulfide bond (C, in red) occurs between C67 and C160. **(B)** Primary sequence alignment of the Can f 6 dog lipocalin, cat (Fel d 4), horse (Equ c 1), and mouse (Mus m 1). There is much more sequence similarity in these lipocalins compared to those of the dog lipocalins. Alignment residues are as: "*" Identical amino acid, ":" Similar amino acid, and "." Slightly similar amino acid[35]. NxS/T (Blue N in blue boxes) are predicted N glycosylation sites (28), **S/T** (purple in black boxes) are predicted O glycosylation sites (29), S/T/T (orange) are predicted phosphorylation sites (30), and **Y** are predicted sulfation sites (31).

https://doi.org/10.1371/journal.pone.0213052.g002

Currently, no structure of Fel d 4 has been reported. Superposition of the structures of Can f 2, Can f 4, and Equ c 1 onto Can f 6 yields an rms deviation for Cα backbone atoms of 1.32, 1.48, and 1.38 Å respectively. The overall topology is similar for all 4 (Fig 4). The main structural differences are the β-strand turns and the folds preceding the C-terminal helices. For example, secondary structure analysis [26] of the lipocalins here, reveals that only Can f 2 has a short 3_{10} helix (res 124–126) preceding the C-terminal α-helix (Fig 2) whereas the differences in secondary structures of the other lipocalins likely interfere with 3_{10} helix formation. In Can f 6, Can f 4, and Equ c 1, the equivalent region is a turn towards the C-terminal α-helix. In both Can f 6 and Equ c 1, the turn has a Pro in β_F (127 in Can f 6 and 143 in Equ c 1) that stacks against a Tyr in β_A (22 in Can f 6 and 38 in Equ c 1). Pro127 stabilizes this turn via a backbone carbonyl hydrogen bond to the side chain of Arg159 (Arg175 in Equ c 1). By comparison Can f 4, has a

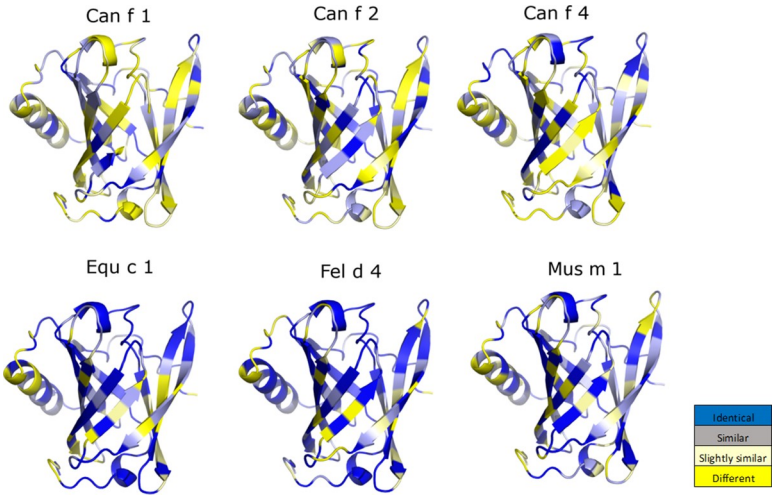


Fig 3. The sequence alignment of lipocalins threaded onto the structure of Can f 6. The sequence homology between the dog lipocalin Can f 6, 4, and 2 are more distinct compared to that between Can f 6 and Equ c 1, Fel d 4, and Mus m 1.

https://doi.org/10.1371/journal.pone.0213052.g003

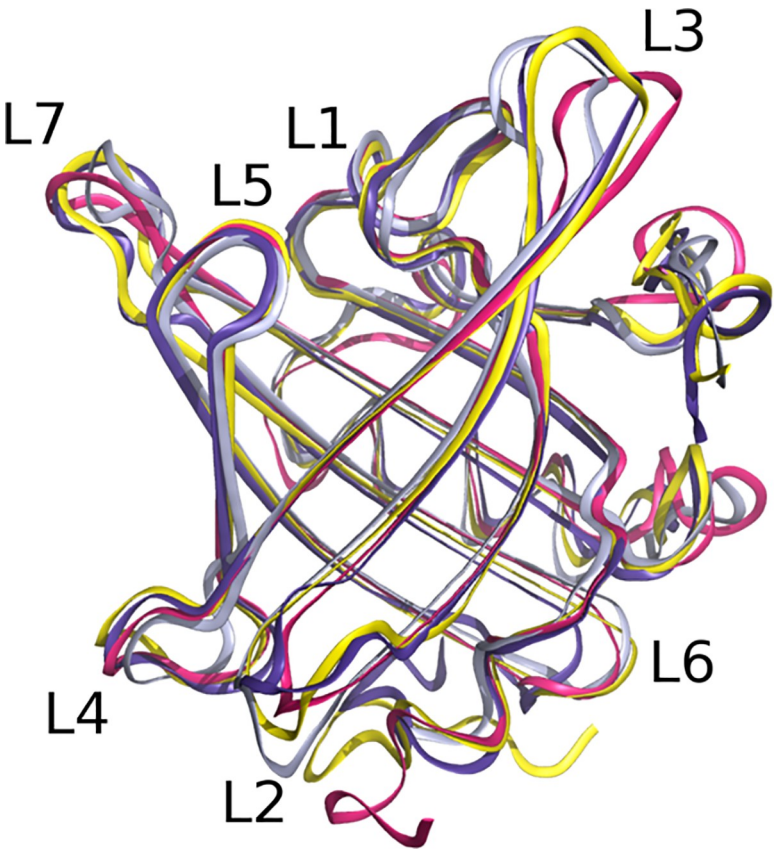


Fig 4. Structural overlay of Can f 6 (yellow) on Can f 2 (red) and Can f 4 (grey) Equ c 1 (purple), and Mus m 1 (brown). The overall architecture of lipocalins is highly conserved with a β-barrel that encloses a hydrophobic core. The major differences are outside the core.

https://doi.org/10.1371/journal.pone.0213052.g004

β_F Gly120 and Ala121 that packs against the β_A Lys16. These loops and turns are flexible in comparison the rigid barrel and though resolved, they are more susceptible to slight changes. In Can f 4's dimeric structure these locations will limit the ability for IgE antibody recognition and therefore binding thereby reducing its allergenicity compared to the monomeric form of Can f 6. However, it is unclear if these conformational differences are normally occurring in nature or due to crystal packing observed in the lattice.

Although lipocalins are of similar size (Table 2 and Fig 2), the residue differences mean they do not necessarily have a similar overall charge[42], immunogenic sites, nor post translational modification patterns. Predicted post translational modifications are shown (Figs 2 and 5). The predicted N-linked glycosylation pattern for Can f 6 is different when compared to the other dog lipocalins here (Figs 2 and 5). Two of the predicted N-linked glycosylation sites [43] of the immunologically cross-reactive Can f 6, Fel d 4, and Equ c 1 are at the similar sites, but are not present in Mus m 1 (Fig 5B). Recombinant Can f 6 was expressed in E. Coli and was not phosphorylated as determined by mass spec. However, the prediction algorithm for Can f 6 suggests numerous potential sites (Figs 2 and 5C).

These differences in glycosylation and phosphorylation will affect lipocalin solubility, ability to bind other proteins and may also be an important feature of antibody recognition and lipocalin immunogenicity. Glycosylation and phosphorylation may mask or alter the surface features of nearby residues[44]. For example, Mouse mAb 65 recognizes the natural glycosylated

Table 2. Molecular masses, numbers of positively and negatively charged amino acids and theoretical pI of Canf 6 and select lipocalins using ProtParam.

Protein	Total Amino Acids	Molecular Mass (kDa)	Positively charged amino acids (Arg + Lys)	Negatively charged amino acids (Asp + Glu)	pI	Resolution (Å)
Can f 6	179	20667	21	34	4.75	2 (Now)
Can f 4	159	17693	19	20	6.17	2.6 (20)
Can f 2	163	18448	14	27	4.82	1.45 (32)
Can f 1	156	17319	18	21	5.66	
Fel d 4	170	19573	19	31	4.76	
Equ c 1	172	20097	20	34	4.51	2.3 (15)
Mus m 1	163	18765	19	31	4.84	1.8 (16)

https://doi.org/10.1371/journal.pone.0213052.t002

Equ c 1 and not the recombinant unglycosylated Equ c 1[26]. Similarly, human tear lipocalins are known to be phosphorylated and changes in phosphorylation may play a role in lipocalin related diseases[43].

Comparison of the ligand binding cavities of some lipocalins

In many cases the biologically relevant lipocalin ligands are as yet unknown. Thus close examination of the properties of their ligand binding cavities, deduced from the presence of unresolved electron density and the shape of the cavity volume in these locations, is of interest. In the structure of Can f 6 there is unmodeled electron density inside the highly hydrophobic ligand binding cavity which is uniquely shaped (Fig 6A–6C). Since Can f 6, as described here,

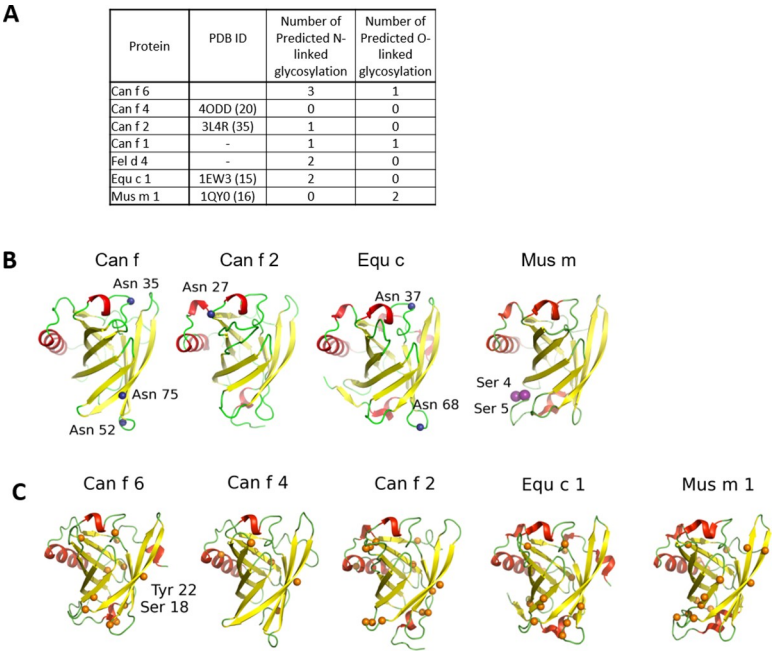
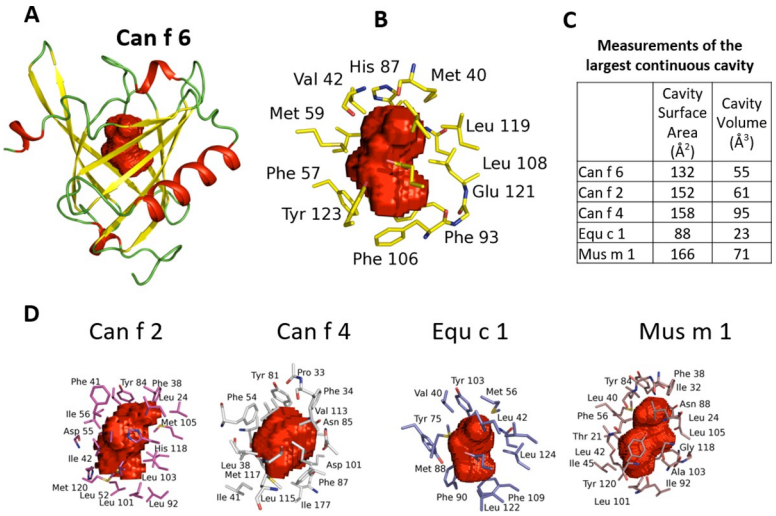


Fig 5. Comparison of predicted post translational modification sites. (A) Summary of glycosylation sites. **(B)** Three N-linked glycosylation residues (blue spheres) are predicted at Asn35, 52, and 75 in Can f 6; one in Can f 2 at Asn27; and two in Equ c 1 at Asn37 and 68. All these sites are on loops. The O-linked glycosylation site predicted in Can f 6 (purple spheres) is at the end of the C-terminus (Ser166) and is not modeled because this region of the structure is disordered. **(C)** Predicted phosphorylation sites (orange spheres). Phosphorylated Tyr22 and Ser18 in human tear lipocalin are present and labeled in Can f 6.

https://doi.org/10.1371/journal.pone.0213052.g005



	Cavity Surface Area (Å^2)	Cavity Volume (Å^3)
Can f 6	132	55
Can f 2	152	61
Can f 4	158	95
Equ c 1	88	23
Mus m 1	166	71

Fig 6. A comparison of the ligand binding cavities in the different lipocalins. Cavity residues and volume maps (red) were calculated using a rolling probe of 1.4Å and displayed as Connolly surfaces (25, 26, 27). **(A)** The overall ligand core is centrally located within the beta barrel and closed at the top by the L1 and L5 loops. **(B)** Residues from multiple strands of the β-barrel contribute to the ligand cavity core with hydrophobic amino acids and aromatic side chains in the interior of the ligand binding surface. **(C)** The ligand cavity volumes are fairly similar except for Equ c 1. **(D)** The ligand cavities of Can f 2 (red), Can f 4 (grey), Equ c 1 (purple), and Mus m 1 (brown) have a variable shape. Can f 6 and Equ c 1 have similar ligand cavity shape, with the slight differences between them likely reflects slightly differently shaped ligands.

https://doi.org/10.1371/journal.pone.0213052.g006

was expressed in the bacterial periplasm and purified in its native form, the electron density observed in the ligand pocket may be from a full or partially occupied crystallization reagent, or metabolite from the bacteria used for expression although the ligand in the cavity is not thought to be from the Poly Ethylene Glycol (PEG) used in the crystallization conditions since its shape is inconsistent. In either case the ligand molecule likely has similar properties to the natural binding ligand.

The central cavities of the lipocalins described here vary in size and shape reflecting differences and likely specificity for their natural ligands (Fig 6D). The cavity of Can f 6 has a solvent accessible surface area of approximately 132 Å^2 with a volume of 55.3 Å^3 calculated by CASTp [45]. These sizes are comparable to that of Can f 2 which has a surface area of 152.0Å^2 with a volume of 60.6Å^3 and smaller than Can f 4 which has a surface area of 158.3 Å^2 with a volume of 95.0Å$^{3.}$ (Fig 6C).

As observed for other lipocalins, the residues of Can f 6 ligand cavity are mainly hydrophobic. A total of 17 residues from the 8 antiparallel β-barrel are involved in defining the shape of the Can f 6 ligand cavity. Twelve of the 17 amino acids that line the cavity are hydrophobic with 4 polar and 1 negatively charged residue (Table 3). A comparison of the homologous

Table 3. Ligand binding pocket analysis of Can f 6 with the 17 main amino acid residues that define the shape of ligand binding cavity compared to homologous proteins.

Can f 6	26L	40M	42V	57F	59M	70L	72L	74C	85V	87H	91N	93F	106F	108L	119L	121E	123Y
Can f 4	Y	F	L	F	F	N	V	A	G	Y	N	F	Y	V	V	L	M
Can f 2	A	F	V	G	L	V	L	A	L	Y	N	L	L	M	V	H	M
Equ c 1	F	M	V	A	Y	F	M	F	L	Y	N	F	L	L	L	E	Y
Fel d 4	L	M	V	F	H	L	L	A	V	Y	N	F	L	L	L	E	Y

Only the amino acid position for Can f 6 is listed.

https://doi.org/10.1371/journal.pone.0213052.t003

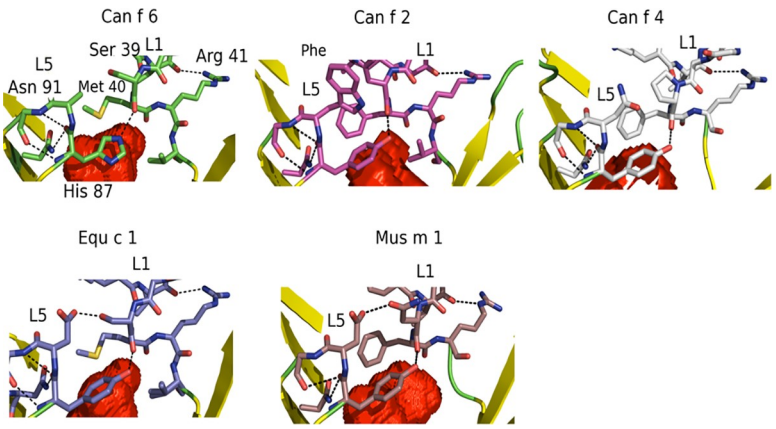


Fig 7. Comparison of the cavity capping residues of lipocalins. In Can f 6, at the entrance of the calyx cavity His87 in L5 stabilizes L1 over the ligand binding mouth. Arg41 and Asp91 stabilizes His87 and the L1 loop and are highly conserved indicating a common capping mechanism between lipocalins. The His in Can f 6 is unique compared to a highly conserved Tyr that is observed in the other lipocalins reflecting a unique ligand binding mechanism. Volume map is portrayed in red as Connolly surfaces (26).

https://doi.org/10.1371/journal.pone.0213052.g007

lipocalins show little sequence conservation of the cavity except for a conserved Asn91 (Table 3) and highly conserved Val42, Phe93, and Leu108 (Fig 6D)

At the cavity opening of the lipocalins discussed here, L1 and L5 close the cavity entrance (Fig 7). Unique to Can f 6 is His87 of L5 that closes the L1 loop over the ligand space via a hydrogen bond to the backbone carbonyl of Ser39 in L1. His87 is itself stabilized by intra L5 hydrogen bonds and L5 Asn91. The L1 lid loop is stabilized by numerous hydrogen bonds to itself and by L1 Arg41 binding to the backbone residues of the L1 3_{10} helix. In contrast, Can f 2, Fel d 4, Equ c 1, Mus m 1, and other lipocalins, have a highly conserved Tyr present instead of a His (Fig 7). Though Tyr and His are similar, this site may be important for ligand recruitment and recognition. Nearby Asn91 and Arg41 are highly conserved within the lipocalins (Fig 2) and alternate residues have strongly similar properties, indicating a highly conserved ligand calyx cavity capping mechanism for recognition and binding of ligands.

The calyx cap conformation in Can f 6, places Met40 of the L1 loop over the ligand cavity where it does not make significant internal interactions. In the other lipocalins, an L1 Phe sits over the ligand space (Fig 7) and these residues are thought to also be involved in ligand recognition and capture.

It is possible that the Can f 6 structure here represents a closed conformation with the unresolved internal electron density that could be a full or partial ligand. Other lipocalins have different conformations of L1 and L5. Compared to the structure of Can f 6, the structure of retinoic acid bound to RBP reveals that L1 and L5 are further apart[37] and a broader opened mouth conformation. The L1 and L5 conformation accommodates the sterol ring of the ligand while the aliphatic tail sits in the ligand cavity. However, the structure of the apo form of RBP does not have significant global changes in L1 and L5. By comparison, in the structure of the Sphingosine 1-Phosphate bound apolioprotein[38], the L1 and L5 are more similarly oriented to that observed for Can f 6 although the L1 loop is rearranged to a more open conformation. The resulting space is filled with waters that make hydrogen bonds to the sphingosine phosphate head group. Only a fully resolved ligand bound structure will reveal whether or not that is the case.

Table 4. Demographics of human serum samples.

	Dog Dander IgE negative (<0.35kU/L)	Dog Dander IgE positive (>5 kU/L)
Age in years ±(Std Dev)	43.3 (15)	41 (12.6)
Males	1	1
Females	4	4
Skin testing Induration in mm ±(Std Dev)	0	10 (2.3) x 8 (2.4)
Mean sIgE kU/L ±(Std Dev)	<0.35	33.4 (12.1)

https://doi.org/10.1371/journal.pone.0213052.t004

Human IgE binds recombinant Can f 6

The ability of recombinant Can f 6 to bind human sera IgE was analyzed in 5 patients who were previously determined to be AP dog dander positive on immediate skin prick testing as well as dog dander IgE positive (>5 kU/L, median 33.4 kU/L, ImmunoCAP) (Table 4). Age matched controls who were immediate skin prick test negative and IgE negative (<0.35kU/L) were identified. The frequency of qualitative IgE binding to Can f 6 in dog dander positive samples was 60% (3 out of 5 patients, p-value = 0.1199) (Fig 8). IgE reactivity verifies the antigencity of recombinant Can f 6.

Although no data was available regarding symptoms with exposure, molecular component resolved testing may be more accurate in the diagnosis of dog allergy especially in the individual who tested dog dander IgE and skin testing negative but who has evidence of IgE binding to Can f 6.

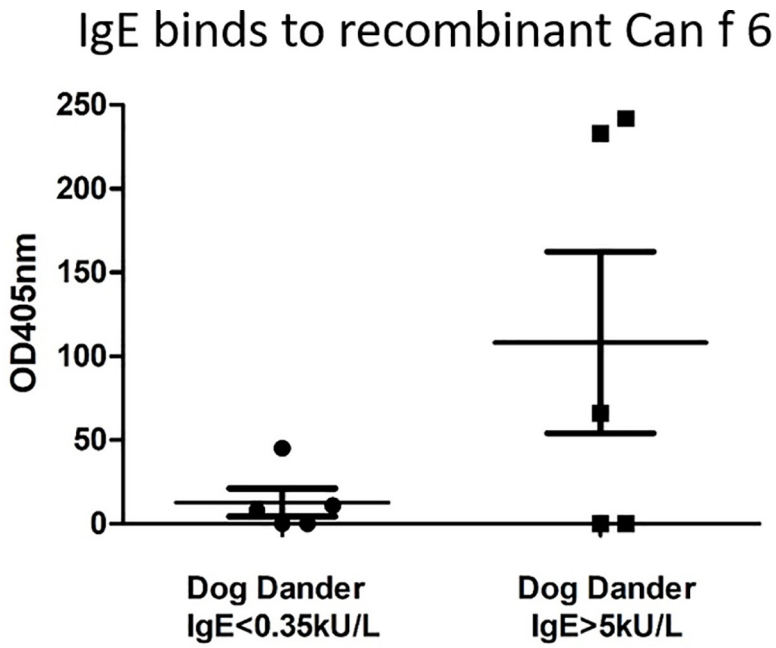


Fig 8. Highly pure recombinant Can f 6 binds human serum IgE in individuals who have been tested for sensitization to unfractionated whole dander IgE. Three of 5 samples from dog dander positive individuals have marked elevation in IgE binding to Can f 6. Mean and standard deviation of OD405nm for patients with dog IgE<0.35: 12.8±8.34 and IgE>5: 108±54.2 for Can f 6. Experiments were completed twice with the trend remaining the same. Overall p-value = 0.1199.

https://doi.org/10.1371/journal.pone.0213052.g008

In summary, the biologically active high resolution x-ray structure of Can f 6 reveals greater insights into the specific roles of critical residues involved in ligand recognition and allergenicity. Further investigation into the role of lipocalins in vivo as well as in clinical allergenic diseases are needed.

Conclusion

We present the crystal structure of the dog lipocalin Can f 6 and provide a detailed analysis of the structure including its binding pocket compared to previously reported select lipocalin allergens and its functional ability to bind human IgE. The structure of recombinant Can f 6 show the classic lipocalin calyx tertiary structure as is typically described for proteins in this family. Interesting, Can f 6 has a primary sequence that is more similar to other mammalian lipocalins than to other dog lipocalins suggesting a role for co-sensitization and cross-reactivity.

Conclusive data regarding clinically relevant cross-reactivity between different animals is lacking. The lipocalin family of proteins constitute the most important group of animal aeroallergens and human sensitization has been identified against lipocalins from dogs, horses, cats, mice, rats, guinea pigs, rabbits, and hamsters[46]. A clinical history of reactions following animal contact and identified IgE sensitization is classically used to identify clinical reactivity. Unfortunately, dog and cat allergens are ubiquitous with exposure that is not limited to direct contact[7] and may be high enough to cause clinical symptoms. Therefore, manifestation of symptoms as a result of exposure to different animals may be difficult to assess above baseline symptoms due to dogs or cats. As more and more allergens are purified, studies can be completed with direct delivery to subjects to illicit reactions thereby identifying conclusively which molecules cause allergic symptoms.

For most, the exact in vivo function of most lipocalins remain unclear although they are classified as odorant and pheromone binding proteins since they are thought to reside within their hydrophobic pocket and they may also function as enzymes or immune related proteins. Highly pure immunologically active proteins will allow in depth molecular analysis and mechanistic studies into defining the role of this protein family in causing allergic disease.

Supporting information

S1 Fig. Omit 2Fo-Fc electron density map of the unmodeled ligand in the Can f 6 binding pocket. The binding surface is predominantly lined by hydrophobic amino acids.
(TIF)

S1 Table. The Can f 6 monomer-monomer surface interfaces within the crystal packing generated in PISA.
(DOCX)

Acknowledgments

We acknowledge and thank the Advanced Photon Source at Argonne National Laboratory for the facilities and staff of the 24-ID-C beamline for their assistance in data-collection. We are grateful to Dr. Jennifer Maynard for the Mopac54 vector plasmid. Samples and data used for this study were downloaded from the National Jewish Health Research Database (https://www.nationaljewish.org/research-science/support/biobanks/bioinfo), supported by National Jewish Health. We thank Dr. Philippa Marrack for her comments during the preparation of this manuscript.

Author Contributions

Conceptualization: Sanny K. Chan.

Data curation: Gina M. Clayton, Janice White, Sanny K. Chan.

Formal analysis: Gina M. Clayton, Janice White, Schuyler Lee, Sanny K. Chan.

Funding acquisition: John W. Kappler, Sanny K. Chan.

Investigation: Gina M. Clayton, Sanny K. Chan.

Methodology: Gina M. Clayton, John W. Kappler, Sanny K. Chan.

Resources: Sanny K. Chan.

Software: Schuyler Lee.

Supervision: Sanny K. Chan.

Validation: Gina M. Clayton, Sanny K. Chan.

Visualization: Sanny K. Chan.

Writing – original draft: Gina M. Clayton, Sanny K. Chan.

Writing – review & editing: Gina M. Clayton, John W. Kappler, Sanny K. Chan.

References

1. Hilger C, Kuehn A, Hentges F. Animal lipocalin allergens. Curr Allergy Asthma Rep. 2012; 12: 438–447. https://doi.org/10.1007/s11882-012-0283-2 PMID: 22791068
2. Jensen-Jarolim E, Pacios LF, Bianchini R, Hofstetter G, Roth-Walter F. Structural similarities of human and mammalian lipocalins, and their function in innate immunity and allergy. Allergy Eur J Allergy Clin Immunol. 2015; 71: 286–294. https://doi.org/10.1111/all.12797 PMID: 26497994
3. Nilsson OB, Van Hage M, Grönlund H. Mammalian-derived respiratory allergens—Implications for diagnosis and therapy of individuals allergic to furry animals. Methods. Elsevier Inc.; 2014; 66: 86–95. https://doi.org/10.1016/j.ymeth.2013.09.002 PMID: 24041755
4. Niemi MH, Rytkönen-Nissinen M, Miettinen I, Jänis J, Virtanen T, Rouvinen J. Dimerization of lipocalin allergens. Sci Rep. Nature Publishing Group; 2015; 5: 13841. https://doi.org/10.1038/srep13841 PMID: 26346541
5. Chan SK, Leung DYM. Dog and Cat Allergies: Current State of Diagnostic Approaches and Challenges. Allergy Asthma Immunol Res. 2018; 10: 97. https://doi.org/10.4168/aair.2018.10.2.97 PMID: 29411550
6. Asher MI, Montefort S, Björkstén B, Lai CK, Strachan DP, Weiland SK, et al. Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multicountry cross-sectional surveys. Lancet. 2006; 368: 733–743. https://doi.org/10.1016/S0140-6736(06)69283-0 PMID: 16935684
7. Konradsen JR, Fujisawa T, Van Hage M, Hedlin G, Hilger C, Kleine-Tebbe J, et al. Allergy to furry animals: New insights, diagnostic approaches, and challenges. J Allergy Clin Immunol. 2015; 135: 616–625. https://doi.org/10.1016/j.jaci.2014.08.026 PMID: 25282018
8. Ukleja-Sokołowska N, Gawrońska-Ukleja E, Żbikowska-Gotz M, Socha E, Lis K, Sokołowski Ł, et al. Analysis of feline and canine allergen components in patients sensitized to pets. Allergy, Asthma Clin Immunol. BioMed Central; 2016; 12: 61. https://doi.org/10.1186/s13223-016-0167-4 PMID: 27956908
9. Arshad SH, Tariq SM, Matthews S, Hakim E. Sensitization to Common Allergens and Its Association With Allergic Disorders at Age 4 Years: A Whole Population Birth Cohort Study. Pediatrics. 2001; 108: e33–e33. https://doi.org/10.1542/peds.108.2.e33 PMID: 11483843
10. Käck U, Asarnoj A, Grönlund H, Borres MP, van Hage M, Lilja G, et al. Molecular allergy diagnostics refine characterization of children sensitized to dog dander. J Allergy Clin Immunol. 2018; https://doi.org/10.1016/j.jaci.2018.05.012 PMID: 29852259
11. Nilsson OB, Binnmyr J, Zoltowska A, Saarne T, van Hage M, Grönlund H. Characterization of the dog lipocalin allergen Can f 6: The role in cross-reactivity with cat and horse. Allergy Eur J Allergy Clin Immunol. 2012; 67: 751–757. https://doi.org/10.1111/j.1398-9995.2012.02826.x PMID: 22515174

12. Wang Y, Li L, Song W, Zhou Y, Cao M. Canis familiaris allergen Can f 6: expression, purification and analysis of B-cell epitopes in Chinese dog allergic children. Oncotarget. 2017; 8: 90796–90807. https://doi.org/10.18632/oncotarget.21822 PMID: 29207604

13. Hilger C, Swiontek K, Arumugam K, Lehners C, Hentges F. Identification of a new major dog allergen highly cross-reactive with Fel d 4 in a population of cat- and dog-sensitized patients. J Allergy Clin Immunol. 2012; 129: 2–6. https://doi.org/10.1016/j.jaci.2011.10.017 PMID: 22104604

14. Saarelainen S, Rytkönen-Nissinen M, Rouvinen J, Taivainen A, Auriola S, Kauppinen A, et al. Animal-derived lipocalin allergens exhibit immunoglobulin E cross-reactivity. Clin Exp Allergy. 2008; 38: 374–381. https://doi.org/10.1111/j.1365-2222.2007.02895.x PMID: 18070162

15. Lascombe MB, Grégoire C, Poncet P, Tavares GA, Rosinski-Chupin I, Rabillon J, et al. Crystal structure of the allergen Equ c 1. A dimeric lipocalin with restricted IgE-reactive epitopes. J Biol Chem. American Society for Biochemistry and Molecular Biology; 2000; 275: 21572–7. https://doi.org/10.1074/jbc.M002854200 PMID: 10787420

16. Bingham RJ, Findlay JBC, Hsieh S-Y, Kalverda AP, Kjellberg A, Perazzolo C, et al. Thermodynamics of Binding of 2-Methoxy-3-isopropylpyrazine and 2-Methoxy-3-isobutylpyrazine to the Major Urinary Protein. American Chemical Society; 2004; https://doi.org/10.1021/JA038461I

17. Yin L, Crawford F, Marrack P, Kappler JW, Dai S. T-cell receptor (TCR) interaction with peptides that mimic nickel offers insight into nickel contact allergy. Proc Natl Acad Sci. 2012; 109: 18517–18522. https://doi.org/10.1073/pnas.1215928109 PMID: 23091041

18. Huynh K and Partch CL. Current Protocols in Protein Science: Analysis of protein stability and ligand interactions by thermal shift assay. Curr Protoc Protein Sci. 2016; 79: 28.9.1–28.9.14. https://doi.org/10.1002/0471140864.ps2809s79.Current

19. Kabsch W. Xds. Acta Crystallogr Sect D Biol Crystallogr. 2010; 66: 125–132. https://doi.org/10.1107/S0907444909047337 PMID: 20124692

20. Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr. International Union of Crystallography; 2013; 69: 1204–14. https://doi.org/10.1107/S0907444913000061 PMID: 23793146

21. Adams PD, Afonine P V., Bunkóczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr Sect D Biol Crystallogr. 2010; 66: 213–221. https://doi.org/10.1107/S0907444909052925 PMID: 20124702

22. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr. International Union of Crystallography; 2007; 40: 658–674. https://doi.org/10.1107/S0021889807021206 PMID: 19461840

23. Niemi MH, Rytkönen-Nissinen MA, Jänis J, Virtanen T, Rouvinen J. Structural aspects of dog allergies: The crystal structure of a dog dander allergen Can f 4. Mol Immunol. Elsevier Ltd; 2014; 61: 7–15. https://doi.org/10.1016/j.molimm.2014.04.003 PMID: 24859823

24. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr Sect D Biol Crystallogr. International Union of Crystallography; 2011; 67: 235–242. https://doi.org/10.1107/S0907444910045749 PMID: 21460441

25. Krissinel E. Stock-based detection of protein oligomeric states in jsPISA. Nucleic Acids Res. 2015; 43: W314–W319. https://doi.org/10.1093/nar/gkv314 PMID: 25908787

26. Heinig M, Frishman D. STRIDE: A web server for secondary structure assignment from known atomic coordinates of proteins. Nucleic Acids Res. 2004; 32: 500–502. https://doi.org/10.1093/nar/gkh429 PMID: 15215436

27. DeLano W. Pymol: An open-source molecular graphics tool. CCP4 Newsl Protein Crystallogr. 2002;700.

28. Tian W, Chen C, Lei X, Zhao J, Liang J. CASTp 3.0: computed atlas of surface topography of proteins. Nucleic Acids Res. 2018; 46: W363–W367. https://doi.org/10.1093/nar/gky473 PMID: 29860391

29. http://xray.bmc.uu.se/usf/voidoo.html.

30. Kleywegt GJ, Jones TA. Detection, delineation, measurement and display of cavities in macromolecular structures. Acta Crystallogr Sect D Biol Crystallogr. 1994; 50: 178–185. https://doi.org/10.1107/S0907444993011333 PMID: 15299456

31. http://www.cbs.dtu.dk/services/NetNGlyc/.

32. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT-BG, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J. 2013; 32: 1478–1488. https://doi.org/10.1038/emboj.2013.79 PMID: 23584533

33. http://www.cbs.dtu.dk/services/NetPhos/.

34. https://web.expasy.org/sulfinator/.

35. Sievers F, Higgins DG. Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences. Methods in molecular biology (Clifton, NJ). 2014. pp. 105–116. https://doi.org/10.1007/978-1-62703-646-7_6

36. Yamamoto K, Ishibashi O, Sugiura K, Ubatani M, Sakaguchi M, Nakatsuji M, et al. Crystal structure of the dog allergen Can f 6 and structure-based implications of its cross-reactivity with the cat allergen Fel d 4. Sci Rep. 2019; 9: 1503. https://doi.org/10.1038/s41598-018-38134-w PMID: 30728436

37. Zanotti G, Marcello M, Malpeli G, Folli C, Sartori G, Berni R. Crystallographic studies on complexes between retinoids and plasma retinol-binding protein. J Biol Chem. 1994; 269: 29613–20. Available: http://www.ncbi.nlm.nih.gov/pubmed/7961949 PMID: 7961949

38. Dahlbäck B, Ahnström J, Christoffersen C, Bo L, Christoffersen C, Nielsen LB. Apolipoprotein M: structure and function. Futur Lipidol. 2008; 3: 495–503.

39. Rouvinen J, Jänis J, Laukkanen M-L, Jylhä S, Niemi M, Päivinen T, et al. Transient dimers of allergens. PLoS One. Public Library of Science; 2010; 5: e9037. https://doi.org/10.1371/journal.pone.0009037 PMID: 20140203

40. Schöll I, Kalkura N, Shedziankova Y, Bergmann A, Verdino P, Knittelfelder R, et al. Dimerization of the major birch pollen allergen Bet v 1 is important for its in vivo IgE-cross-linking potential in mice. J Immunol. 2005; 175: 6645–50. Available: http://www.ncbi.nlm.nih.gov/pubmed/16272319 https://doi.org/10.4049/jimmunol.175.10.6645 PMID: 16272319

41. Madhurantakam C, Nilsson OB, Uchtenhagen H, Konradsen J, Saarne T, Högbom E, et al. Crystal Structure of the Dog Lipocalin Allergen Can f 2: Implications for Cross-reactivity to the Cat Allergen Fel d 4. J Mol Biol. Elsevier Ltd; 2010; 401: 68–83. https://doi.org/10.1016/j.jmb.2010.05.043 PMID: 20621650

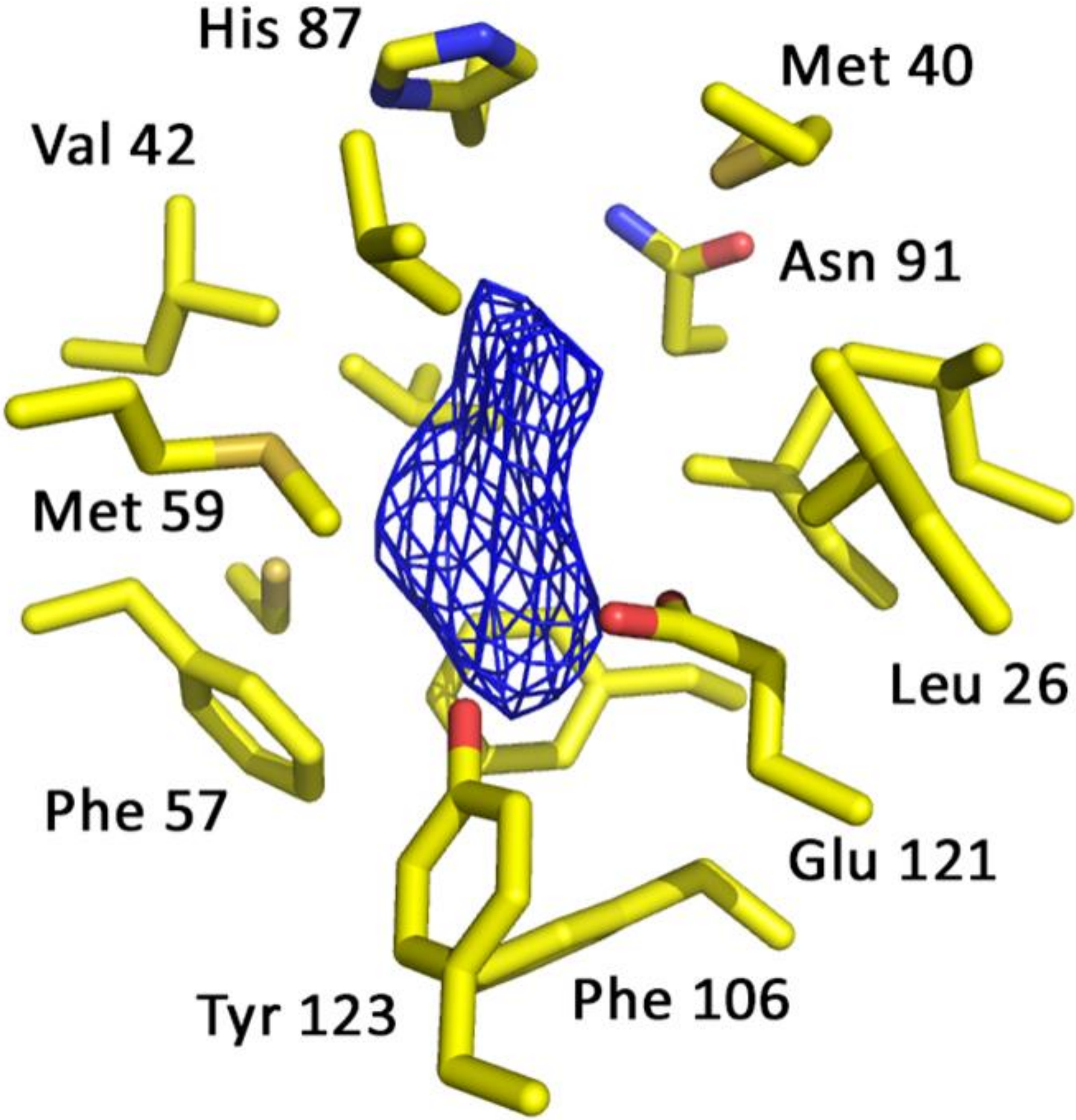
42. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, et al. Protein identification and analysis tools in the ExPASy server. Methods Mol Biol. 1999; 112: 531–52. Available: http://www.ncbi.nlm.nih.gov/pubmed/10027275 https://doi.org/10.1385/1-59259-584-7:531 PMID: 10027275

43. Zhao Z, Liu J, Wasinger VC, Malouf T, Nguyen-Khuong T, Walsh B, et al. Tear lipocalin is the predominant phosphoprotein in human tear fluid. Exp Eye Res. Elsevier Ltd; 2010; 90: 344–349. https://doi.org/10.1016/j.exer.2009.11.013 PMID: 19951704

44. Loebel D, Scaloni A, Paolini S, Fini C, Ferrara L, Breer H, et al. Ligand-Binding of Boar Salivary Lipocalin. Biochem Soc. 2000; 379: 369–379. https://doi.org/10.1016/S0176-6724(86)80137-7

45. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res. Oxford University Press; 2006; 34: W116–8. https://doi.org/10.1093/nar/gkl282 PMID: 16844972

46. http://www.allergen.org/index.php [Internet].


His 87
Met 40
Val 42
Asn 91
Met 59
Leu 26
Phe 57
Glu 121
Tyr 123
Phe 106

S1 Table. The Can f 6 monomer-monomer surface interfaces within the crystal packing generated in PISA.

Surface residues	148 (92%)			
Symmetry Operation	y, x,-z	-y+1, x-y+1, z+1/3	-x, -x+y, -z+1/3	x-y, -y+1, -z-1/3
Interface residues	27 (17%)	17 (10%)	16	4 (3%)
Buried ASA (Å^2)	736 (9.4%)	510	420 (5%)	69 (1%)
Hydrogen bonds	14	2	0	1
Salt bridges	0	0	0	0
Complex Formation Significance Score (CSS)	0.1	0	0	0
Total ASA	7826			