Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Production, crystallization and preliminary X-ray diffraction analysis of the allergen Can f 2 from *Canis familiaris*

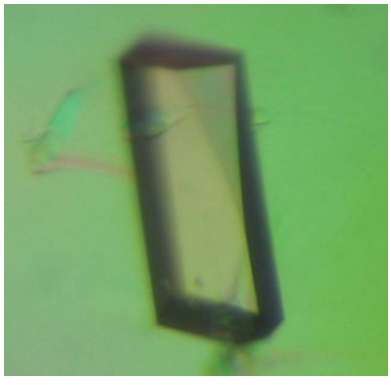
Chaithanya Madhurantakam,[a] Ola B. Nilsson,[b] Klas Jönsson,[b] Hans Grönlund[b]‡ and Adnane Achour[a]*‡

[a]Centre for Infectious Medicine, F59, Department of Medicine Huddinge, Karolinska University Hospital Huddinge, Karolinska Institutet, SE-141 86 Stockholm, Sweden, and [b]Department of Medicine, Clinical Immunology and Allergy Unit, Karolinska Institutet, Karolinska University Hospital Solna, Stockholm, Sweden

‡ These authors should be considered joint last authors.

Correspondence e-mail: adnane.achour@ki.se

Received 3 February 2009
Accepted 24 March 2009



The allergen Can f 2 from dog (*Canis familiaris*) present in saliva, dander and fur is an important cause of allergic sensitization worldwide. Here, the production, isolation, crystallization and preliminary X-ray diffraction analysis of two crystal forms of recombinant Can f 2 are reported. The first crystal form belonged to space group *C*222, with unit-cell parameters $a = 68.7$, $b = 77.3$, $c = 65.1$ Å, and diffracted to 1.55 Å resolution, while the second crystal form belonged to space group *C*2, with unit-cell parameters $a = 75.7$, $b = 48.3$, $c = 68.7$ Å, $\beta = 126.5^{\circ}$, and diffracted to 2.1 Å resolution. Preliminary data analysis indicated the presence of a single molecule in the asymmetric unit for both crystal forms.

1. Introduction

Exposure to indoor allergens may lead to allergic sensitization as well as the development of bronchial hyper-reactivity and asthma (Munir *et al.*, 1993). Cats and dogs, the two most popular home-environment pets, are the sources of potent allergens. Allergy to the domestic dog (*Canis familiaris*) and cat (*Felis domesticus*) is prominent worldwide. We have previously determined crystal structures of the main cat allergen Fel d 1, demonstrating its striking structural resemblance to human uteroglobin (Kaiser, Grönlund, Sandalova, Ljunggren, Achour *et al.*, 2003; Kaiser, Grönlund, Sandalova, Ljunggren, van Hage *et al.*, 2003; Kaiser *et al.*, 2005, 2007). Although crude preparations of dog and cat allergens have shown a certain level of immunological cross-reactivity in which the hair- and dander-specific IgEs of dog-allergic subjects cross-reacted with allergenic cat fur (Boutin *et al.*, 1988), the sequence identity between cat and dog allergens is negligible (7%). *Canis familiaris* allergens 1 and 2 (Can f 1 and Can f 2, respectively) are major allergens found in dog dander (Saarelainen *et al.*, 2004; de Groot *et al.*, 1991; Schou *et al.*, 1991; Ford *et al.*, 1989) and saliva (de Groot *et al.*, 1991) and they share 24.5% sequence identity (Kamata *et al.*, 2007). Both dog allergens belong to the lipocalin family of extracellular proteins that facilitate the binding and transport of small lipophilic molecules (Konieczny *et al.*, 1997). Can f 2 is produced from the tongue and parotid gland of *C. familiaris*. Trichosurin, the milk whey protein from the metatherian *Trichosurus vulpecula* (common brushtail possum; Watson *et al.*, 2007), and α_{2u}-globulin (A2U), the major urinary protein from adult rat (Böcskei *et al.*, 1992), display sequence identities of 34% and 30%, respectively, to Can f 2.

Here, we report the cloning, expression, purification, crystallization and preliminary diffraction data collection of recombinant Can f 2 allergen (rCan f 2). The rCan f 2 protein is composed of 170 amino acids including an *Xho*I restriction site and a His tag (a signal peptide of 19 amino-acid residues in length has been removed) with a calculated molecular mass of 18.9 kDa. The two crystals obtained using the sitting-drop vapour-diffusion method belonged to space groups *C*222 and *C*2 and diffracted to 1.55 and 2.1 Å resolution, respectively. Determination of the three-dimensional structure of rCan f 2 will provide insight into its function and facilitate the search for potential structural resemblances to other protein families.

Furthermore, the structure of rCan f 2 will be used as a template for the mapping of B-cell and T-cell epitopes.

2. Materials and methods

2.1. Cloning and expression of rCan f 2

A synthetic gene encoding rCan f 2 was produced by PCR using overlapping oligonucleotides (Table 1) and DNAWorks (http://molbio.info.nih.gov/dnaworks). The synthetic gene was assembled by PCR and the resulting fragment was ligated into the expression vector pET20b (Novagen, EMD Chemicals Inc., Darmstadt, Germany) using the *Nde*I and *Xho*I restriction sites. The expressed protein has a 6×His tag located at the C-terminus (Fig. 1*a*). The rCan f 2 protein sequence ends with amino-acid residue Asp162, which is followed by the residues LE encoded by the *Xho*I restriction site that link rCan f 2 and the His tag. The expression plasmid was transformed into *Escherichia coli* BL21(DE3)pLysS (Novagen) and plated onto a

Table 1
List of primers.

rCan f 2 primer	Sequence
1	GGAATTCCATATGGAGGGTAACCAT
2	CCGCTTAACTCTTCCAGACCACCTTGTGGCTCTTCATGGTTACCCTCCATATGGAATTCC
3	TGGTCTGGAAGAGTTAAGCGGCCGTTGGCACTCTGTCGCCTTAGCATCTAACAAAAGCGA
4	GATAAAAACACGAAAATGACCCCATGGCTTGATTAAGTCGCTTTTGTTAGATGCTAAGGC
5	CATGGGGTCATTTTCGTGTTTTTATCCACAGCATGAGCGCCAAGGATGGCAACTTACATG
6	CTTTTCGCACTGGCCATCTTGCGGAATTAAGATATCGCCATGTAAGTTGCCATCCTTGGC
7	CAAGATGGCCAGTGCGAAAAGGTTTCTCTGACCGCCTTCAAAACGGCCACCAGCAACAAG
8	CCGCCAGGTATAAGTCATTGTGGCCCCAGTACTCCAGGTCGAACTTGTTGCTGGTGGCCG
9	CACAATGACTTATACCTGGCGGAGGTTGATCCGAAAAGCTATCTGATCCTGTATATGATT
10	CAACCAGAGACGTATCGTCGTTGTATTGGTTAATCATATACAGGATCAGATAGCTTTTCG
11	ACGACGATACGTCTCTGGTTGCGCACTTAATGGTCCGCGATCTGAGCCGCCAGCAGGATT
12	GTGCAGACCAATGTCTTCACACACAGACTCAAATGCTGGCAGAAAATCCTGCTGGCGGCT
13	TGTGAAGACATTGGTCTGCACAAAGATCAAATTGTGGTGCTGAGCGACGACGACCGCTGT
14	CGCGCCGCTCGAGGTCACGGCTACCTTGACAGCGGTCGTCGTCG
5′ extension primer	GGAATTCCATATGGAGGGAAACCATGAGGAGCCC
3′ extension primer	CGCGCCGCTCGAGGTCTCTGGAACCCTGGCAGCGATC

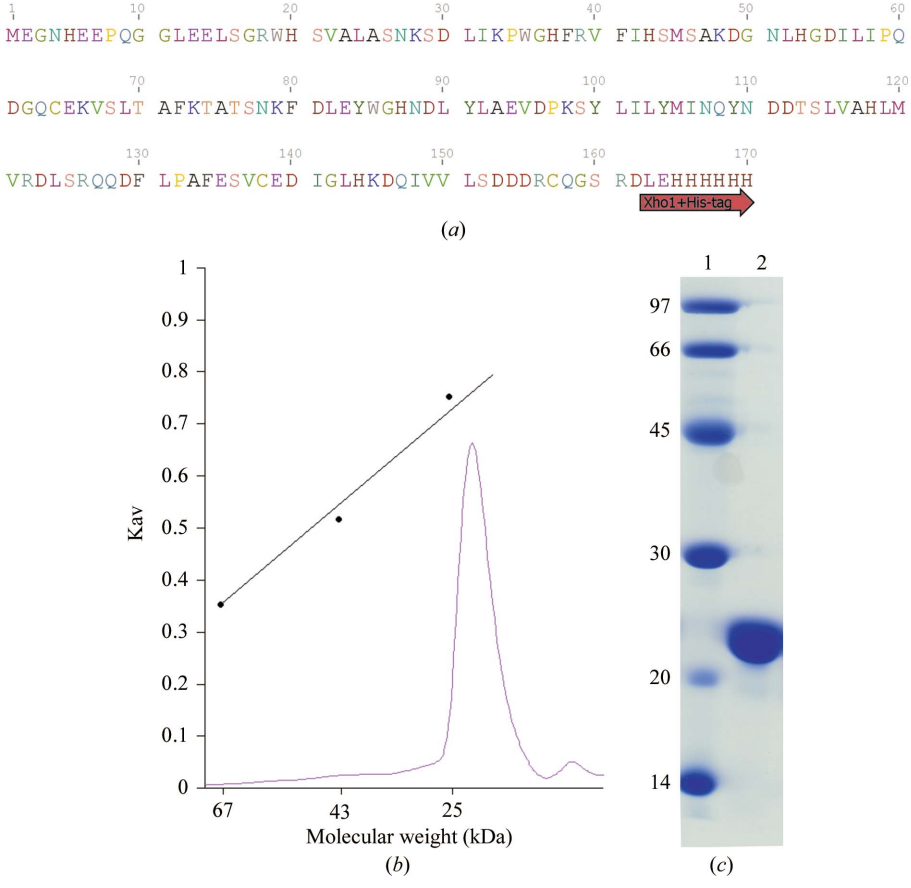


Figure 1
(*a*) Primary structure of His-tagged rCan f 2 fusion molecule with the signal peptide removed. The rCan f 2 protein sequence ends with amino-acid residue Asp162 and the residues LE encoded by the *Xho*I restriction site, linking rCan f 2 and the His tag. The *Xho*I restriction site and the His tag derived from the vector are indicated by an arrow. (*b*) Size-exclusion chromatography of rCan f 2. The markers correspond to the following: 67 kDa, bovine serum albumin; 43 kDa, ovalbumin; 25 kDa, chymotrypsinogen A. (*c*) Protein separation by 12% SDS–PAGE under reducing conditions stained with Coomassie Brilliant Blue. Lanes 1 and 2 contain molecular-weight markers (kDa) and 1 µg purified rCan f 2, respectively.

Table 2
Data-collection statistics for rCan f 2 crystal forms *A* and *B*.

Values in parentheses are for the highest resolution shell.

Crystal form	*A*	*B*
Wavelength (Å)	0.93	0.93
Resolution (Å)	33.94–1.55 (1.63–1.55)	38.55–2.10 (2.21–2.10)
Space group	*C*222	*C*2
Unit-cell parameters (Å, °)	a = 68.7, b = 77.3, c = 65.1	a = 75.7, b = 48.3, c = 68.7, β = 126.5
V_M (Å^3 Da^{-1})	2.54	2.66
Solvent content (%)	51.6	53.7
No. of molecules in ASU	1	1
No. of observed reflections	88720 (12794)	42636 (6229)
No. of unique reflections	24608 (3497)	11765 (1704)
Redundancy	3.6 (3.7)	3.6 (3.7)
Completeness (%)	99.6 (99.1)	100.0 (100.0)
R_{merge}† (%)	4.6 (15.0)	10.1 (20.2)
$\langle I/\sigma(I)\rangle$	17.2 (7.6)	10.9 (5.5)
Mosaicity (°)	0.72	0.65

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith observation of reflection hkl and $\langle I(hkl)\rangle$ is the weighted average intensity for all observations i of reflection hkl.

Luria–Bertani broth (LB) agar plate containing 100 mg l^{-1} ampicillin (amp) and 35 mg l^{-1} chloramphenicol (chl) overnight at 310 K. A single picked colony was cultured in 15 ml LB medium with amp and chl overnight at 310 K and inoculated in 2 l LB medium containing amp and chl. When the cell concentration reached log phase (OD of 0.6 at 600 nm), protein production was induced by the addition of 0.4 m*M* IPTG (Sigma, St Louis, Missouri, USA) and the simultaneous addition of 100 mg l^{-1} amp. After 3 h incubation, the cultures were centrifuged at 15 000*g* for 10 min (Beckman Coulter). Cell pellets were resuspended in 30 ml PBS and stored frozen at 253 K.

2.2. Purification of recombinant Can f 2

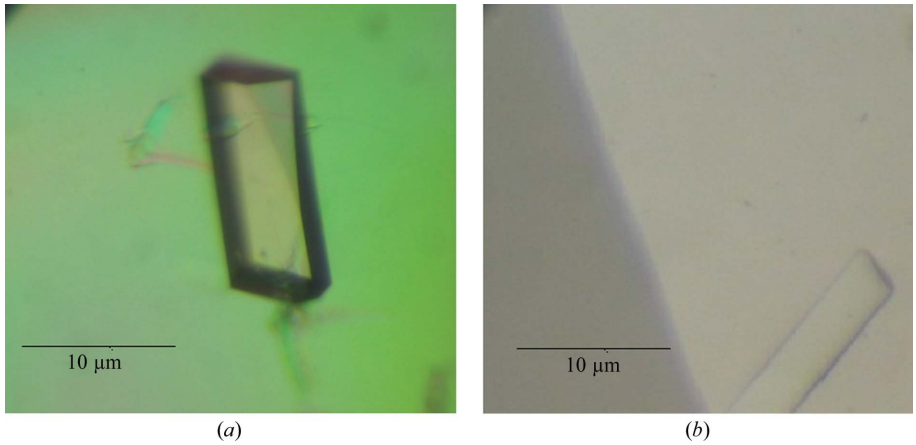
Frozen *E. coli* cells were disrupted by 10 × 15 s ultrasonication bursts on ice (Soniprep 150 ultrasonic disintegrator, Sanyo Gallenkamp) and centrifuged for 10 min at 15 000*g*. The supernatant was filtered (0.2 µm) and run on a 5 ml Ni^{2+}-HiTrap affinity column (GE Healthcare, Sweden). The column was washed using ten column volumes of 20 m*M* Tris–HCl, 0.5 *M* NaCl, 20 m*M* imidazole pH 7.4. His-tagged rCan f 2 was eluted using 20 m*M* Tris–HCl, 0.5 *M* NaCl, 0.5 *M* imidazole pH 7.4. The eluted sample was subjected to size-exclusion chromatography (16/60 Superdex 200 prep-grade column, GE Healthcare) equilibrated with PBS. Monomeric rCan f 2 was isolated after size-exclusion chromatography (Fig. 1*b*), pooled and diluted in a 1:1 ratio with 20 m*M* Tris–HCl pH 8.7 prior to anion-exchange chromatography. The proteins were loaded onto a 10 ml Q-Sepharose HP column equilibrated with 20 m*M* Tris–HCl pH 8.7. Elution was achieved with a ten column-volume linear gradient using 20 m*M* Tris–HCl, 1 *M* NaCl pH 8.7, giving a NaCl concentration of 250 m*M*. Eluted protein was desalted on a PD-10 column (GE Healthcare) to 20 m*M* Tris–HCl, adjusted to pH 7.4, aliquoted and stored at 193 K. The protein concentration was 4.6 mg ml^{-1} as determined using the BCA protein assay (Pierce). The purity was estimated by SDS–PAGE under reducing conditions (Fig. 1*c*).

2.3. Crystallization of rCan f 2

Preliminary crystallization conditions were screened using sparse-matrix screens (Hampton Research, California, USA), PACT and JCSG+ (Molecular Dimensions Ltd, Suffolk, UK) in 96-well Corning plates (Corning Incorporated, New York, USA) at 277 and 293 K. rCan f 2 (4.6 mg ml^{-1}) protein-screening assays were set up by the sitting-drop vapour-diffusion method using a Phoenix crystallization robot (Art Robbins Instruments). Nanodrops containing 100 nl protein solution and 100 nl crystallization solution were equilibrated against reservoir containing 50 µl crystallization solution using the sitting-drop vapour-diffusion method. Fine screening of initial conditions were performed by equilibrating 2 µl protein solution and 2 µl crystallization solution against 1 ml crystallization solution using the hanging-drop vapour-diffusion method.

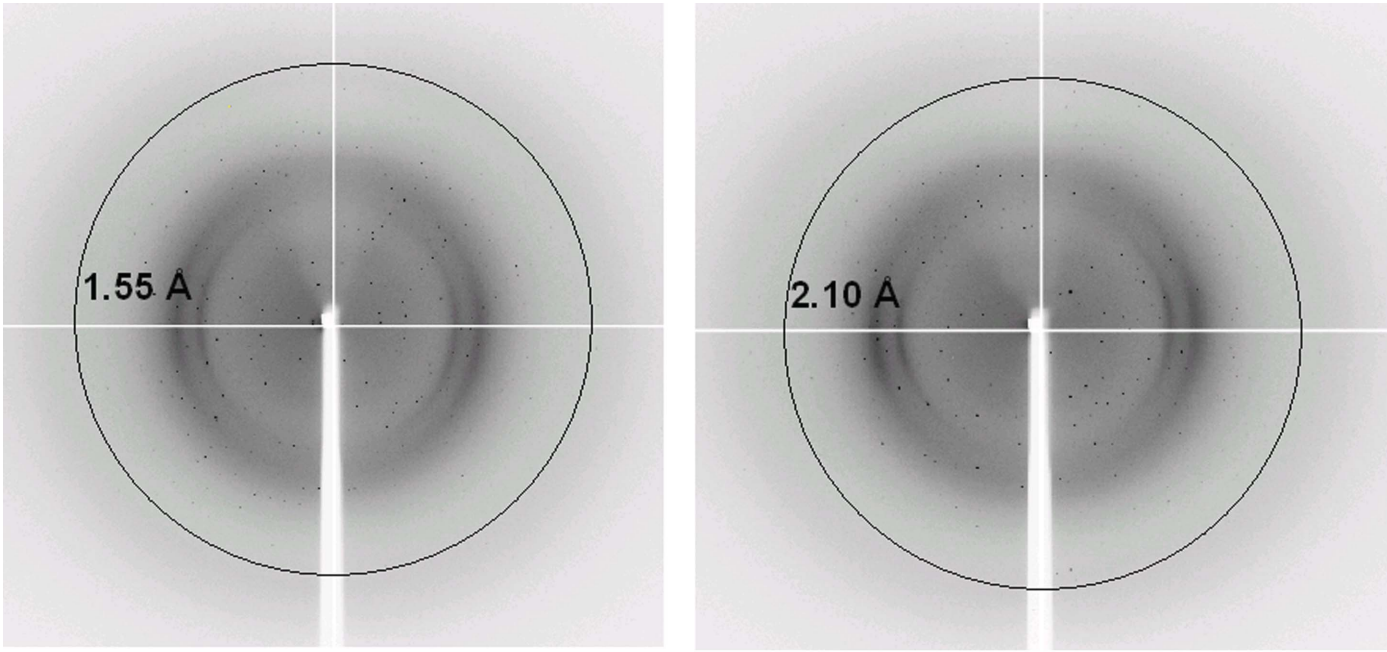
2.4. Data collection and processing

Crystals were soaked in cryoprotectant (80% crystallization solution and 20% glycerol) before flash-freezing in a cold nitrogen stream. Data sets were collected on beamline ID14-2 at the synchrotron-radiation facility at the ESRF (Grenoble, France) to resolutions of 1.55 and 2.1 Å for crystal forms *A* and *B*, respectively. A monochromatic X-ray beam of fixed wavelength 0.93 Å and an ADSC Q4 CCD detector were used for data collection. A total of 540 images were collected with 0.5° oscillation per frame for crystal form *A*, of which 360 images were processed, and a total of 360 images were collected with 0.5° oscillation per frame for crystal form *B*. Space group and unit-cell parameters were determined using the auto-indexing option in the program *MOSFLM* (Leslie, 1992). Scaling and reduction of the data sets were performed using the program *SCALA* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).



(*a*) (*b*)

Figure 2
rCan f 2 crystals were grown using the sitting-drop method. (*a*) Crystal form *A* (10 × 4 µm) was obtained from 50%(*w*/*v*) Tacsimate pH 6.0 at 277 K. (*b*) Crystal form *B* (10 × 3 µm) was obtained from 0.2 *M* magnesium chloride hexahydrate, 0.1 *M* Tris–HCl pH 8.0 and 35%(*w*/*v*) PEG 4000 at 293 K.



(*a*) (*b*)

Figure 3
X-ray diffraction patterns from rCan f 2 crystal forms *A* (*a*) and *B* (*b*).

3. Results

The recombinant form of the Can f 2 allergen protein is composed of 170 amino acids with a calculated molecular mass of 18 935 Da and a theoretical pI of 4.94 (Fig. 1*a*). The protein was produced in *E. coli* and isolated using a combination of Ni^{2+}-HiTrap affinity-column, gel-filtration (Fig. 1*b*) and ion-exchange chromatography. A minimum of 70 mg highly purified monomeric rCan f 2 was obtained from 2 l bacterial culture (Fig. 1*c*).

Two crystal forms (*A* and *B*) were observed that were thin plate-like and thin rod-shaped with dimensions of 10 × 4 and 10 × 3 µm, respectively (Fig. 2). Crystal form *A* appeared after 20 d at 293 K in 60%(*w*/*v*) Tacsimate pH 6.0. Crystal form *B* appeared after a month at 293 K in 0.2 *M* magnesium chloride hexahydrate, 0.1 *M* Tris–HCl pH 8.5 and 30%(*w*/*v*) PEG 4000. Further fine screening was performed using hanging-drop vapour diffusion to improve the size and the quality of the crystals; the final crystals of forms *A* and *B* were obtained in 50%(*w*/*v*) Tacsimate pH 6.0 at 277 K and 0.2 *M* magnesium chloride hexahydrate, 0.1 *M* Tris–HCl pH 8.0 and 35%(*w*/*v*) PEG 4000 at 293 K, respectively. Both forms were cryoprotected with 20% glycerol followed by vitrification in liquid nitrogen.

Crystal forms *A* and *B* diffracted to 1.55 and 2.1 Å resolution, respectively (Fig. 3). The diffraction data statistics for rCan f 2 crystal forms *A* and *B* are given in Table 2. Crystal form *A* is orthorhombic, with unit-cell parameters a = 68.7, b = 77.3, c = 65.1 Å, while crystal form *B* is monoclinic, with unit-cell parameters a = 75.7, b = 48.3, c = 68.7 Å, β = 126.5°. The average *B*-factor values derived from Wilson plots are 13.6 and 20.9 $Å^2$, respectively (French & Wilson, 1978). Matthews coefficient analysis (Matthews, 1968) for the two crystal forms indicated that both forms are most likely to contain a single molecule in the asymmetric unit (Matthews coefficient of V_M = 2.54 and 2.66 $Å^3$ Da^{-1} corresponding to 51.6 and 53.7% solvent content for crystal forms *A* and *B*, respectively).

rCan f 2 shows a sequence identity of 34% with the milk whey protein trichosurin from the metatherian *T. vulpecula* and hence the structure of this protein (PDB code 2r73; Watson *et al.*, 2007) is considered to be a suitable search model for determination of the structure of rCan f 2 by molecular replacement.

We acknowledge ESRF beamline ID14-2 (Grenoble, France), beamline I711 at MAX II (Lund, Sweden), beamline MX 14-2 at BESSY (Berlin, Germany) for access to synchrotron-radiation facilities and the Structural Genomics Consortium Stockholm (SGC), Department of Medical Biochemistry and Biophysics at the Karolinska Institute for access to the Phoenix robot automated crystallization facility. This work was supported by grants from the Swedish Cancer and Allergy Foundation, the Swedish Asthma and Allergy Association's Research Foundation, Hesselman's, Magnus Bergvall's and Konsul Th. C. Bergh's Research Foundations and the Karolinska Institutet.

References

Böcskei, Z., Groom, C. R., Flower, D. R., Wright, C. E., Phillips, S. E. V., Cavaggioni, A., Findlay, J. B. C. & North, A. C. T. (1992). *Nature (London)*, **360**, 186–188.
Boutin, Y., Hébert, H., Vrancken, E. R. & Mourad, W. (1988). *Clin. Allergy*, **18**, 287–293.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Ford, A. W., Alterman, L. & Kemeny, D. M. (1989). *Clin. Exp. Allergy*, **19**, 183–190.
French, S. & Wilson, K. (1978). *Acta Cryst.* A**34**, 517–525.
Groot, H. de, Goei, K. G. H., van Swieten, P. & Aalberse, R. C. (1991). *J. Allergy Clin. Immunol.* **87**, 1056–1065.
Kaiser, L., Grönlund, H., Sandalova, T., Ljunggren, H. G., Achour, A., Schneider, G. & van Hage, M. (2003). *Int. Arch. Allergy Immunol.* **132**, 25–26.
Kaiser, L., Grönlund, H., Sandalova, T., Ljunggren, H. G., van Hage, M., Achour, A. & Schneider, G. (2003). *J. Biol. Chem.* **278**, 37730–37735.
Kaiser, L., Grönlund, H., van Hage-Hamsten, M. & Achour, A. (2005). *Acta Cryst.* F**61**, 232–234.

Kaiser, L., Velickovic, T. C., Badia-Martinez, D., Adedoyin, J., Thunberg, S., Hallén, D., Berndt, K., Grönlund, H., Gafvelin, G., van Hage, M. & Achour, A. (2007). *J. Mol. Biol.* **370**, 714–727.
Kamata, Y., Miyanomae, A., Nakayama, E., Miyanomae, T., Tajima, T. & Hoshi, H. (2007). *Int. Arch. Allergy Immunol.* **142**, 291–300.
Konieczny, A., Morgenstern, J. P., Bizinkauskas, C. B., Lilley, C. H., Brauer, A. W. & Bond, J. F. (1997). *Immunology*, **92**, 577–586.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Munir, A. K. M., Einarsson, R., Schou, C. & Dreborg, S. K. G. (1993). *J. Allergy Clin. Immunol.* **91**, 1067-1074.
Saarelainen, S., Taivainen, A., Rytkönen-Nissinen, M., Auriola, S., Immonen, A., Mäntyjärvi, R., Rautiainen, J., Kinnunen, T. & Virtanen, T. (2004). *Clin. Exp. Allergy*, **34**, 1576–1582.
Schou, C., Svendsen, U. G. & Lowenstein, H. (1991). *Clin. Exp. Allergy*, **21**, 321–328.
Watson, R. P., Demmer, J., Baker, E. N. & Arcus, V. L. (2007). *Biochem. J.* **408**, 29–38.