The EMBO Journal Vol. 21 No. 19 pp. 5007–5016, 2002

The cross-reactive calcium-binding pollen allergen, Phl p 7, reveals a novel dimer assembly

Petra Verdino, Kerstin Westritschnig[1], Rudolf Valenta[1] and Walter Keller[2]

Institute of Chemistry, Structural Biology Group, Karl-Franzens-University Graz, Heinrichstrasse 28, A-8010 Graz and [1]Department of Pathophysiology, Molecular Immunopathology Group, General Hospital, University of Vienna, Waehringerguertel 18–20, A-1090 Vienna, Austria

[2]Corresponding author
e-mail: walter.keller@uni-graz.at

The timothy grass pollen allergen Phl p 7 assembles most of the IgE epitopes of a novel family of 2 EF-hand calcium-binding proteins and therefore represents a diagnostic marker allergen and vaccine candidate for immunotherapy. Here we report the first three-dimensional structure of a representative of the 2 EF-hand allergen family, Phl p 7, in the calcium-bound form. The protein occurs as a novel dimer assembly with unique features: in contrast to well known EF-hand proteins such as calmodulin, parvalbumin or the S100 proteins, Phl p 7 adopts an extended conformation. Two protein monomers assemble in a head-to-tail arrangement with domain-swapped EF-hand pairing. The intertwined dimer adopts a barrel-like structure with an extended hydrophobic cavity providing a ligand-binding site. Calcium binding acts as a conformational switch between an open and a closed dimeric form of Phl p 7. These findings are interesting in the context of lipid- and calcium-dependent pollen tube growth. Furthermore, the structure of Phl p 7 allows for the rational development of vaccine strategies for treatment of sensitized allergic patients.

Keywords: calcium-binding protein/crystal structure/domain-swapping/EF-hand allergen/vaccine design

Introduction

IgE-mediated allergies represent a major health problem in the industrialized world as they affect almost 25% of the population (Kay, 2001a,b; Wills-Karp *et al*., 2001). The immediate symptoms of the disease (e.g. allergic rhinoconjunctivitis, dermatitis, bronchial asthma and anaphylactic shock) are caused by the cross-linking of effector cell-bound IgE antibodies by allergens, which leads to the release of biological mediators such as histamine or leukotrienes. In order to induce strong effector cell activation and thus inflammatory responses, an allergen must be able to cross-link effector cell-bound IgE antibodies efficiently. This process requires the presence of at least two IgE epitopes on the allergen surface (Segal *et al*., 1977). IgE antibodies of allergic patients may recognize either 'continuous epitopes' consisting of a row of consecutive amino acids or 'discontinuous epitopes', which are composed of amino acids from different portions of the allergen brought into proximity by the molecule fold (Valenta and Kraft, 2001). For the precise analysis of the surface-exposed IgE epitopes, the three-dimensional structures of allergens are the most important source of information (Valenta *et al*., 1998a). The detailed knowledge derived from the three-dimensional structures together with immunological data allows for the rational development of strategies to convert allergen molecules into hypoallergenic derivatives through destruction or reorientation of IgE epitopes. Such allergen derivatives may act as potent vaccines for immunotherapy as they exhibit a reduced risk of inducing anaphylactic side effects, but still preserve their immunogenic features (Valenta *et al*., 1999; Valenta, 2002).

Recently, members of a novel family of 2 EF-hand calcium-binding proteins have been identified as potent allergens originating from pollens of a variety of plant species such as trees, grasses and weeds (Valenta *et al*., 1998b). This particular allergen family is of interest for several reasons: its members are expressed in pollens but were not detected in other plant tissues (Engel *et al*., 1997; Hayek *et al*., 1998; Niederberger *et al*., 1999). Due to the strong IgE cross-reactivity within the 2 EF-hand allergen family, allergic patients are polysensitized to pollens of various plants (Batanero *et al*., 1996; Engel *et al*., 1997; Smith *et al*., 1997; Suphioglu *et al*., 1997; Twardosz *et al*., 1997; Hayek *et al*., 1998; Valenta *et al*., 1998b; Niederberger *et al*., 1999; Okada *et al*., 1999; Tinghino *et al*., 2002). IgE antibodies of patients allergic to 2 EF-hand allergens recognize preferentially the calcium-bound form of the protein, whereas the calcium-depleted apo form exhibits a strongly reduced IgE-binding capacity (Valenta *et al*., 1998b).

The fact that 2 EF-hand allergens share structural and immunological features suggests the use of representative members of this protein family for diagnosis and immunotherapy (Tinghino *et al*., 2002). The timothy grass pollen protein Phl p 7 (Niederberger *et al*., 1999) contains most of the relevant IgE epitopes of this allergen family and is thus the most promising candidate for vaccine design (Tinghino *et al*., 2002). In this study, we have determined the three-dimensional structure of Phl p 7 by X-ray crystallography at 1.75 Å resolution. Our experiments reveal an unexpected intertwined dimer assembly yielding novel implications for the biological function of 2 EF-hand allergens and allowing for a rational design of hypoallergenic vaccines for immunotherapy.

Results and discussion

Overall structure of Phl p 7

The structure reveals a functional dimer with interchain EF-hand pairing (Figure 1A and B). Each monomer (chain A and B, respectively) consists of two calcium-binding

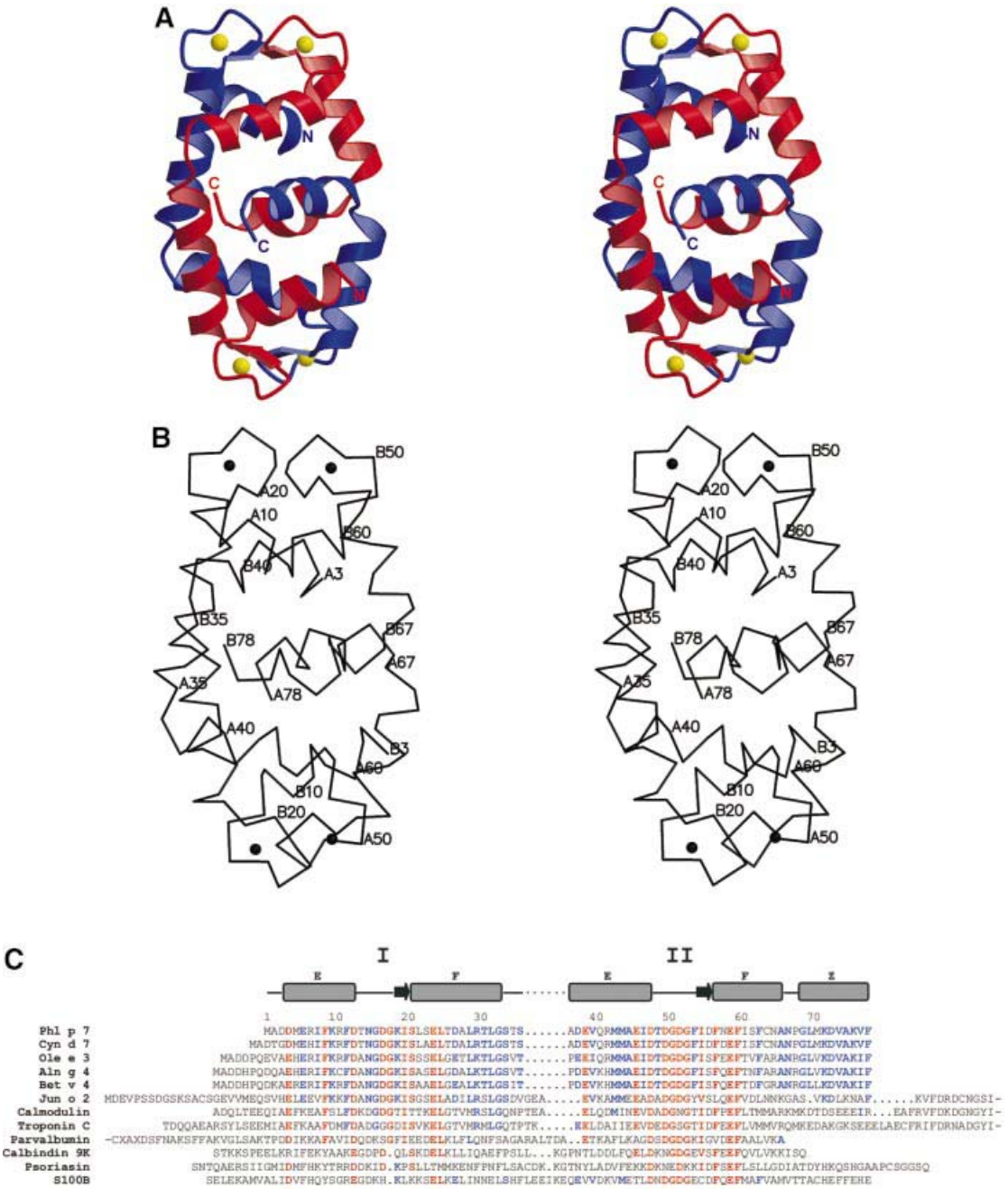

Fig. 1. (**A**) Stereo ribbon representation of the intertwined Phl p 7 dimer. Chain A is coloured in blue, chain B is shown in red and the calcium ions are highlighted in yellow. The N-terminal EF-hand of chain A (AI) and the C-terminal EF-hand of chain B (BII) comprise the upper EF-hand pair, whereas the C-terminal EF-hand of chain A (AII) and the N-terminal EF-hand of chain B (BI) form the lower one. The C-terminal Z-helices comprise an equatorial belt. Figures were prepared with MOLSCRIPT (Kraulis, 1991) and rendered using Raster3D (Merritt and Murphy, 1994). (**B**) Stereo C_α trace of the Phl p 7 dimer, labelled at every tenth amino acid. (**C**) Sequence alignment of five representative 2 EF-hand pollen allergens (from the top: timothy grass, bermuda grass, olive, alder, birch), the 4 EF-hand pollen allergen Jun o 2 (cypress) and various representative EF-hand proteins. The secondary structure plot is derived from the three-dimensional structure of Phl p 7. The helices of the N- and C-terminal EF-hands (EFI and EFII) and the Z-helix are represented by grey tubes and the short β-strands by black arrows. Gaps are symbolized as dots; hyphens indicate that only parts of the whole sequences are shown. More than 75% homology (conserved exchanges: RK, DE, NQ, TS) within all proteins is shown in red, and residues which exhibit 100% homology within the 2 EF-hand pollen allergens are pictured in blue.

helix–loop–helix domains (the N- and C-terminal EF-hands, Figure 1C) connected via a three amino acid hinge region. In Phl p 7 this linker is much shorter than in classical EF-hand proteins (where the hinge region comprises 6–14 residues), enforcing an extended conformation of the two EF-hand domains. This novel arrangement is unexpected as it stands in contrast to the globular shape of domains with sequential EF-hand pairs as observed throughout the large EF-hand protein family. The superimposition of a Phl p 7 monomer with representative EF-hand proteins demonstrates the difference between the extended and the globular EF-hand conformations (Figure 2A): the N- and C-terminal EF-hands of Phl p 7 are arranged spatially distant, whereas the

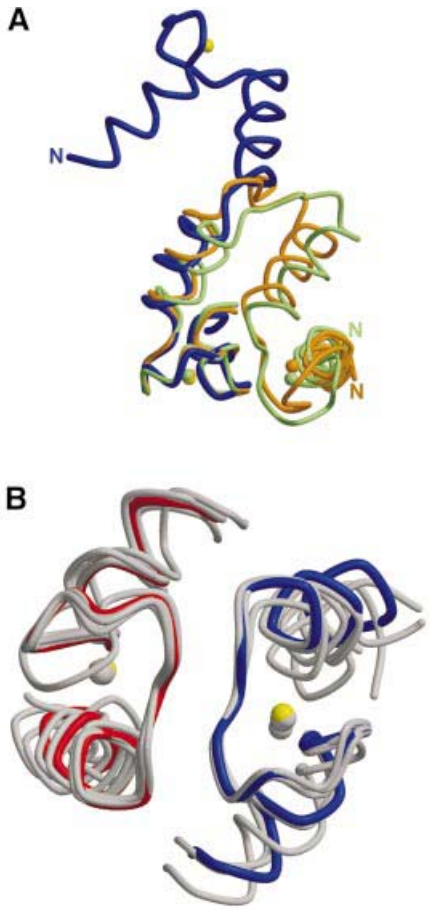

Fig. 2. (**A**) Superimposition of a Phl p 7 monomer with two representative calcium-binding EF-hand proteins. The backbone of Phl p 7 chain A (except the Z-helix) is shown in blue; its calcium ions are pictured in yellow. The calcium-binding loop of the C-terminal EF-hand AII was superimposed with the corresponding loops of the C-terminal EF-hand of a calmodulin domain (1OSA; light green) and S100B (1MHO; orange), respectively. Whereas the two EF-hands of Phl p 7 adopt an extended conformation, the EF-hands of representative EF-hand proteins are arranged in close spatial proximity. In contrast to the rather short three amino acid hinge loop between the calcium-binding sites of Phl p 7, a much longer loop (6–14 residues) allows for the sequential EF-hand pairing in well-known EF-hand proteins. (**B**) Superimposition of the pairing calcium-binding loops of Phl p 7 and representative EF-hand proteins. The chains of Phl p 7 are shown in blue (AII) and red (BI), respectively, and the calcium ions are represented in yellow. The sequentially pairing EF-hand loops of calmodulin (1OSA), troponin C (1TOP), parvalbumin (1PVB) and S100B (1MHO) are represented in grey. Although in Phl p 7 the pairing occurs between the two monomer chains, the spatial arrangement of the loops is identical to the intrachain paired loops of standard EF-hand proteins.

intradomain calcium-binding sites of the EF-hand representatives calmodulin and S100B are in close proximity to each other.

C-terminal to the second EF-hand in Phl p 7 there is an 11 amino acid distorted helix (Z-helix). Due to a sharp kink at residues 67/68, this helix protrudes from the C-terminal EF-hand motif with an angle of ~100° (Figure 1A). This is in contrast to structures of other EF-hand proteins in which the region adjacent to the C-terminal F-helix is either a continuous helical connection to the next domain (e.g. calmodulin) or missing altogether (e.g. calbindin D_{9k}). The Phl p 7 monomers assemble as an intimate dimer in a head-to-tail arrangement stabilized via helix–helix interactions between EF-hand motifs across the entire dimer interface. Furthermore, each of the C-terminal Z-helices intertwines with both E-helices of the opposite monomer. The F-helices of both the N- and C-terminal EF-hands come into close contact with their non-crystallographic symmetry mate where they approach the molecular 2-fold axis close to the hinge regions on one side (left in Figure 1A) and at the kink regions on the opposite side (right in Figure 1A).

The overall structure of Phl p 7 is barrel shaped, with approximate dimensions of 35 × 45 × 35 Å. The calcium-coordinating loops form the top and the bottom of the barrel, the E- and F-helices accommodate the upper and lower side walls, whereas the C-terminal Z-helices constitute an equatorial belt. The extended conformation and the intertwined head-to-tail arrangement of the protein monomers lead to the formation of an extended cavity within the interior of the dimer. This cavity is lined with hydrophobic side chains and provides an enclosed ligand-binding site.

Dimer formation through a novel way of EF-hand pairing

All four calcium-binding sites are fully occupied and well ordered within the Phl p 7 dimer. The atomic *B*-factors of the calcium ions are lower than the average value observed for the main chain atoms (Table I). The calcium coordination in the N- and C-terminal binding sites follows the 'canonical' EF-hand motif (Kretsinger and Nockolds, 1973; Lewit-Bentley and Rety, 2000): the calcium ion is enclosed by a 12 amino acid loop and coordinated by seven oxygen atoms of asparagines or aspartic acids, a peptide carbonyl, a water molecule and a bidentate glutamic acid at the vertices of a pentagonal bipyramid. The mean distance between the calcium ions and the oxygen atoms in Phl p 7 is 2.38 Å (2.23–2.60 Å) (Figure 3A and B).

In most calcium-binding proteins, EF-hands occur in sequential pairs (Lewit-Bentley and Rety, 2000) which are arranged in spatial vicinity and connected via a flexible hinge loop of 6–14 amino acids (Figure 2A). The EF-hand pairs are stabilized by the formation of a hydrophobic core between the E- and F-helices and by a short antiparallel β-sheet between the calcium-binding loops. The so-called 'rule of EF-hand pairing' (Lewit-Bentley and Rety, 2000) is also fulfilled for Phl p 7. However, in contrast to intramolecular pairing of sequential EF-hands within one protein chain, intermolecular pairing occurs between the C- and N-terminal domains of opposite chains of the Phl p 7 dimer (Figure 1A and B). Although this assembly comprises a completely different domain arrangement of the single monomer chains, it is most remarkable how well the interchain paired calcium-binding loops of Phl p 7 superimpose with the intrachain paired loops of representative EF-hand proteins such as calmodulin, troponin C, parvalbumin or S100B (r.m.s.d. values in the range of 0.3–0.6 Å) (Figure 2B). The interactions between the EF-hands AI/BII and AII/BI in Phl p 7 are nearly identical and comparable with the well established interactions between sequential intramolecular EF-hand pairs (Babu *et al.*, 1988; Matsumura *et al.*, 1998): in addition to a hydrophobic core formed between the E- and F-helices, two highly conserved key residues provide directed

Table I. Data collection and refinement statistics

Data collection	Native	EMP
Wavelength (Å)	0.8482	0.9950
Resolution range (Å)	14–1.75	25–1.75
Measured reflections	122 317	143 392
Unique reflections	13 072	12 985
Completeness (%)	100.0 (100.0)[a]	99.5 (98.8)[a]
$\langle I/\sigma(I)\rangle$ (final shell)	7.3	3.7
R_{sym} (%)[b]	4.7 (18.5)[a]	3.4 (18.3)[a]
Refinement		
Resolution	14–1.75	
R_{cryst} (%)[c]	19.84	
R_{free} (%)[d]	22.52	
R.m.s.d. bonds (Å)	0.005	
R.m.s.d. angles (°)	0.97	
$\langle B\rangle_{main\ chain\ atoms}$ (Å^2)[e]	19.1	
$\langle B\rangle_{side\ chain\ atoms}$ (Å^2)[e]	23.1	
$\langle B\rangle_{calcium\ atoms}$ (Å^2)[e]	15.3	
$\langle B\rangle_{solvent\ atoms}$ (Å^2)[e]	33.6	

[a]Values in parentheses refer to the highest resolution shell.
[b]$R_{sym} = \sum|I-\langle I\rangle| / \sum I$, where I is the intensity of an individual measurement and $\langle I\rangle$ is the corresponding mean value.
[c]$R_{cryst} = \sum||F_o|-|F_c|| / \sum|F_o|$, where $|F_o|$ is the observed and $|F_c|$ the calculated structure factor amplitude.
[d]R_{free} is the same as R_{cryst} calculated with a randomly selected test set of 5.0% of all reflections that was never used in the refinement calculations.
[e]$\langle B\rangle$ is the average atomic B-factor.

interactions between the calcium-binding loops of the monomers. I20 and I55 of the opposite protein chains form a short antiparallel β-sheet. Hydrophobic interactions occur between the side chains of the latter residues as well as between the side chains of K19 and F54, whose backbone carbonyl oxygens additionally coordinate the calcium ions (Figure 3C).

Dimer interactions in Phl p 7 are optimized, furthermore, by adjusting the interhelical angles within the EF-hand motifs: the N-terminal EF-hands exhibit angles of 117° and 110°, which are well within the reported range for calcium-bound EF-hand motifs (Matsumura *et al.*, 1998). In contrast, the value of 82° found in the C-terminal domains differs significantly. This adjustment seems to be necessary for the formation of the nearly globular overall structure of the dimer, especially for the interdigitating arrangement of the Z-helices with the E-helices of the opposite monomer. The Z-helices, whose sequences are highly conserved throughout the 2 EF-hand pollen allergens (Figure 1C), are unusually hydrophobic. A distorted hydrogen bonding pattern within the Z-helices and a 3_{10}-helix turn at the very C-terminus of the protein allow for placement of all hydrophobic side chains either at the dimer interface or in the hydrophobic interior.

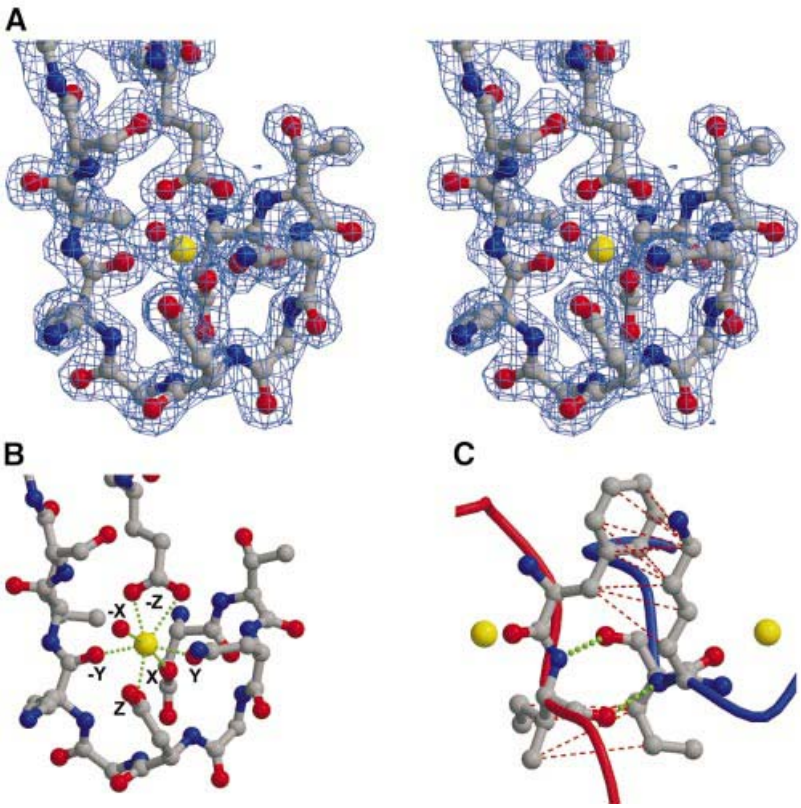

Fig. 3. (**A**) Stereo diagram of the $(3F_o-2F_c)$ electron density map at 1.75 Å resolution contoured at 1σ around the 12 amino acid calcium-binding loop of AI (residues 13–24) with the calcium ion (yellow sphere) positioned in the middle. (**B**) The 7-fold coordination of the calcium ion by oxygen atoms (red) of asparagines or aspartic acids, a peptide carbonyl oxygen, a water molecule and a bidentate glutamic acid in the positions X Y Z –Y –X –Z according to the nomenclature of Kretsinger (Kretsinger and Nockolds, 1973; Lewit-Bentley and Rety, 2000). (**C**) The loop pairing interactions between AI (blue) and BII (red): two hydrogen bonds (green dotted lines) are formed between the peptide planes of I20 (chain A) and I55 (chain B). Furthermore, this short β-sheet assembly is stabilized by hydrophobic interactions (red dotted lines) between the side chains of I20 and I55 as well as K19 and F54. The latter residues are also involved in the calcium coordination via their backbone carbonyl oxygen atoms [in position –Y in (B)].

A novel family of EF-hand proteins comprising domain-swapped dimers

Phl p 7 is a representative of a new class of EF-hand proteins exploiting domain swapping (Bennett *et al*., 1995; Liu and Eisenberg, 2002) for dimer formation. This exchange of (sub)domains between different protein chains was suggested as an evolutionary mechanism for oligomer assembly (Bennett *et al*., 1995; Heringa and Taylor, 1997). Furthermore, domain-swapped dimers which differ in function from the monomers were reported, consistent with the hypothesis that domain swapping can be a means of regulating activity (Bennett *et al*., 1995; Schymkowitz *et al*., 2000; Liu and Eisenberg, 2002). No example of a naturally occurring domain-swapped EF-hand protein has been reported yet; however, in the case of the monomeric 2 EF-hand protein calbindin D_{9k}, an artificial domain-swapped assembly was enforced by a mutation in the hinge region (Hakansson *et al*., 2001).

Since domain swapping can be initiated by extreme environmental conditions or by mutations in several proteins (Liu and Eisenberg, 2002), we applied size exclusion chromatography experiments to confirm that the Phl p 7 dimerization may also occur under physiological conditions. At neutral (pH 7.1) as well as at acidic (pH 4.7) conditions, Phl p 7 eluted as a dimer with an apparent mol. wt of 14 kDa (Figure 4A). Three further structure-based findings support the assumption that the Phl p 7 protein is a naturally occurring domain-swapped dimer. (i) In Phl p 7, the hinge region between the two EF-hand domains comprises only three amino acids (residues 34–36) whereas in intradomain paired EF-hands of various representative proteins this loop comprises between six and 14 residues (calculated with PROCHECK; Laskowski *et al*., 1993). Artificial shortening of hinge loops was shown to force monomers of staphylococcal nuclease (Green *et al*., 1995) to form domain-swapped dimers. Furthermore, the domain-swapped interleukin-5 dimer (Milburn *et al*., 1993) was suggested to have evolved from a monomeric form by loop deletion (Bennett *et al*., 1995). (ii) The hydrogen bonding network between side chains and the polypeptide backbone in the hinge region stabilizes the extended conformation of the Phl p 7 monomers. The involved residues, R30, S34, S36 and E39, are conserved throughout all the members of the 2 EF-hand pollen allergen group (Figure 1C). (iii) The domain-swapped arrangement enables the Phl p 7 dimer to assume an energetically very favourable globular shape in which the mainly hydrophobic Z-helices interdigitate with both amphiphilic E-helices of the opposite monomer, stabilized by numerous hydrophobic interactions. Within a compact monomer with sequential EF-hand pairing, these dimer-specific interactions would be impossible.

The hydrophobic cavity forms a potential ligand-binding site

The characteristic assembly of Phl p 7 leads to the formation of an extended hydrophobic cavity occupying a volume of 800 $Å^3$ (calculated with VOIDOO; Kleywegt and Jones, 1994) in the centre of the protein dimer (Figure 5A). This cavity is lined by hydrophobic side chains, and the exchange of solvent or ligands appears impossible for the assembled dimer as it is closed to a spherical probe of 1.4 Å radius. The cavity adopts an

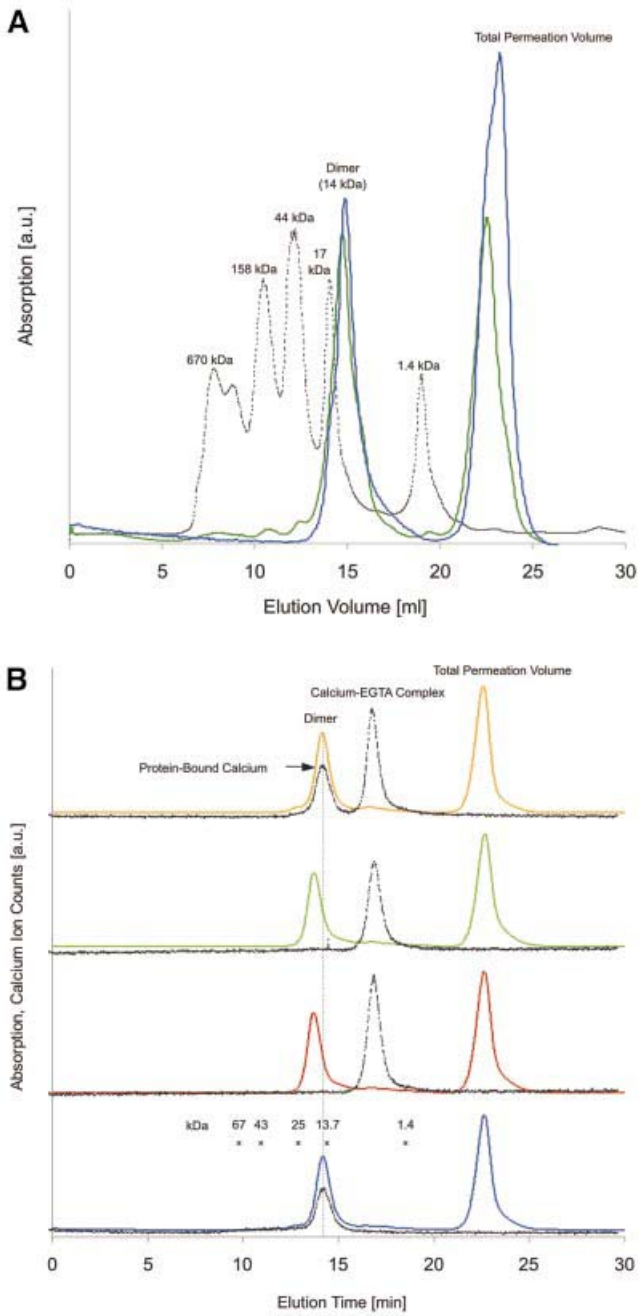

Fig. 4. (**A**) pH-dependent size exclusion chromatography. The protein standard is shown as a black dotted line. Phl p 7 assembles as a dimer with an apparent mol. wt of 14 kDa at both pH 7.1 (green) and pH 4.7 (blue). (**B**) Size exclusion chromatography coupled to ICP-MS (inductively coupled plasma mass spectrometry). The UV absorption traces at 210 nm (coloured lines) and the corresponding online calcium ion count mass spectrometry traces (black lines) are shown. The native Phl p 7 protein and the heat-treated sample after re-annealing in the presence of calcium yield identical elution profiles; the coincident traces are coloured in blue. The heat-treated protein refolded in the presence of EGTA is pictured in red. Room temperature incubation of Phl p 7 in the presence of EGTA is shown in green. The yellow trace represents the protein after heat denaturation and refolding in the presence of EGTA followed by subsequent addition of excessive $CaCl_2$. Phl p 7 remains dimeric under all applied conditions, even in the completely calcium-free form. Calcium depletion results in a shorter elution time, corresponding to an increase of the hydrodynamic radius of the dimer. Addition of calcium to the EGTA-treated protein leads to re-incorporation of the metal ion and an identical elution time as for the untreated calcium-bound protein.

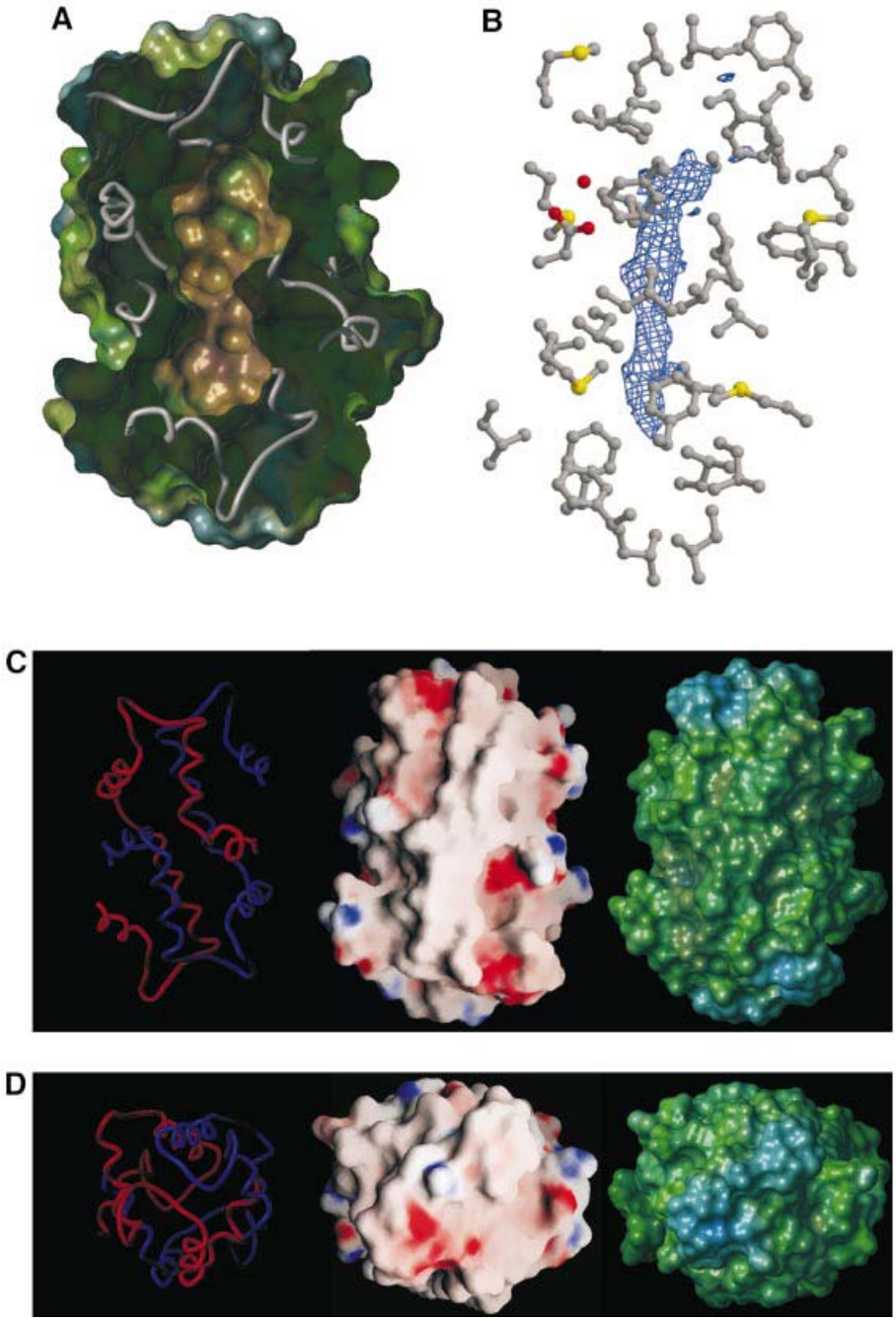

Fig. 5. (**A**) Intersection through the Phl p 7 dimer. The polypeptide backbone is represented in light grey. Surfaces were calculated and coloured according to their hydrophobicity, from brown (hydrophobic) through green to blue (hydrophilic) with SYBYL® 6.7.1 (Tripos Inc.). The protein's outer surface is hydrophilic (especially in the calcium-binding regions), whereas the cavity is predominantly hydrophobic and completely sealed by hydrophobic side chains. (**B**) Zooming into the hydrophobic cavity: the $(3F_o - 2F_c)$ electron density contoured at 1σ of the ligand (blue map) surrounded by the lining residues. Eight isoleucine, eight leucine, two alanine and one valine residue comprise the lateral walls of the cavity, whereas clusters of phenylalanine residues form the top and the bottom. The only polar side chain within the hydrophobic assembly, an aspartic acid coordinating a solvent molecule, is located in the more globular upper part of the cavity. (**C** and **D**) Surface plots of key features of the Phl p 7 dimer assembly: the ridge evolving at the 'kink' region and the calcium-coordinating site AI/BII. The left pictures represent the polypeptide backbones of chain A (blue) and chain B (red). The electrostatic potential distribution coloured from red (negative) through white to blue (positive) with GRASP (Nicholls *et al*., 1991) is shown in the middle. The right pictures show the surface hydrophobicity that is colour-coded analogously to (A).

asymmetric shape with a more globular upper part and a narrower, elongated lower part. Within the cavity, we observe the electron density of a potential hydrophobic ligand (Figure 5B). Due to the mainly hydrophobic environment within the cavity and the fact that the cavity adopts an asymmetric shape, we built models of fatty acids with chain lengths between 12 and 16 carbon atoms into the elongated electron density. All these models fit well within the cavity without major clashes. However, the exact nature of the ligand could not be determined due to the limited quality of the electron density. This might be due to the fact that the ligand exhibits rotational and

conformational flexibility, or may be due to heterogeneity. It also should be considered that we used the recombinant protein expressed in a heterologous expression system for crystallization and therefore various hydrophobic ligands other than the specific one might be bound within the cavity.

It is well established that pollen germination and pollen tube growth critically depend on calcium (Brewbaker and Kwack, 1963; Pierson *et al*., 1996; Franklin-Tong, 1999). Furthermore, it was shown that lipids are required for pollen tube guidance (Wolters-Arts *et al*., 1998; Franklin-Tong, 1999). The exact biological functions of the pollen-specific 2 EF-hand proteins are unknown to date (Valenta *et al*., 1998b), but due to their calcium-binding ability they were suggested to be involved in the control of calcium metabolism in pollen germination and pollen tube growth (Batanero *et al*., 1996; Engel *et al*., 1997; Smith *et al*., 1997; Suphioglu *et al*., 1997; Niederberger *et al*., 1999; Okada *et al*., 1999). The hydrophobic cavity revealed by the three-dimensiuonal structure of Phl p 7 suggests a ligand-binding function rather than a mere calcium-buffering function, corresponding well to the observation that Phl p 7 and its homologous proteins are expressed tissue specifically in pollens (Valenta *et al*., 1998b) and that they occur in very low amounts (<1% of the total pollen protein content; Niederberger *et al*., 1999; K.Westritschnig and R.Valenta, unpublished data).

A calcium-dependent conformational switch

The EF-hand motif is well known to undergo conformational changes upon calcium binding (Strynadka and James, 1991). This feature is exploited functionally by a multitude of EF-hand proteins (e.g. calmodulin, aequorin or S100 proteins) for signal transmission and calcium-dependent target binding (Yap *et al*., 1999). As the cavity in the Phl p 7 dimer is inaccessible to solvent exchange, the uptake or the release of the ligand must be coupled to transient opening or dimer dissociation.

In order to investigate the influence of calcium on the structural assembly, we tested the oligomerization and calcium-binding status of Phl p 7. The protein was subjected to thermal denaturation and refolding under calcium supplementation and calcium depletion conditions and compared with the corresponding untreated samples. In the presence of calcium, Phl p 7 regained its native fold as shown by CD spectroscopy (Niederberger *et al*., 1999; P.Verdino and W.Keller, unpublished data), and in size exclusion experiments it eluted as a calcium-bound dimer with an apparent mol. wt of 14 kDa identical to the untreated protein (blue trace in Figure 4B). This result clearly demonstrates that the dimer is the thermodynamically preferred oligomerization state of Phl p 7 in the calcium-bound form. Performing the same experiment under calcium depletion conditions, Phl p 7 still elutes as a dimer (red trace in Figure 4B) but its fold is reduced as shown by CD spectroscopy. The depleted dimer is completely free of calcium and exhibits an increased hydrodynamic radius observed as a significantly shorter elution time. Complete calcium depletion is also achieved by incubation with chelating agents at room temperature (green trace in Figure 4B). Upon addition of calcium (yellow trace in Figure 4B), the metal ion is re-incorporated into the protein and the elution profile is identical to that of the native calcium-bound dimer.

Our results show that Phl p 7 exists in two distinct conformational states that can be converted reversibly by the addition or depletion of calcium. As manifested in the crystal structure, the calcium-bound dimer exhibits a closed protein shell impenetrable to the ligand. We thus envisage that a calcium-dependent conformational change of the EF-hands, corresponding to the observed change of the hydrodynamic radius, would provide sufficient flexibility for ligand uptake or release.

Towards a structure-based design of allergy vaccines

IgE inhibition experiments have demonstrated that calcium-binding pollen allergens of the 2 EF-hand family are highly cross-reactive (Engel *et al*., 1997; Smith *et al*., 1997; Suphioglu *et al*., 1997; Twardosz *et al*., 1997; Hayek *et al*., 1998; Valenta *et al*., 1998b; Niederberger *et al*., 1999). They were also shown to share epitopes with members of the 3 EF- as well as the 4 EF-hand pollen allergen families (Tinghino *et al*., 2002), but as yet no cross-reactivity was observed with EF-hand allergens from other sources [e.g. human or spinach calmodulin (Smith *et al*., 1997) or the major fish allergen, parvalbumin (I.Swoboda and R.Valenta, unpublished data)]. These distinct allergenic properties may be explained on the basis of the dimeric Phl p 7 structure and its unique surface features that are predicted to be very similar for the highly conserved members of the 2 EF-hand pollen allergens but quite distinct for EF-hand proteins from other sources. One of these features is a completely uncharged ridge extending over the whole length of the dimer (Figure 5C), which is composed of the C-terminal F-helices and the kink regions of both monomers. The calcium-binding loops at the top and the bottom of the dimer represent two very hydrophilic poles and exhibit a remarkable asymmetry: the C-terminal loop is negatively charged whereas the N-terminal loop is uncharged except for a conserved lysine (Figure 5D).

The finding that Phl p 7 occurs as a domain-swapped dimer may explain its high allergenic activity. In fact, we have noted previously that Phl p 7 is extremely potent regarding the induction of effector cell activation as studied by basophil histamine release experiments and skin testing (Niederberger *et al*., 1999). It is thus possible that the domain-swapped dimeric form may expose more IgE-binding sites as compared with a monomeric form of the molecule and thus is more potent regarding the cross-linking of effector cell-bound IgE antibodies. Since Phl p 7 contains most of the IgE epitopes present in other calcium-binding pollen allergens (Tinghino *et al*., 2002), its three-dimensional structure will be the basis for the development of hypoallergenic vaccines for the treatment of allergies to most of the EF-hand pollen allergens. Such derivatives with a reduced risk of inducing anaphylactic side effects may be obtained by strategies that disrupt the characteristic three-dimensional structure and thus the conformational IgE epitopes. We propose several structure-based strategies for the design of such hypoallergenic Phl p 7 derivatives. For the first two strategies, we provide experimental evidence that they can indeed strongly reduce or even abolish the allergenic activity of Phl p 7

(see Supplementary figure 6 available at *The EMBO Journal* Online).

(i) It is well documented that the IgE-binding capacity of the 2 EF-hand pollen allergens is strongly modulated by the presence or absence of protein-bound calcium (Valenta *et al*., 1998b). In most cases, the IgE binding was greatly reduced or even completely lost for the calcium-depleted proteins in the presence of chelating agents (Engel *et al*., 1997; Suphioglu *et al*., 1997; Twardosz *et al*., 1997; Hayek *et al*., 1998; Niederberger *et al*., 1999). IgE binding could be regained by addition of calcium (Smith *et al*., 1997). The most obvious strategy to obtain hypoallergenic derivatives of Phl p 7 is to mutate those amino acids that are involved in calcium coordination (e.g. the aspartates, asparagines and the highly conserved glutamate as shown in Figure 3B). In this context, we produced a Phl p 7 derivative containing two mutations in each of the two calcium-binding loops. This mutant exhibits strongly reduced or abolished IgE-binding capacity when tested with sera from Phl p 7 allergic patients. The protein lacks its calcium-binding ability as shown by size exclusion chromatography coupled to ion-specific mass spectrometry experiments, and occurs mainly as a dimer but also partly as a tetrameric and monomeric form (data not shown).

(ii) Another possibility for the generation of hypoallergenic Phl p 7 derivatives is the production of allergen fragments containing the isolated N- or C-terminal calcium-binding domain. We have constructed two such synthetic Phl p 7 fragments that indeed have completely lost their IgE-binding capacity (see Supplementary figure 6).

Other structure-based strategies which may lead to the generation of hypoallergenic Phl p 7 derivatives include the following.

(iii) Based on the three-dimensional structure of Phl p 7, it may be considered to target surface-exposed amino acids which potentially are involved in epitope formation (rather than disrupting discontinuous epitopes by introducing conformational changes). Promising candidates are the highly conserved residues K19 and F54 as well as the amino acids within the uncharged ridge.

(iv) Another promising approach may be the disruption of the dimer assembly. This strategy targets the high cross-linking activity of Phl p 7 on the effector cell-bound IgE antibodies. As a result of the domain swapping, the protein assembles into a highly symmetric dimer. This leads to the doubling of (identical) IgE epitopes. A conversion of the domain-swapped dimer into the corresponding monomer would probably maintain the IgE epitopes, but would strongly diminish the allergy eliciting cross-linking activity. We suggest mutations of the hinge loop that elongate the linker region between the N- and C-terminal calcium-binding domains. This would enable the formation of a monomeric Phl p 7 with intramolecular EF-hand pairing.

(v) Finally, it may be considered to disrupt the dimeric structure by mutations of the hydrophobic residues that provide interaction between the E-helices and the Z-helix of the opposite monomer in the domain-swapped dimer.

In conclusion, we have revealed the three-dimensional structure of a highly cross-reactive pollen allergen, Phl p 7, and thus provide the basis for the rational design of hypoallergenic vaccines (Valenta, 2002). Such vaccines may be used for the treatment of sensitized allergic patients and, perhaps, for prophylactic allergy vaccination.

Materials and methods

Purification of recombinant Phl p 7

Recombinant Phl p 7 was prepared as described (Niederberger *et al*., 1999) with the modification that no boiling step was applied for the enrichment of the protein. The final dialysate was lyophilized, re-suspended in double-distilled water, and stored at –20°C. Mass spectrometry revealed that the N-terminal methionine residue was processed in the heterologous expression system.

Crystallization

Crystals were grown at 20°C by sitting drop vapour diffusion. Drops consisted of 3.0 µl of 5.0 mg/ml Phl p 7 protein in double-distilled water mixed with 0.7 µl of the 0.5 ml well liquor (0.2 M ammonium sulfate, 24–28% PEG 4000, 0.1 M alanine/HCl, adjusted to pH 3.2–3.4). Crystals took 2–5 days to appear and grew to a size of 0.3 × 0.2 × 0.04 mm within 10 days. The crystals grew as rhombic plates out of a common nucleation centre and had to be broken out of the clusters using a micro-scalpel. An ethyl-mercury-phosphate (EMP) heavy atom derivative was prepared by applying a very small portion of solid EMP directly to a drop containing well-shaped crystals followed by incubation for 4 days.

Data collection

Crystals were cryo-protected via short soaks in a solution consisting of 10% PEG 200 and 90% crystallization well liquor, and were flash frozen in liquid nitrogen. Native data were collected at the EMBL X11 beamline at the DORIS storage ring, DESY, Hamburg. Heavy atom derivative data were collected at the X-ray diffraction beamline XRD1 at the ELETTRA Synchrotron Light Laboratory, Trieste, at a wavelength close to the mercury LIII absorption edge (Table I). All data were processed and reduced using DENZO and SCALEPACK (Otwinowski, 1993). The space group of the crystals was determined to be $P2_1$. Unit cell dimensions were $a = 31.08$ Å, $b = 56.21$ Å, $c = 37.24$ Å and $\beta = 90.23°$ for the native crystal, and $a = 31.05$ Å, $b = 56.44$ Å, $c = 37.27$ Å and $\beta = 90.27°$ for the mercury derivative, respectively.

Phasing

Initially, we tried to solve the structure by molecular replacement using EF-hand pairs, but also lone EF-hands of various calcium-binding proteins as search models. No clear solutions could be obtained with the programs AMORE (Navaza, 1994) and MOLREP (Vagin and Teplyakov, 1997) [both part of the CCP4 suite (CCP4, 1994)]. However, the best solutions were subjected to rigid-body refinement, but the resulting electron density maps were not interpretable. Thus we employed heavy atom methods for phasing. Initial anomalous Patterson maps calculated from the mercury single-wavelength anomalous dispersion (SAD) data set showed two heavy atom sites in the asymmetric unit, consistent with two molecules each bearing a heavy atom site. Phases were determined with the program SOLVE (Terwilliger and Berendzen, 1999) and the data subjected to density modification using the program RESOLVE (Terwilliger, 2000). Initial electron density maps indicated α-helices and molecular boundaries. Using the structure factors of the native data (excluding a random set comprising 5% of the reflections) and the phases obtained from RESOLVE, 131 out of 154 residues (connectivity index of 0.94) could be traced automatically by the program ARP/wARP (Perrakis *et al*., 1999) using the warpNtrace mode. Further tracing and model building was done manually with the program O (Jones *et al*., 1991). Four calcium ions could be identified due to their significantly smaller *B*-factors as compared with the mean values for solvent atoms.

Refinement

The initial model of the dimer without any solvent molecules was refined against the native data (resolution 14.0–1.75 Å) using CNS (Brünger *et al*., 1998). Refinement including simulated annealing starting at 4000 K resulted in an initial $R_{cryst} = 29.68\%$ and $R_{free} = 32.66\%$. Further cycles of rebuilding (especially in the terminal regions) and refinement, as well as water picking and the incorporation of three sulfate molecules, decreased the R_{cryst} to 21.57% (R_{free} = 23.78%). Simulated annealing omit maps were used as criteria for the chain tracing in the N- and C-termini, but the density corresponding to Ala2 remained of poor quality. An elongated electron density portion not belonging to the protein monomers was clearly visible in the middle of the dimer interface. During the refinement

process, the quality of the density did not increase well enough to permit the exact identification of the ligand. The final model (R_{cryst} = 19.84%; R_{free} = 22.52%) included two Phl p 7 monomers (residues 3–78), four calcium ions, three sulfate molecules and 157 solvent molecules. Residues 5, 20, 30, 42, 44, 62 and 71 of monomer A and residues 5, 10, 19, 23 and 62 of monomer B were built in alternate conformations. The Ramachandran plot calculated with PROCHECK (Laskowski *et al*., 1993) showed 93.3% of all residues in the core regions and no residues in the disallowed regions.

pH-dependent size exclusion chromatography

Volumes of 10 µl of 5 mg/ml Phl p 7 in double-distilled water were incubated for several hours at 20°C with an equal volume of 20 mM potassium phosphate, 150 mM NaCl pH 7.1 or 20 mM sodium acetate, 150 mM NaCl pH 4.7. The samples were applied to a Superose 12 HR10/30 column (Amersham Biosciences) equilibrated with the corresponding buffer and eluted isocratically at a flow rate of 0.5 ml/min. Due to the lack of any tryptophan and tyrosine residues in the Phl p 7 protein, detection was conducted at 210 nm. The BIO-RAD Gel Filtration Standard was used for calibration under identical conditions at pH 7.1. The apparent mol. wt of 14 kDa for Phl p 7 was determined from the linear regression of log molecular weight versus elution volume. A Phl p 7 mutant with two mutations in each calcium-binding domain assembles in various oligomerization states and eluted as a tetramer (apparent mol. wt of 26 kDa), dimer (14 kDa) and monomer (8 kDa) (data not shown).

Size exclusion chromatography coupled to ICP-MS

Volumes of 5 µl of 5 mg/ml Phl p 7 in double-distilled water were incubated overnight with 15 µl of 20 mM Tris–HCl, 50 mM NH_4Cl pH 7.0 or with the same buffer containing 2 mM EGTA. Identical samples were prepared, boiled for 5 min at 100°C in a water bath and slowly re-folded. A further sample consisting of 5 µl of 5 mg/ml Phl p 7 and 13 µl buffer with 2 mM EGTA was subjected to heat denaturation and annealing. After reaching room temperature, $CaCl_2$ was added to a final concentration of 4 mM and incubated for several hours. Aliquots of 5 µl of the protein samples were applied to a Superdex 75 PC 3.2/30 column (Amersham Biosciences) equilibrated with the corresponding buffer with or without 2 mM EGTA, and eluted isocratically at a flow rate of 0.1 ml/min. Detection was performed at 210 nm and the column was calibrated using the Amersham Pharmacia Low Molecular Weight Gel Filtration Calibration Kit. The Agilent 1100 HPLC system was coupled to an Agilent 7500c ICP-MS (inductively coupled plasma mass spectrometer) where the following masses were monitored simultaneously: 44 (calcium) and 34 (sulfur). The sulfur signal was used as an internal standard for the protein's elution.

Coordinates

The atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession number 1K9U.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online.

Acknowledgements

We would like to thank Christoph Kratky for supplying financial support and the infrastructure which made this work possible, Karl Gruber and Ulrike Wagner for helpful comments and discussions, and Winfried Mosler for assistance. We acknowledge the skilful help of Walter Goessler with the HPLC-ICP-MS experiments and thank all the people who were involved in data collection at the synchrotron beamlines X11 EMBL DESY Hamburg and XRD1 ELETTRA Trieste. This research was supported by the Austrian Science Fund FWF (P13283, P12559, Y078GEN, F018), the ICP program of the Austrian Ministry for Education and Science and by a research grant from BIOMAY, Vienna, Austria.

References

Babu,Y.S., Bugg,C.E. and Cook,W.J. (1988) Structure of calmodulin refined at 2.2 Å resolution. *J. Mol. Biol.*, **204**, 191–204.

Batanero,E., Villalba,M., Ledesma,A., Puente,X.S. and Rodriguez,R. (1996) Ole e 3, an olive-tree allergen, belongs to a widespread family of pollen proteins. *Eur. J. Biochem.*, **241**, 772–778.

Bennett,M.J., Schlunegger,M.P. and Eisenberg,D. (1995) 3D domain swapping: a mechanism for oligomer assembly. *Protein Sci.*, **4**, 2455–2468.

Brewbaker,J.L. and Kwack,B.H. (1963) The essential role of calcium ion in pollen germination and pollen tube growth. *Am. J. Bot.*, **50**, 859–865.

Brünger,A.T. *et al*. (1998) Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D*, **54**, 905–921.

CCP4 (1994) Collaborative Computational Project No. 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D*, **50**, 760–763.

Engel,E. *et al*. (1997) Immunological and biological properties of Bet v 4, a novel birch pollen allergen with two EF-hand calcium-binding domains. *J. Biol. Chem.*, **272**, 28630–28637.

Franklin-Tong,V.E. (1999) Signaling and the modulation of pollen tube growth. *Plant Cell*, **11**, 727–738.

Green,S.M., Gittis,A.G., Meeker,A.K. and Lattman,E.E. (1995) One-step evolution of a dimer from a monomeric protein. *Nat. Struct. Biol.*, **2**, 746–751.

Hakansson,M., Svensson,A., Fast,J. and Linse,S. (2001) An extended hydrophobic core induces EF-hand swapping. *Protein Sci.*, **10**, 927–933.

Hayek,B. *et al*. (1998) Molecular and immunologic characterization of a highly cross-reactive two EF-hand calcium-binding alder pollen allergen, Aln g 4: structural basis for calcium-modulated IgE recognition. *J. Immunol.*, **161**, 7031–7039.

Heringa,J. and Taylor,W.R. (1997) Three-dimensional domain duplication, swapping and stealing. *Curr. Opin. Struct. Biol.*, **7**, 416–421.

Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A*, **47**, 110–119.

Kay,A.B. (2001a) Allergy and allergic diseases. First of two parts. *N. Engl. J. Med.*, **344**, 30–37.

Kay,A.B. (2001b) Allergy and allergic diseases. Second of two parts. *N. Engl. J. Med.*, **344**, 109–113.

Kleywegt,G.J. and Jones,T.A. (1994) Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallogr. D*, **50**, 178–185.

Kraulis,P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.

Kretsinger,R.H. and Nockolds,C.E. (1973) Carp muscle calcium-binding protein. II. Structure determination and general description. *J. Biol. Chem.*, **248**, 3313–3326.

Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283–291.

Lewit-Bentley,A. and Rety,S. (2000) EF-hand calcium-binding proteins. *Curr. Opin. Struct. Biol.*, **10**, 637–643.

Liu,Y. and Eisenberg,D. (2002) 3D domain swapping: as domains continue to swap. *Protein Sci.*, **11**, 1285–1299.

Matsumura,H., Shiba,T., Inoue,T., Harada,S. and Kai,Y. (1998) A novel mode of target recognition suggested by the 2.0 Å structure of holo S100B from bovine brain. *Structure*, **6**, 233–241.

Merritt,E.A. and Murphy,M.E.P. (1994) Raster3D Version 2.0: a program for photorealistic molecular graphics. *Acta Crystallogr. D*, **50**, 869–873.

Milburn,M.V., Hassell,A.M., Lambert,M.H., Jordan,S.R., Proudfoot, A.E., Graber,P. and Wells,T.N. (1993) A novel dimer configuration revealed by the crystal structure at 2.4 Å resolution of human interleukin-5. *Nature*, **363**, 172–176.

Navaza,J. (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A*, **50**, 157–163.

Nicholls,A., Sharp,K.A. and Honig,B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins*, **11**, 281–296.

Niederberger,V. *et al*. (1999) Calcium-dependent immunoglobulin E recognition of the apo- and calcium-bound form of a cross-reactive two EF-hand timothy grass pollen allergen, Phl p 7. *FASEB J.*, **13**, 843–856.

Okada,T., Zhang,Z., Russell,S.D. and Toriyama,K. (1999) Localization of the Ca^{2+}-binding protein, Bra r 1, in anthers and pollen tubes. *Plant Cell Physiol.*, **40**, 1243–1252.

Otwinowski,Z. (1993) Oscillation and data reduction program. In Sawyers,L., Isaacs,N. and Bailey,S. (eds), *Proceedings of the CCP4*

Study Weekend: Data Collection and Processing. SERC Daresbury Laboratory, Warrington, UK, pp. 56–62.
Perrakis,A., Morris,R. and Lamzin,V.S. (1999) Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.*, **6**, 458–463.
Pierson,E.S., Miller,D.D., Callaham,D.A., van Aken,J., Hackett,G. and Hepler,P.K. (1996) Tip-localized calcium entry fluctuates during pollen tube growth. *Dev. Biol.*, **174**, 160–173.
Schymkowitz,J.W., Rousseau,F., Irvine,L.R. and Itzhaki,L.S. (2000) The folding pathway of the cell-cycle regulatory protein p13suc1: clues for the mechanism of domain swapping. *Struct. Fold. Des.*, **8**, 89–100.
Segal,D.M., Taurog,J.D. and Metzger,H. (1977) Dimeric immunoglobulin E serves as a unit signal for mast cell degranulation. *Proc. Natl Acad. Sci. USA*, **74**, 2993–2997.
Smith,P.M., Xu,H., Swoboda,I. and Singh,M.B. (1997) Identification of a Ca^{2+} binding protein as a new Bermuda grass pollen allergen Cyn d 7: IgE cross-reactivity with oilseed rape pollen allergen Bra r 1. *Int. Arch. Allergy Immunol.*, **114**, 265–271.
Strynadka,N.C.J. and James,M.N.G. (1991) Towards an understanding of the effects of calcium on protein structure and function. *Curr. Opin. Struct. Biol.*, **1**, 905–914.
Suphioglu,C., Ferreira,F. and Knox,R.B. (1997) Molecular cloning and immunological characterisation of Cyn d 7, a novel calcium-binding allergen from Bermuda grass pollen. *FEBS Lett.*, **402**, 167–172.
Terwilliger,T.C. (2000) Maximum-likelihood density modification. *Acta Crystallogr. D*, **56**, 965–972.
Terwilliger,T.C. and Berendzen,J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr. D*, **55**, 849–861.
Tinghino,R. *et al*. (2002) Molecular, structural and immunologic relationships between different families of recombinant calcium-binding pollen allergens. *J. Allergy Clin. Immunol.*, **109**, 314–320.
Twardosz,A., Hayek,B., Seiberler,S., Vangelista,L., Elfman,L., Gronlund,H., Kraft,D. and Valenta,R. (1997) Molecular characterization, expression in *Escherichia coli* and epitope analysis of a two EF-hand calcium-binding birch pollen allergen, Bet v 4. *Biochem. Biophys. Res. Commun.*, **239**, 197–204.
Vagin,A. and Teplyakov,A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.*, **30**, 1022–1025.
Valenta,R. (2002) The future of antigen-specific immunotherapy of allergy. *Nature Rev. Immunol.*, **2**, 446–453.
Valenta,R. and Kraft,D. (2001) Recombinant allergen molecules: tools to study effector cell activation. *Immunol. Rev.*, **179**, 119–127.
Valenta,R. *et al*. (1998a) The immunoglobulin E allergen interaction: a target for therapy of type I allergic diseases. *Int. Arch. Allergy Immunol.*, **116**, 167–176.
Valenta,R. *et al*. (1998b) Calcium-binding allergens: from plants to man. *Int. Arch. Allergy Immunol.*, **117**, 160–166.
Valenta,R., Vrtala,S., Focke Tejkl,M., Bugajska Schretter,A., Ball,T., Twardosz,A., Spitzauer,S., Gronlund,H. and Kraft,D. (1999) Genetically engineered and synthetic allergen derivatives: candidates for vaccination against Type I allergy. *Biol. Chem.*, **380**, 815–824.
Wills-Karp,M., Santeliz,J. and Karp,C.L. (2001) The germless theory of allergic disease: revisiting the hygiene hypothesis. *Nat. Immunol.*, **1**, 69–75.
Wolters-Arts,M., Lush,W.M. and Mariani,C. (1998) Lipids are required for directional pollen-tube growth. *Nature*, **392**, 818–821.
Yap,K.L., Ames,J.B., Swindells,M.B. and Ikura,M. (1999) Diversity of conformational states and changes within the EF-hand protein superfamily. *Proteins*, **37**, 499–507.

Received February 4, 2002; revised June 28, 2002;
accepted August 14, 2002