

A major IgE epitope-containing grass pollen allergen domain from Phl p 5 folds as a four-helix bundle

Ornella Maglio^{1,2}, Jose W.Saldanha¹, Susanne Vrtala³,
Susanne Spitzauer⁴, Rudolf Valenta³ and
Annalisa Pastore^{1,5}

¹NIMR, The Ridgeway, Mill Hill, London NW7 1AA, UK, ²Department of Chemistry, University of Napoli Federico II, Via Cynthia, I-80134 Naples, Italy, ³Department of Pathophysiology, Vienna General Hospital, Vienna and ⁴Clinical Institute for Medical and Chemical Laboratory Diagnostics, Vienna, Austria

⁵To whom correspondence should be addressed.
E-mail: apastor@nimr.mrc.ac.uk

Phl p 5, a 29 kDa major allergen from timothy grass pollen, is one of the most reactive members of group 5 allergens. Its sequence comprises two repeats of a novel alanine-rich motif (AR) whose structure and allergenic response are still mostly unknown. We report here a structural characterization of an immunodominant fragment of Phl p 5, Phl p 5(56–165) which comprises the first AR repeat. Recombinant (r)Phl p 5(56–165) was expressed in *Escherichia coli*, purified to homogeneity and shown to be sufficient to react with serum IgE from 90% of grass pollen allergic patients. Using NMR spectroscopy, we show conclusively that the fragment forms a compact globular domain which is, however, prone to degradation with time. The rPhl p 5(56–165) fold consists of a four-helix bundle held together by hydrophobic interactions between the aromatic rings and aliphatic side chains. This evidence gives clear indications about the structure of the full-length Phl p 5 and provides a rational basis for finding ways to stabilize the fold and designing therapeutic vaccines against grass pollen allergy.

Keywords: allergen/four-helix bundle/grass pollen/NMR/structure

Introduction

Grasses are among the most potent sources of allergens both because they produce large amounts of pollen and because they are widely distributed in temperate climate areas where a higher population density is usually concentrated (Esch, 1999). Allergens from various grasses and corn pollens have been classified according to their cross-reactive potential into different groups in order of their discovery (King *et al.*, 1995). Group 5 contains some of the most important grass pollen allergens which are recognized by IgE antibodies of almost all grass pollen allergic patients and induce strong IgE antibody responses and heavy clinical reactions (e.g. asthma) in sensitized patients (Suphioglu *et al.*, 1992; Vrtala *et al.*, 1993, 1998; Bufe *et al.*, 1998). cDNA cloning and sequence comparison of group 5 allergens from various grass and corn species have revealed high sequence homology as the molecular basis for the cross-reactivity of IgE antibody and T cell responses to group 5 allergens (Valenta *et al.*, 1996; Müller *et al.*, 1998; Niederberger *et al.*, 1998; Burton *et al.*, 1999). Phl p 5, a group 5 allergen from timothy grass, is a protein of 29 kDa

whose cDNA was first identified by Vrtala *et al.* (Vrtala *et al.*, 1993). Since then, it was shown that Phl p 5 exists as different isoforms (isoallergens) which can be classified into two major groups, Phl p 5a and b, with high sequence homology and similar allergenic activities (van Neerven *et al.*, 1999). Recombinant Phl p 5 (rPhl p 5) proteins have been produced and their immunological properties shown to be comparable to those from natural sources (Niederberger *et al.*, 1998). Ribonuclease activity has been attributed to Phl p 5b and localised in the C-terminus of the protein (Bufe *et al.*, 1996).

Understanding the structural rules that govern allergen–antibody recognition is a crucial step for designing recombinant versions of hypoallergenic variants which could then help to develop therapeutic strategies and/or specific vaccines against Type I allergies (Valenta *et al.*, 1999). The molecular basis of the Phl p 5/IgE recognition can therefore be clarified only by a detailed description of the structure of this allergen both as an isolated protein and in its IgE complex. However, only very little structural information is currently available for Phl p 5. This may be partially due to the instability of certain recombinant group 5 allergens: in a crystallization trial, the full-length rPhl p 5b degraded into smaller fragments, only one of which, spanning the last 133 C-terminal residues, would crystallize (Bufe *et al.*, 1996). Circular dichroism studies have shown that Phl p 5a has prevalently a helical secondary structure (Flicker *et al.*, 2000).

In this work, we characterized an N-terminal fragment of Phl p 5a, Phl p 5(56–165), comprising about one-third of the full-length Phl p 5. The fragment was identified by a monoclonal IgE isolated from a combinatorial IgE library that had been constructed from lymphocytes of a grass pollen allergic patient (Steinberger *et al.*, 1996). Phl p 5(56–165) has been shown to represent an immunodominant portion of Phl p 5 and to contain IgE and T cell epitopes (Steinberger *et al.*, 1996; Burton *et al.*, 1999). The corresponding region of the rye grass homologue (Lol p 5) is also a major IgE-reactive fragment (Vrtala *et al.*, 1998). A recombinant version of Phl p 5, rPhl p 5(56–165), strongly induces basophil histamine release and immediate type skin reactions in grass pollen allergic individuals (Flicker *et al.*, 2000). Here, we analysed IgE and IgG subclass recognition of purified rPhl p 5(56–165) using sera from grass pollen allergic patients and provided, for the first time, direct evidence supported by extensive NMR structure analysis combined with molecular modelling, that rPhl p 5(56–165) forms a compact structural domain with a four-helix bundle fold.

Materials and methods

Expression and purification of rPhl p 5(56–165)

rPhl p 5(56–165) was produced as described previously (Flicker *et al.*, 2000). In short, the construct was amplified by PCR techniques, using the full-length cDNA clone as a template and expressed in *Escherichia coli* BL21 (DE3) (Novagen). Cells were suspended in 25 mM imidazole, pH 7.4, 0.1% Triton

X-100, lysed by lysozyme (20 µg/g cells) and freeze–thawing. Bacterial DNA was digested by DNaseI (0.1 mg/g cells) incubated at room temperature for 20 min. After centrifugation of the lysate (20 000 g for 20 min), ammonium sulphate (50%, w/v) was added to the supernatant. The precipitate was resuspended in 10 mM Na₂PO₄ at pH 7.0 and desalted with a Sephadex PD10 column (Pharmacia). The eluate was loaded on an Sp-Sepharose column (Pharmacia) and eluted with a 0–0.5 M NaCl gradient. Fractions enriched for the protein were dialysed against 10 mM Tris buffer at pH 10, loaded on a diethylaminoethylcellulose-Sepharose column (Pharmacia) and eluted with a 0–0.5 M NaCl gradient. rPhl p 5(56–165), as pure as estimated both by 12% SDS–PAGE and by mass spectrometry, was desalted using a Sephadex PD10 column. Protein concentrations were determined by UV spectrophotometry using the protein extinction coefficient calculated according to von Hippel and Gill (von Hippel and Gill, 1989).

IgE and IgG subclass reactivity to rPhl p 5(56–165) and IgE competition experiments

Enzyme-linked immunosorbent assay (ELISA) plates (Nunc, Roskilde, Denmark) were coated with freshly purified rPhl p 5(56–165) (5 µg/ml PBS) at 4°C overnight. The plates were first washed twice with PBS, 0.05% Tween-20 and blocked for 3 h with PBS, 1% BSA, 0.05% Tween-20 at room temperature. They were then incubated with sera from grass pollen allergic patients which were diluted 1:10 and 1:50 in PBS, 0.5% BSA, 0.05% Tween-20 for detection of IgE and IgG1–4 antibodies, respectively, at 4°C overnight. The plates were finally washed five times with PBS, 0.5% BSA, 0.05% Tween-20. Bound immunoglobulins were detected with mouse monoclonal anti-human IgE and IgG1–4 antibodies (Pharmingen, San Diego, CA) and a horseradish peroxidase-coupled sheep anti-mouse antiserum (Amersham, Amersham, Buckinghamshire, UK) by ELISA (Vrtala *et al.*, 1996). Serum samples (50 µl aliquots) were diluted 1:1 with PBS and pre-incubated with 3 µg of freshly purified rPhl p 5(56–165) or, for control, with 3 µg of BSA (diluted in PBS to a concentration of 1 µg/µl) overnight at 4°C. The levels of specific IgE directed against the complete rPhl p 5 allergen from timothy grass, rye grass and Kentucky bluegrass pollen extracts were quantified by CAP-fluoro enzyme immunoassay (CAP-FEIA) measurements (Pharmacia, Uppsala, Sweden) as described (Kazemi-Shirazi *et al.*, 2000). The percentage inhibition of IgE binding was calculated under the assumption that pre-incubation with a protein that is immunologically unrelated to grass pollen proteins (i.e. BSA) will not reduce reactivity of IgE to grass pollen allergens and hence corresponds to 100% binding. The reduction of binding achieved by pre-incubation of sera with rPhl p 5(56–165) versus the binding obtained after pre-incubation of sera with a control protein (BSA) is expressed as percentage inhibition. The percentage inhibition was calculated as follows:

$$\% \text{ inhibition} = \frac{100 - \text{kUA/l} [\text{serum with rPhl p 5(56–165)}]}{100/\text{kUA/l} (\text{serum with BSA})} \times 100$$

where the levels of bound IgE are measured in kUA/l (kilounits of antigen per liter).

NMR analysis

NMR samples typically consisted of 0.5–1 mM protein samples dissolved in H₂O–D₂O (9:1) with 20 mM phosphate buffer at pH 7.0. The spectra were recorded at 25°C using either

Table I. Differential IgE and IgG subclass recognition of the Phl p 5(56–165) domain; the optical density values as determined by ELISA are shown for a few representative patients

Patient	IgE	IgG1	IgG2	IgG3	IgG4
1	0.889	0.578	0.391	0.067	0.083
5	0.116	0.720	0.323	0.066	1.042
11	0.997	0.558	0.363	0.101	0.104
14	1.369	0.430	0.167	0.092	0.138
25	0.182	0.483	0.191	0.237	0.076
29	1.133	0.544	0.223	0.156	0.082

Table II. Percentage inhibition of IgE binding to purified full-length rPhl p 5 to rye grass and Kentucky blue grass pollen extracts obtained after pre-adsorption of sera from grass pollen allergic patients with rPhl p 5(56–165)

Patient	Inhibition of IgE binding (%)		
	rPhl p 5	Rye grass	Kentucky blue grass
1	14	13	2
2	–	35	24
3	42	16	5
4	40	20	18
5	34	11	21
6	21	30	23
7	6	–	–
8	42	17	16
9	46	48	31
10	–	16	28
11	32	17	–
12	58	21	26
13	25	20	–
14	35	47	8
15	26	–	11
16	32	28	18
17	44	25	23
18	44	18	12
19	30	2	11
20	43	18	6
21	32	28	30
22	15	6	10
23	8	51	1
24	9	7	10
25	44	25	18
26	26	13	–
27	50	27	25
28	35	26	11
29	31	26	30
Av.	29	21	14

IgE binding was measured by CAP-FEIA (Pharmacia). The reduction of binding achieved by pre-incubation of sera with rPhl p 5(56–165) versus the binding obtained after pre-incubation of sera with a control protein (BSA) is expressed as percentage inhibition. Average inhibitions (Av.) for the 29 sera are displayed at the bottom. The patients are numbered as in Table I.

a Varian 600 MHz UNITY-PLUS or a Bruker 800 MHz spectrometer. 2D NOESY and TOCSY experiments were performed using 80 and 70 ms mixing times, respectively. A 3D NOESY-HSQC experiment with 80 ms mixing time was performed on a ¹⁵N-labelled sample. The water signal was suppressed using a WATERGATE sequence (Piotto *et al.*, 1992). The spectra were transformed and analysed using FELIX software.

Structure prediction and modelling

A multiple alignment of the homologous allergen sequences was obtained using CLUSTALX (Thompson *et al.*, 1997) and modified manually when necessary. Secondary structure

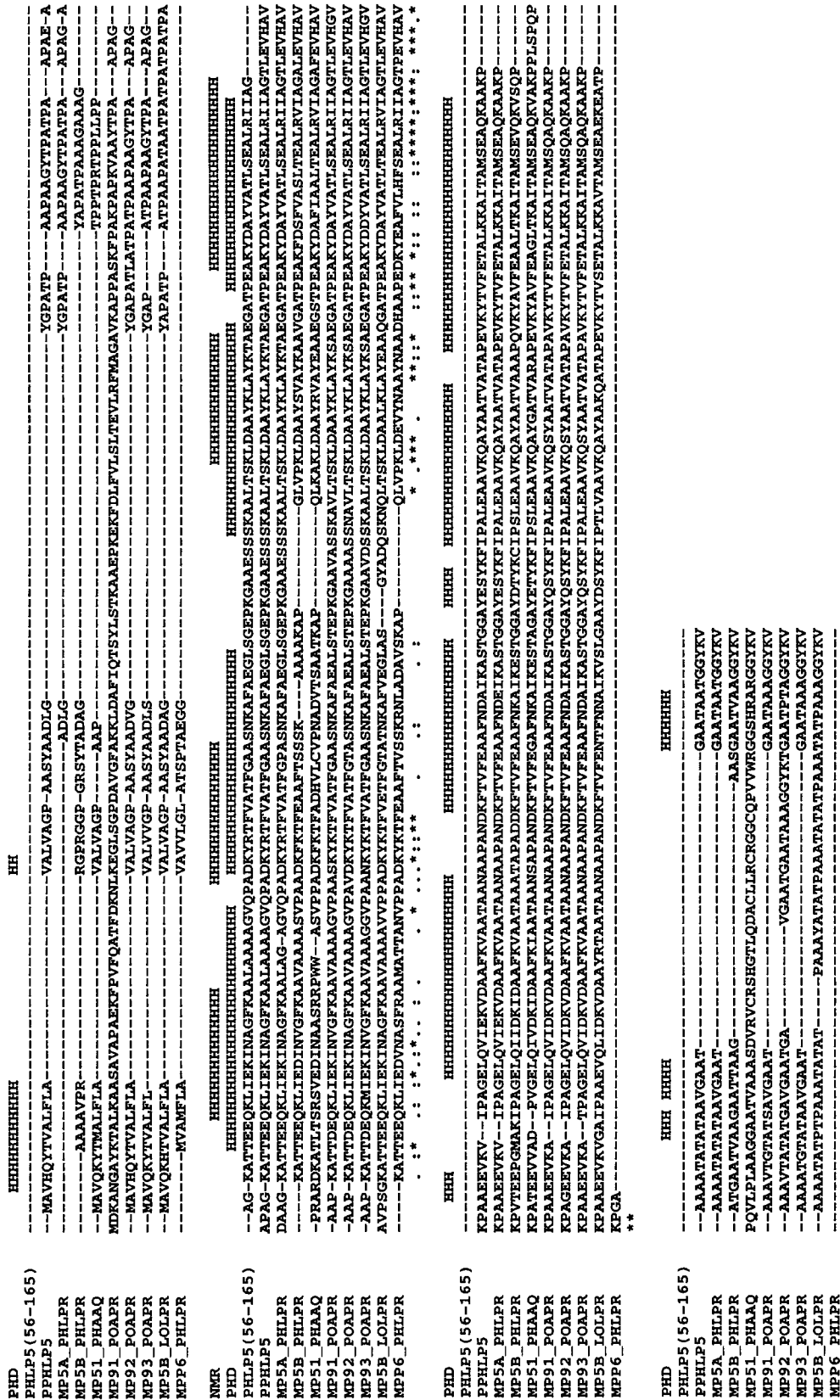


Figure 1. Multiple alignment of Phl p 5 sequences from different species. The alignment was generated by CLUSTALX (Thompson *et al.*, 1997). The sequences are identified with their SWISSPROT entry names. The sequence of the construct used for this work is indicated [rPhl p 5(56--165)]. The PHD secondary structure prediction (Rost and Sander, 1993) and the secondary structure as observed experimentally by NMR are indicated in the PHD and NMR lines, respectively. In the bottom line, total and partial conservation are indicated with stars, colons and dots.

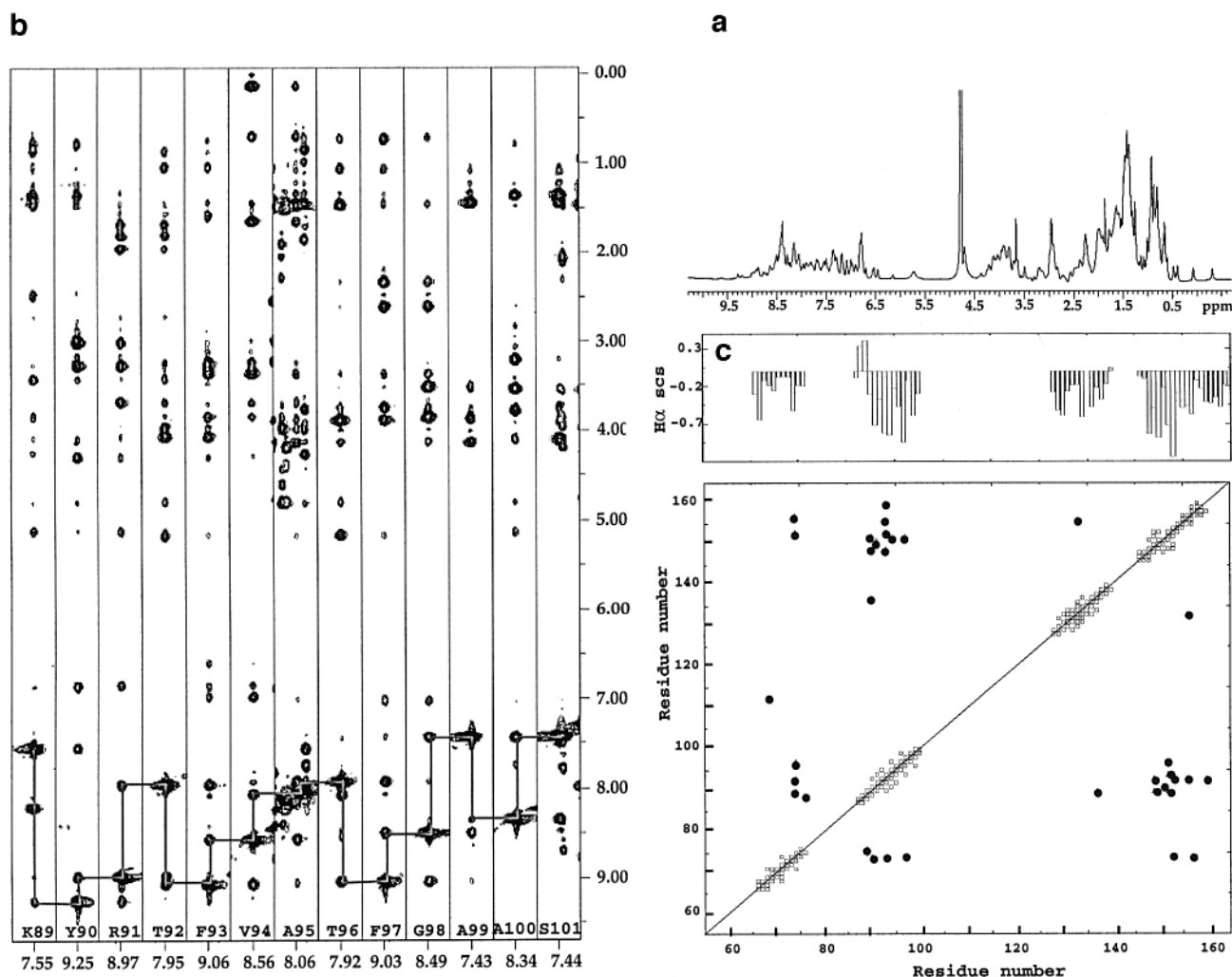


Figure 2. (a) ^1H NMR spectrum of rPhl p 5(56–165) recorded at 600 MHz and 27°C. The resonances show excellent spectral dispersion typical of a well defined 3D structure. (b) 3D ^1H - ^{15}N NOESY-HSQC of ^{15}N -labeled rPhl p 5(56–165) recorded at 600 MHz and 27°C. (c) Contact map showing medium and long-range connectivities identified unambiguously by NMR.

predictions were attempted by submitting the multiple alignment to the PHD server (Rost and Sander, 1993). The two sequence repeats were identified by psi-Blast (Altschul and Koonin, 1998).

The following procedure was adopted to identify a suitable template for modelling. First, extensive search through the sequence databases was carried out by psi-Blast (Altschul and Koonin, 1998) but no homology could be spotted outside the Phl p 5 family. The sequence of Phl p 5(56–165) was then used as a query sequence to the UCLA–DOE structure prediction server at URL <http://www.doe-mbi.ucla.edu/~frsvr/frsvr.html> (Fischer and Eisenberg, 1996). This query gave a fold prediction of cytochrome *c'* from *Rhodospseudomonas palustris* (PDB code 1A7V) (Shibata *et al.*, 1998) together with an initial alignment of the sequences (Z score 2.6).

A structural model of Phl p 5(56–165) was built on 1A7V using QUANTA (Release 4.0, MSI). The experimental NMR restraints were analysed with respect to this model, which suggested how the sequence alignment could be amended. This procedure was iterated, changing the alignment at each step in an attempt to produce a model which satisfied all the NMR restraints.

Results

rPhl p 5(56–165) represents a major IgE epitope-containing grass pollen allergen domain

The sera from 48 grass pollen allergic patients were analysed by ELISA for the presence of IgE antibodies against both the full-length protein and the rPhl p 5(56–165) fragment. Forty-three of these sera (89.6%) exhibited IgE reactivity to both molecules whereas shorter synthetic peptides (e.g. a peptide comprising residues 97–128 of Phl p 5) did not bind IgE (data not shown). When IgE and IgG subclass reactivity to the rPhl p 5 (56–165) fragment was analysed, an interesting diversification of the IgE and IgG subclass recognition of the allergen domain was observed as exemplified in Table I for six representative patients. Patient 1 exhibits IgE, IgG1 and IgG2 subclass but almost no IgG3 or IgG4 reactivity to the molecule. Certain patients contained IgE and IgG antibodies of each of the subclasses to rPhl p 5 (56–165) (e.g. patient 11), whereas others (e.g. patients 25 and 29) lacked specific IgG4 or (e.g. patient 14) had very low specific IgG3 antibodies. This finding demonstrates that IgE and IgG subclass responses in allergic patients can be directed to different epitopes on the

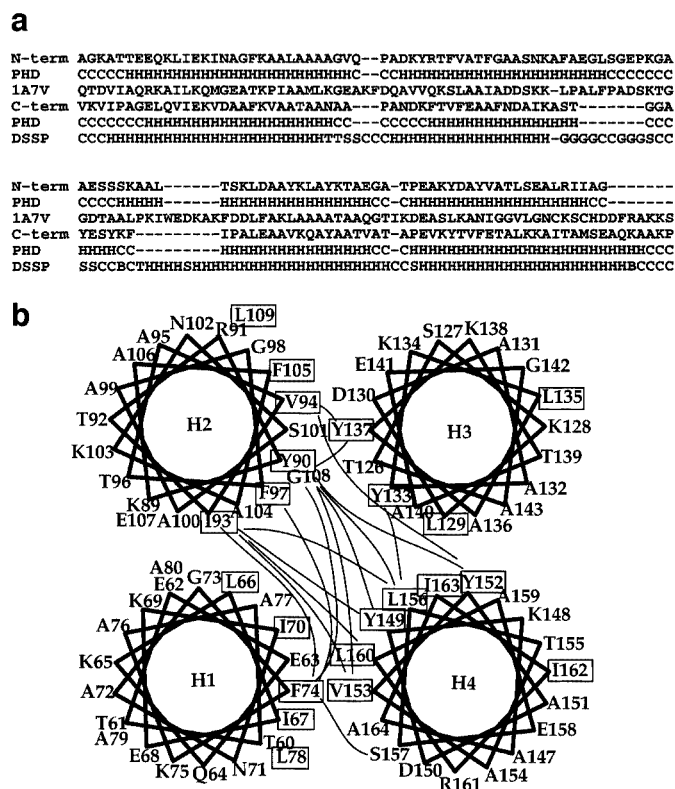


Figure 3. (a) Alignment between the Phl p 5 and its template. Sequence alignment of the first domain of Phl p 5 with the sequence of cytochrome *c'* from *Rhodopseudomonas palustris* (Shibata *et al.*, 1998) (PDB code 1A7V). The secondary structure of Phl p 5 was predicted using PHD (Rost and Sander, 1993) and the secondary structure of 1A7V was assigned using DSSP (Kabsch and Sander, 1983). H, helix; C, coil; G, 3/10 helix; S, bend; T, turn; B, bridge. (b) Helical wheel representation of the structure of Phl p 5(56–165). The helices are assembled in an up–down–up–down topology. The handedness of the bundle is chosen arbitrarily. Hydrophobic and aromatic residues are boxed. The contacts observed by NMR are shown explicitly.

allergen molecule. The latter can be explained by the activation of a great variety of B cell clones by different epitopes of the allergen during the sensitization process. Since allergen-specific IgE antibodies react to epitopes different from those recognized by IgG, it is unlikely that they have evolved by sequential class switch from B cells producing initially allergen-specific IgG.

Next, we determined the percentage of Phl p 5 specific IgE response which can be pre-adsorbed from sera with rPhl p 5(56–165). Quantitative IgE competition experiments performed by 29 Phl p 5 reactive sera showed that rPhl p 5(56–165) accounts on average for 29% of the Phl p 5 reactive IgE antibodies. Moreover, rPhl p 5 (56–165) cross-reacted with allergens present in pollens of other grass species. Preadsorption with the domain depleted on average 21% of rye grass pollen and 14% of Kentucky blue grass pollen-specific IgE from the sera (Table II).

These results prove that the rPhl p 5(56–165) fragment binds a significant percentage of allergen-specific IgE of grass pollen allergic patients and thus the fragment contains a relevant percentage of the full-length Phl p 5 IgE reactivity.

rPhl p 5(56–165) has a well defined tertiary structure

Sequence analysis of Phl p 5 reveals that the protein, and the whole group 5 allergens, are assembled by two repeats of an AR motif (alanines account for ~30% of the amino acid composition) (Figure 1). Only one AR repeat is observed in

group 6 allergens. In Phl p 5, the two repeats comprise approximately residues 55–174 and 183–284 and are probably the consequence of gene duplication. Besides close Phl p 5 homologues, the AR motif does not share detectable sequence homology with any other known protein, making it difficult to predict its structure. However, alanines have a very strong tendency to form helices (Vila *et al.*, 2000) and such an unusual composition is found for instance in trans-membrane helices and in antifreeze proteins (Davies and Sykes, 2000; Eilers *et al.*, 2000). Accordingly, the far-UV circular dichroism spectra of both rPhl p 5 and of rPhl p 5(56–165) are typical of α -helical proteins (Flicker *et al.*, 2000), with an estimated helical content of ~60%. Secondary structure prediction based on a multiple alignment of group 5 allergens from different species suggests the approximate position of the helical regions along the sequence (Figure 1).

The presence of tertiary structure was probed by NMR spectroscopy. The proton spectrum of rPhl p 5(56–165) shows excellent signal dispersion typical of globular domains (Figure 2a). In particular, the presence of several resonances shifted in the region 0.5–0.0 ppm is typically caused by the persistent interaction of aliphatic chains with aromatic rings expected for a well-folded protein (Wütrich, 1986). We can therefore conclude that rPhl p 5(56–165) is an independently folded globular domain.

As the full-length protein, also rPhl p 5(56–165) degrades with time

Although thermal denaturation curves show that rPhl p 5(56–165) is thermally stable and that its unfolding is mostly reversible (Flicker *et al.*, 2000), the protein degrades quickly with time. Four samples from independent batches of preparation were tested. The average lifetime of each sample was 1–2 weeks. During this time, the solutions became progressively cloudy and a precipitate appeared. The samples were analysed by mass spectrometry and N-terminal sequencing typically at time 0 and 1–2 weeks after sample preparation. The pattern of degradation observed by mass was consistent with progressive degradation from the N-terminus (data not shown). Such a behaviour, which could not be prevented by varying experimental conditions, was reflected directly by the NMR spectra. Analysis of 2D NMR experiments recorded on independently prepared batches of protein showed the same general features but spectra recorded on a protein batch a few days old progressively contained additional peaks which were absent in fresh samples (data not shown).

Direct NMR evidence shows that rPhl p 5(56–165) is a four-helix bundle

Because of the progressive degradation, spectral assignment constituted a real challenge. A set of homonuclear 2D NMR experiments and a ^{15}N NOESY-HSQC spectrum were recorded on fresh aliquots of three independent batches of unlabelled protein and on a ^{15}N -labelled sample, respectively. The quality of the heteronuclear spectrum can be appreciated in Figure 2b. Reliable resonance assignment could be achieved unambiguously only for the regions L66–A76, A87–A100, K128–A140 and E146–G165 (52% of the total sequence) for which a clear match to NOESY and TOCSY connectivities was possible. Interruption of the assignment was caused by peak overlap and could not be confidently extended further. All the residues in the assigned regions are in an α -helical conformation as directly supported by the $\text{H}\alpha$ secondary chemical shifts and by the presence of HN–HN(*i*, *i* + 1), $\text{H}\alpha$ –HN(*i*, *i* + 3),

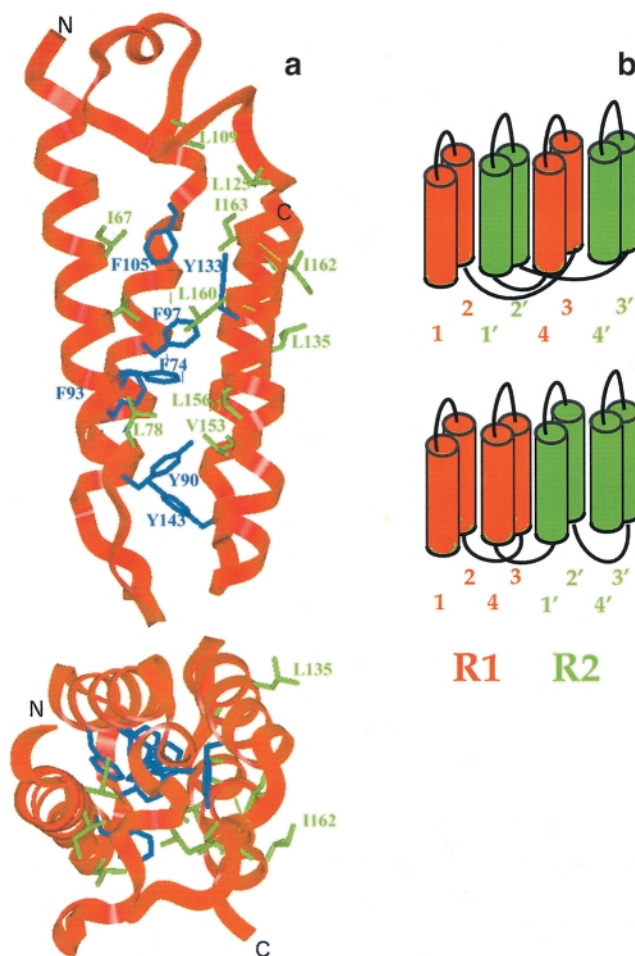


Figure 4. (a) Two orthogonal views of the 3D model of Phl p 5(56–165) based on 1A7V using the alignment shown in Figure 3a. The side chains of hydrophobic and aromatic residues are shown in green and blue, respectively. The 3D model is in excellent agreement with what is expected from the helix wheel diagram obtained independently. (b) Schematic representation of the predicted Phl p 5 fold. The two AR repeats could either pack side by side or they could swap one or two helices. One example of a swapped topology is given here.

$H\alpha-HN(i, i + 3)$ and $H\alpha-HN(i, i + 3)$ connectivities (Figure 2c). These results are in excellent agreement with the predicted location of secondary structure elements (Figure 1). The unassigned regions are in predicted loops and at the N-terminus of the predicted helix 3.

Direct indication about the tertiary fold could be obtained from the unambiguous assignment of several long-range contacts (38 restraints) involving eight out of the nine aromatic side chains (Figure 2c). These contacts are consistent with the domain being formed by four interrupted helices which fold into a globular four-helix bundle domain held together by hydrophobic interactions between the aromatic protons and other aliphatic side chains.

A molecular model of the first AR motif can be inferred from experimental evidence

The modelling of the Phl p 5(56–165) fold was based on the following facts. (a) The domain folds into a compact (although prone to degradation) structure formed by not fewer than four helices as supported by the experimentally identified long-range contacts. (b) Of the several topologies in principle possible for a four-helix bundle, we can exclude those that require a long connectivity between adjacent helices 1–2 and 3–4 (Figure 3a). These linkers are not long enough to allow inversion of the helix direction. (c) Further restraints on the possible topologies are imposed by the length of the linker

between helices 2 and 3 in the second AR repeat (Figures 1 and 3a). We must therefore conclude that the domain consists of a four-helix bundle with an up–down–up–down topology, a fold never observed before for an allergen. A helix wheel representation based on the NMR restraints was used to predict the relative orientation of the helices and their packing (Figure 3b). Orientation of the helices in such a way to satisfy the experimental long-range contacts observed experimentally also leads to maximization of the number of buried hydrophobic residues.

Although no known structure shows appreciable sequence homology with Phl p 5, a sound, albeit low-resolution 3D model of the domain can be prepared using as a template the structure of cytochrome *c'* from *Rhodospseudomonas palustris* (Shibata *et al.*, 1998). This structure was the highest hit suggested by a threading algorithm and, since its secondary structure elements are of length comparable to those of Phl p 5(56–165), the alignment of the two sequences requires relatively few insertions/deletions. The resulting model, which derives from the alignment shown in Figure 3a, satisfies all the NMR restraints (Figure 4a). Interestingly, when adopting this alignment, while the hydrophobic core is buried in the domain interior, two adjacent hydrophobic residues (Leu135 and Ile162) remain exposed, suggesting a potential surface of interaction with the second AR repeat.

Discussion

We have characterized rPhl p 5(56–165), a fragment of Phl p 5 which comprises about one-third of the full-length protein and contains one of the two repeats of a novel sequence AR motif. The allergenic properties of rPhl p 5(56–165) were measured quantitatively and compared with those of the full-length protein. We show that almost all grass pollen allergic patients contain IgE antibodies against rPhl p 5(56–165) and a high percentage of Phl p 5-specific IgE is recognized also by rPhl p 5(56–165), in agreement with previous work which demonstrated that rPhl p 5 induces effector cell activation, basophil histamine release and immediate-type skin reactions (Flicker *et al.*, 2000). The rPhl p 5(56–165) fragment is therefore a major allergenic agent on its own.

We have further shown that rPhl p 5(56–165) is folded and acts as an independently folded unit. Phl p 5 provides another example of preferential recognition of conformational epitopes by IgE antibodies as demonstrated for an increasing number of respiratory allergens (Vrtala *et al.*, 1997, 1999; Focke *et al.*, 2001): only the full-length rPhl p 5(56–165) is recognized by IgE antibodies whereas smaller Phl p 5-derived peptides failed to react with IgE antibodies (data not shown).

Structure determination of both the full-length protein and/or its fragments has proved difficult. In our hands, rPhl p 5(56–165) is unstable and prone to degradation. This behaviour did not, however, come as a surprise: already in 1996, it was observed that recombinant Phl p 5 is unstable in solution and degrades into fragments of variable length between 10 and 29 kDa (Bufo *et al.*, 1996). Of these, only the fragment rPhl p 5(180–313) was able to crystallize while the fragment described in the present study was not identified as a direct product of degradation. These observations strongly suggest that the instability is a property intrinsic to recombinant Phl p 5 rather than being accidentally related to the experimental conditions. The behaviour observed for rPhl p 5(180–313) is strongly reminiscent of apolipoprotein III, a lipid-associated four-helix bundle, which also folds as a four-helix bundle prone to degradation in solution (Wang *et al.*, 1997). It is likely that *in vivo* also Phl p 5 needs to be stabilized either by a cofactor or by a specific environment, which is lost in the recombinant protein.

We have presented the first direct experimental evidence to allow reconstruction of the main features of its fold using an approach that combines experimental restraints with molecular modelling. The NMR observables (i.e. secondary chemical shifts, short-range NOE connectivities), together with a number of unambiguous long-range side chain contacts, conclusively show that rPhl p 5(56–165) forms a compact four-helix bundle with an up–down–up–down topology. The fold is held together by an intricate network of aromatic and hydrophobic interactions which form the hydrophobic core (Figures 3b and 4a). This conclusion is in excellent agreement with a preliminary report on the crystal structure of a rPhl p 5b(180–313) fragment (which spans the second AR repeat), also described to form a four-helix bundle with similar topology (Bufo *et al.*, 2000). We can therefore conclude that Phl p 5 is an all- α protein mostly assembled by two four-helix bundles packing against each other or intertwined by domain swapping (Figure 4b). Interactions between the two repeats are supported by the reported capability of rPhl p 5(180–313) to dimerize (Bufo *et al.*, 2000) and by the presence of conserved exposed hydrophobic side chains on the surface of Phl p 5(56–165)

(Figure 4). Interestingly, despite the fold duplication, the N- and C-terminal domains seem to have a different IgE response (Bufo *et al.*, 1994). A helical bundle fold, never described before in allergens, will therefore need to be added to the list of observed topologies (Aalberse, 2000). Because of the high sequence homology, group 5 and 6 grass pollen allergens are all expected to share the same fold.

While further studies are still necessary to explore and characterize the structural basis of the allergenic response of Phl p 5, the information presented in this paper should provide useful guidelines for designing both more stable and hypoallergic versions of Phl p 5 or its parts which can then be used as vaccines.

Acknowledgements

This study was supported by grant Y078GEN from the Austrian Science Fund and by the ICP project of the Austrian Ministry of Education and Science. O.M. was supported by a NATO fellowship.

References

- Aalberse, R.C. (2000) *J. Allergy Clin. Immunol.*, **106**, 228–238.
- Altschul, S.F. and Koonin, E.V. (1998) *Trends Biochem. Sci.*, **23**, 444–447.
- Bufo, A., Becker, W.M., Schramm, G., Petersen, A., Mamat, U. and Schaak, M. (1994) *J. Allergy Clin. Immunol.*, **94**, 173–181.
- Bufo, A., Betzel, C., Schramm, G., Petersen, A., Becker, W.M., Schlaak, M., Perbandt, M., Dauter, Z. and Weber, W. (1996) *J. Biol. Chem.*, **271**, 27193–27196.
- Bufo, A.K., Gehlar, G., Schramm, M., Schlaak, M. and Becker, W.M. (1998) *Am. J. Respir. Crit. Care Med.*, **157**, 1269–1276.
- Bufo, A., Rajashankar, K., Eschenburg, S., Lindmer, B., Betzel, C. and Weber, W. (2000) *Allergy, Suppl.* **63**, 55, 32.
- Burton, M.D., Blaher, B., Suphioglu, C., O’Hehir, R.E., Carbone, F.R. and Rolland, J.M. (1999) *J. Allergy Clin. Immunol.*, **103**, 255–261.
- Davies, P.L. and Sykes, B.D. (2000) *Curr. Opin. Struct. Biol.*, **7**, 828–834.
- Eilers, M., Shekar, S.C., Shieh, T., Smith, S.O. and Fleming, P.J. (2000) *Proc. Natl Acad. Sci. USA*, **97**, 5796–5801.
- Esch, R.E. (1999) In Lockey, R.F. and Bukantz, S.C. (eds), *Allergens and Allergen Immunotherapy*. Marcel Dekker, New York, pp. 103–120.
- Fischer, D. and Eisenberg, D. (1996) *Protein Sci.*, **5**, 947–955.
- Flicker, S. *et al.* (2000) *J. Immunol.*, **165**, 3849–3859.
- Focke, M., Mahler, V., Ball, T., Sperr, W.R., Majlesi, Y., Valent, P., Kraft, D. and Valenta, R. (2001) *FASEB J.*, **15**, 2042–2044.
- Kabsch, W. and Sander, C. (1983) *Biopolymers*, **22**, 2577–2637.
- Kazemi-Shirazi, L., Pauli, G., Purohit, A., Spitzauer, S., Froschl, R., Hoffmann-Sommergruber, K., Breiteneder, H., Scheiner, O., Kraft, D. and Valenta, R. (2000) *J. Allergy Clin. Immunol.*, **105**, 116–125.
- King, T.P., Hoffman, D., Lowenstein, H., Marsch, D.G., Platts-Mills, T.A.E. and Thomas, W. (1995) *J. Allergy Clin. Immunol.*, **96**, 3–14.
- Müller, W.D., Karamfilov, T., Kahlert, H., Stuewe, H.T., Fahlbusch, B., Cromwell, O., Fiebig, H. and Jäger, L. (1998) *Clin. Exp. Allergy*, **28**, 1538–1548.
- Niederberger, V., Laffer, S., Froschl, R., Kraft, D., Rumpold, H., Kapiotis, S., Valenta, R. and Spitzauer, S. (1998) *J. Allergy Clin. Immunol.*, **101**, 258–264.
- Piotto, M., Saudek, V. and Sklenar, V. (1992) *J. Biomol. NMR*, **2**, 661–665.
- Rost, B. and Sander, C. (1993) *J. Mol. Biol.*, **232**, 584–599.
- Shibata, N., Iba, S., Misaki, S., Meyer, T.E., Bartsch, R.G., Cusanovich, M.A., Morimoto, Y., Higuchi, Y. and Yasuoka, N. (1998) *J. Mol. Biol.*, **284**, 751–760.
- Steinberger, P., Kraft, D. and Valenta, R. (1996) *J. Biol. Chem.*, **271**, 10967–10972.
- Suphioglu, C., Singh, M.B., Taylor, P., Bellomo, R., Holmes, P., Puy, R. and Knox, R.B. (1992) *Lancet*, **339**, 569–572.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) *Nucleic Acids Res.*, **25**, 4876–4882.
- van Neerven, R.J.J., Wissenbach, M., Ipsen, H., Bufo, A., Arnved, J. and Wuertzen, P.A. (1999) *Int. Arch. Allergy Immunol.*, **118**, 125–128.
- Valenta, R., Steinberger, P., Duchene, M., Kraft, D. (1996) *Immunol. Cell Biol.*, **74**, 187–194.
- Valenta, R., Vrtala, S., Focke-Tejkl, M., Bugajska-Schretter, A., Twardosz, A., Spitzauer, S., Groenlund, H. and Kraft, D. (1999) *J. Biol. Chem.*, **380**, 815–824.

- Vila,J.A., Ripoll,D.R. and Scheraga H.A. (2000) *Proc. Natl Acad. Sci. USA*, **97**, 13075–13079.
- Vrtala,S. *et al.* (1993) *J. Immunol.*, **151**, 4773–4781.
- Vrtala,S., Susani,M., Sperr,W.R., Valent,P., Laffer,S., Dolecek,C., Kraft,D. and Valenta,R. (1996) *J. Allergy Clin. Immunol.*, **97**, 781–787.
- Vrtala,S., Hirtenlehner,K., Vangelista,L., Pastore,A., Eichler,H.G., Sperr,W.R., Valent,P., Ebner,C., Kraft,D. and Valenta,R. (1997) *J. Clin. Invest.*, **99**, 1673–1681.
- Vrtala,S., Ball,T., Spitzauer,S., Pandjaitan,B., Suphioglu,C., Knox,B., Sperr,W.R., Valent,P., Kraft,D. and Valenta,R. (1998) *J. Immunol.*, **160**, 6137–6144.
- Vrtala,S., Fischer,S., Grote,M., Vangelista,L., Pastore,A., Sperr,W.R., Valent,P., Reichelt,R., Kraft,D. and Valenta,R. (1999) *J. Immunol.*, **163**, 5489–5496.
- von Hippel,P.H. and Gill,S.C. (1989) *Anal. Biochem.*, **182**, 319–322.
- Wang,J., Gagne,S.M., Sykes,B.D. and Ryan,R.O. (1997) *J. Biol. Chem.*, **272**, 17912–17920.
- Wütrich,K. (1986) *NMR of Proteins and Nucleic Acids*. Wiley, New York.

Received January 21, 2002; revised April 22, 2002; accepted May 21, 2002