Cross-reactivity of pollen and food allergens: soybean Gly m 4 is a member of the Bet v 1 superfamily and closely resembles yellow lupine proteins

Hanna BERKNER*1, Philipp NEUDECKER*1,2, Diana MITTAG†3, Barbara K. BALLMER WEBER†, Kristian SCHWEIMER*, Stefan VIETHS‡ and Paul RÖSCH*4

*Department of Biopolymers and Research Center for Bio-Macromolecules, Universität Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany, †Allergy Unit, Department of Dermatology, University Hospital Zürich, Gloriastr. 31, 8091 Zürich, Switzerland, and ‡Department of Allergology, Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51–59, 63225 Langen, Germany

Synopsis

In many cases, patients allergic to birch pollen also show allergic reactions after ingestion of certain fruits or vegetables. This observation is explained at the molecular level by cross-reactivity of IgE antibodies induced by sensitization to the major birch pollen allergen Bet v 1 with homologous food allergens. As IgE antibodies recognize conformational epitopes, a precise structural characterization of the allergens involved is necessary to understand cross-reactivity and thus to develop new methods of allergen-specific immunotherapy for allergic patients. Here, we report the three-dimensional solution structure of the soybean allergen Gly m 4, a member of the superfamily of Bet v 1 homologous proteins and a cross-reactant with IgE antibodies originally raised against Bet v 1 as shown by immunoblot inhibition and histamine release assays. Although the overall fold of Gly m 4 is very similar to that of Bet v 1, the three-dimensional structures of these proteins differ in detail. The Gly m 4 local structures that display those differences are also found in proteins from yellow lupine with known physiological function. The three-dimensional structure of Gly m 4 may thus shed some light on the physiological function of this subgroup of PR10 proteins (class 10 of pathogenesis-related proteins) and, in combination with immunological data, allow us to propose surface patches that might represent cross-reactive epitopes.

Key words: allergen, epitope, histamine release, NMR, protein structure, soybean

INTRODUCTION

15–20% of the population of industrialized countries, with tendency to increasing numbers, suffer from pollinosis, which is often caused by sensitization to birch pollen in moderate climate zones of Europe [1]. In approx. 90% of birch pollen-allergic patients, the major birch pollen allergen Bet v 1 is responsible for allergic symptoms such as hay fever or asthma [2]. Additionally, birch pollen allergy is often accompanied by allergic reactions after ingestion of specific plant-derived food, including fruits, nuts and vegetables, and this association finds its molecular correlate in structural similarities between respective allergens. A type I allergic reaction like hay fever is mediated by cross-linking of allergen-specific IgE antibodies bound to mast cells or basophils by allergens, which, in turn, leads to degranulation of effector cells and release of histamine and other mediators provoking allergic symptoms. Allergens that show sufficiently similar IgE epitopes may react with the same IgE molecules. IgE epitopes are thought to be conformational rather than sequential [3], and patients sensitized to one allergen may show allergic reactions to proteins with similar three-dimensional structures. Interestingly, patients in birch pollen-related food allergy frequently show multiple food allergies to several but not all foods known to be...
involved in this clinical syndrome, indicating that differences in the epitope structure of the causative allergen molecules may exist.

The identification of proteins highly homologous with the major birch pollen allergen Bet v 1 in pollen allergy-related foods like cherry, apple, hazelnut, carrot, celery or soybean (Figure 1) was the initial step towards an explanation for allergic cross-reactivity at the molecular level.

According to their sequence similarities and related joint characteristics (molecular mass: approx. 17 kDa; acidic pl: approx. 6), these proteins are members of the Bet v 1 protein superfamily and the family of PR10 proteins (class 10 of pathogenesis-related proteins) according to the structural classification of proteins [4]. The question concerning the physiological function of PR10 proteins is still unanswered. Recently, this open question was addressed with the discovery that lupine protein LI PR10.2B serves as a cytokinin reservoir [5] and with a structural study suggesting the norcoclaurine synthase to be a true member of the PR10 family [6]. Clearly, detailed structural studies of the PR10 proteins may contribute to unravel their relationships to one another and thus add to the knowledge of functional features of this class of proteins.

The notion that the similarity of the three-dimensional structures of allergens is key to understanding cross-reactivity was confirmed by comparison of the conformations of allergens from various plants: the three-dimensional structures of the major birch pollen allergen Bet v 1 [7], the major cherry allergen Pru av 1 [8] and the celery allergen Api g 1 [9] not only comprise identical arrangements of secondary structure elements, but also they are virtually identical, explaining the phenomenon of cross-reactivity at a molecular level. Detailed analysis of common features and differences regarding primary and tertiary structures of Bet v 1 homologous allergens can thus help us to identify specific and cross-reactive epitopes on the molecular surface. Although mutational epitope analysis and mimotope selection can furnish information on putative IgE epitopes of Bet v 1 and its homologues, the complete picture necessary to construct hypoallergenic protein variants suitable for immunotherapy requires structural information in atomic detail.

The Bet v 1 homologous protein Gly m 4 from soybean was identified at the mRNA level in 1992 [10], and the putative protein was shown to be induced under stress conditions like starvation, appropriating the protein nomenclature SAM22 (starvation-associated message 22). Later, SAM22 was suggested to be the molecule responsible for allergic reactions to soybean in birch pollen-allergic patients sensitized to Bet v 1 [11], which could be confirmed in subsequent clinical and immunological studies [12]. In consequence, the protein was renamed Gly m 4 according to the general rules for allergen nomenclature [13]. Remarkably, in contrast with other Bet v 1 homologous proteins, the allergic reactions caused by Gly m 4 tend to be severe [14].

As food containing soy protein becomes more and more popular worldwide, soy allergy in patients sensitized to birch pollen represents a major health problem, rendering the detailed analysis of the Gly m 4 structure and its immunological properties an important step towards understanding the molecular basis of this phenomenon and towards the development of safe and efficacious immunotherapy of pollen-related food allergy.

# EXPERIMENTAL

## Structure determination of Gly m 4

### Sample preparation

rGly m 4 (recombinant Gly m 4) was expressed, purified and stored as freeze-dried powder as described in [15]. NMR samples were prepared by dissolving either $^{15}$N- or $^{13}$C-labelled Gly m 4 in 10–50 mM potassium phosphate buffer (pH 7.0), 10% $^2$H$_2$O and 0.02% sodium azide. For specific measurement of NOEs (nuclear Overhauser effects) involving aromatic protons, unlabelled Gly m 4 was dissolved in 50 mM potassium phosphate buffer (pH 7.0) in $^2$H$_2$O. The final sample concentrations were 0.7–1.8 mM.

### NMR experiments

All NMR experiments were performed at 298 K on Bruker DRX 600 MHz, AV 600 MHz and DMX 750 MHz spectrometers with pulsed field-gradient capabilities according to standard experimental procedures [16]. In addition to the experiments for resonance assignment [15], the following NOESY experiments were performed: two-dimensional [1H,1H]-NOESY using unlabelled Gly m 4 in $^2$H$_2$O; 15N-NOESY-HSQC (heteronuclear single-quantum coherence); 13C-NOESY HSQC with 120 ms mixing time; and 15N,13N-HMQC (heteronuclear single quantum correlation)-NOESY-HSQC, 13C,15N-HMQC-NOESY-HSQC with 150 ms mixing time. Slowly exchanging amide protons involved in H-bond formation were identified from a 15N-HSQC spectrum recorded 1 h after the freeze-dried protein was dissolved in $^2$H$_2$O. $^3$JNHx scalar coupling constants were derived from the HNHA experiment. $^2$D$_{\text{HN}}$ RDC (residual dipolar coupling) constants were measured with a Gly m 4 sample weakly aligned by addition of 18 mg/ml Pf1 phage particles using IPAP (inphase antiphase) experiments. [1H]$^3$N steady state NOE values were measured applying standard pulse sequences [17]. NMR data were processed with in-house written software and analysed using NMRView [18].

### Structure calculation

NOESY cross peaks were classified according to their relative intensities and converted into distance restraints with upper limits of 3.0 Å (1 Å = 0.1 nm) (strong), 4.0 Å (medium), 5.0 Å (weak) and 6.0 Å (very weak). For ambiguous distance restraints, the $r^{-6}$ summation over all assigned possibilities defined the upper limit. Distances introduced to account for the involvement of an amide proton in an H-bond were set to 2.3 Å between acceptor and amide proton, and 3.5 Å between acceptor and amide nitrogen. Dihedral restraints for $\Phi$ angles were derived from $^3$JNHx scalar coupling constants including a $\Phi$ angle between $-80^\circ$ and $-40^\circ$ for coupling constants smaller than...
6.0 Hz, whereas values greater than 8.0 Hz were converted into a Φ angle between −160° and −80°. Additional dihedral restraints for Φ and ψ angles were derived from chemical shifts using the TALOS program [19] (Table 1). Structure calculations were performed with the program XPLOR-NIH 1.2.1 [20] using a three-step simulated annealing protocol with floating assignment of prochiral groups including a conformational database potential. The 25 structures with the lowest total energy were then refined using the RDCs with a harmonic potential. RDCs measured for amino acids located in areas flexible on the picosecond to nanosecond timescale (\( ^1H \)\(^{15}N\)-NOE < 0.65) were excluded from the calculations. The 20 structures showing the lowest values of the target function excluding the database potential were further analysed with the programs XPLOR [20], PROCHECK 3.5.4 [21] and MOLMOL (DeLano Scientific; http://pymol.sourceforge.net/). Surface areas of putative cross-reactive IgE epitopes were calculated using the software nACCESS (S.J. Hubbard and J.H. Thornton, Department of Biochemistry and Molecular Biology, University College London, London, U.K.) with a probe radius of 2.5 Å. Co-ordinates were deposited in the PDB (accession code 2K7H).

Soybean allergic patients

Sera of six patients (age: 22 ± 5; 5 females and 1 male) with food allergy to soy as confirmed by double-blind, placebo-controlled food challenge in the Allergy Unit at the University Hospital Zürich, Switzerland [12,24], were used in the present study. All patients reacted to ingestion of soy with an OAS (oral allergy syndrome) involving symptoms like itching and swelling of the oral cavity. In addition, more systemic symptoms such as urticaria and throat tightness occurred in two patients. All patients had concomitant OAS to apple, cherry, carrot, celery (except patient I) and peanut.

Immunological characterization of Gly m 4

Histamine release

The histamine release assay was performed using the Histamine Enzyme Immunoassay kit (Immunotech, Marseilles, France) according to the manufacturer’s instructions. Whole peripheral blood cells of soybean allergic subjects were stimulated after dilution in histamine-release buffer (1:7) provided in the kit. Diluted blood (200 μl) was incubated with 100 μl of different dilutions of allergens (0.0001–100 μg/ml) in PBS for 30 min at 37°C. After centrifugation at 900 g for 5 min at 5°C, the released histamine was measured in the supernatant using the Histamine Enzyme Immunoassay kit. Total histamine was determined after cell lysis by freeze-thawing three times. For analysis of spontaneous histamine release, cells were incubated with PBS instead of diluted allergen.

Immunoblot inhibition

SDS/PAGE [25] was performed using a 15% separation and a 5% stacking gel. rGly m 4 was reduced by heating with dithiothreitol (Sigma–Aldrich, Deisenhofen, Germany), loaded on to the gel at a concentration of 1 μg/cm, and transferred on to 0.2 μm pore-sized nitrocellulose membranes by semi-dry blotting using a Novablot apparatus (Amersham Biosciences, Uppsala, Sweden). Remaining binding sites were blocked twice by incubation with...
RESULTS AND DISCUSSION

Allergenic activity of rGly m 4 evaluated by histamine release and immunoblot inhibition

Histamine release (maximum 23–69%) from blood cells after stimulation with rGly m 4 confirmed the biological activity of the recombinant allergen in 5/6 soy allergic subjects (Figure 2A). The similarity between the curves resulting from stimulation with soy isolate and those resulting from stimulation with rGly m 4 clearly indicated that Gly m 4 was the protein responsible for soy allergy in these patients and it proved the equivalence of rGly m 4 as used in the present study and authentic Gly m 4 occurring in soybean extract. Bet v 1 elicited histamine release at 10–100-fold lower protein concentrations than Gly m 4, indicating that Bet v 1 to be a more potent allergen in these patients. These findings are consistent with Bet v 1 being the primary sensitizing allergen in patients with birch pollen-associated soy allergy.

In immunoblot inhibition, Bet v 1 was found to be the most potent inhibitor of IgE binding to Gly m 4, followed by Gly m 4 and Mal d 1 (Figure 2B). Bet v 1 showed complete inhibition of IgE to Gly m 4 in 4/4 patients that were tested (Figure 2B), whereas Gly m 4 itself showed strong partial inhibition. This finding is consistent with the view that Bet v 1 is the sensitizing allergen. Mal d 1 showed similar strong partial inhibition in all patients. Pru av 1 inhibited IgE binding to Gly m 4 in 3/4 patients (II and III), whereas no inhibition was observed for Api g 1 and Dau c 1 in either patient tested. This difference may be explained by the substitution of the crucial Glu44/Glu45 in Gly m 4, Bet v 1, Mal d 1 and Pru av 1 for Lys44 in Api g 1 and Dau c 1 [26,27].

Solution structure of Gly m 4

Complete backbone amide resonance assignment and assignment of most side chain resonances [15] was the basis for the derivation of 2906 experimental restraints used for structure calculation (Table 1). The resulting ensemble of 20 structures (Figure 3A) is well defined with an overall backbone rmsd (root mean square deviation) of 0.60 Å and only insignificant violations of experimental restraints. Analysis of the stereochemical properties with PROCHECK 3.5.4 [21] confirms good stereochemical quality of the structure ensemble as mirrored by the fact that more than 90% of the amino acid residues are located in the most favoured regions of the Ramachandran plot.

The solution structure of Gly m 4 is composed of a seven-stranded antiparallel β-sheet and two short V-shaped α-helices α1 and α2 that pack together against the long C-terminal helix α3. The β-sheet consists of strands 2 → 11 (β1), 120 → 111 (β7), 97 → 105 (β6), 87 → 97 (β5), 62 → 74 (β4), 59 → 52 (β3) and 37 → 44 (β2). Three of the strands contain β-bulges, namely strand β2 at residues Phe37/Lys38 and Val43/Glu44, strand β4 at Glu71, and strand β5 at Val85. Strand β6 is interrupted at Leu103. The consequence of these irregularities of the β-strands is curving of the whole β-sheet around helix α3 that spans 24 residues from Gly129 to Ala152. Helices α1 and α2 comprise residues Pro13–Val23 and Ala26–Ala33 respectively (Figure 3B).

Table 1 Structure calculation statistics

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental restraints</td>
<td></td>
</tr>
<tr>
<td>Distance restraints</td>
<td></td>
</tr>
<tr>
<td>NOE distance restraints</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2481</td>
</tr>
<tr>
<td>Intraresidual</td>
<td>786</td>
</tr>
<tr>
<td>Sequential</td>
<td>674</td>
</tr>
<tr>
<td>Medium range (2 ≤ Δj ≤ 4)</td>
<td>331</td>
</tr>
<tr>
<td>Long range (Δj ≥ 5)</td>
<td>673</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>17</td>
</tr>
<tr>
<td>Hydrogen bonds*</td>
<td>24</td>
</tr>
<tr>
<td>Dihedral restraints</td>
<td></td>
</tr>
<tr>
<td>Scalar couplings [δ(J(HN,Hα))]</td>
<td>64</td>
</tr>
<tr>
<td>TALOS</td>
<td>228</td>
</tr>
<tr>
<td>Residual dipolar couplings (1Dōnj)</td>
<td>109</td>
</tr>
</tbody>
</table>

Properties of the accepted structure ensemble

- Restraint violation
  - Average distance restraint violation: 0.004 Å ± 0.001 Å
  - Maximum distance restraint violation: 0.24 Å
  - Average dihedral restraint violation: 0.47° ± 0.05°
  - Maximum dihedral restraint violation: 7.3°
  - Average RDC restraint violation: 0.14 Hz ± 0.02 Hz
  - Maximum RDC restraint violation: 1.35 Hz
- Deviation from ideal geometry
  - Bond length: 0.00050 Å ± 0.00005 Å
  - Bond angle: 0.108° ± 0.007°
- Co-ordinate precision
  - Overall rmsd: 0.60 Å
  - Backbone: 0.60 Å
  - Heavy atoms: 0.88 Å
  - Secondary structure rmsd amino acids: 3.3 Å
- Ramachandran plot statistics
  - Most favoured regions: 92.9%
  - Additional allowed regions: 7.1%
  - Generously allowed regions: 0.0%
  - Disallowed regions: 0.0%

*Two distance restraints per hydrogen bond (N_donor − O_acceptor ≤ 3.3 Å; HN donor − O_acceptor ≤ 2.3 Å).

0.3% Tween 20 in TBS (Tris-buffered saline: 50 mM Tris/HCl and 150 mM NaCl, pH 7.4). The nitrocellulose strips were incubated overnight with 130 μl of serum of four patients diluted to 1 ml in incubation buffer (TBS containing 0.05% Tween 20 and 0.1% BSA). For inhibition, strips were incubated with 130 μl of serum and additionally 20 μg of rGly m 4, recombinant Bet v 1, recombinant Api g 1, recombinant Dau c 1, recombinant Pru av 1 and recombinant Mal d 1 respectively.
Comparison with Bet v 1 and other homologues

The conformations of three allergenic proteins of the PR10 family are known: Bet v 1 [7], the major cherry allergen Pru av 1 [8] and the celery allergen Api g 1 [9]. Also known are three structures of Bet v 1 homologous proteins from yellow lupine, LIPR10.1A, LIPR10.1B and LIPR10.2A [28,29]. Gly m 4, Bet v 1 and its homologues feature identical secondary structure elements in identical order. As already evident from the homology model [15], the overall Gly m 4 fold is virtually identical with that of Bet v 1 and includes the hydrophobic cavity formed by the β-sheet and helix α3 that has been suggested to be important for the physiological function of the Bet v 1 family of proteins [8,30]. Superposition of the backbone of the actual experimentally determined Gly m 4 secondary structure elements, however, yields a surprisingly high rmsd of 2.60 Å for Bet v 1, 1.95 Å for Pru av 1 and 3.98 Å for Api g 1 (Figure 5A). The high rmsd values originate from differences in the connection of strands β3 and β4 in Gly m 4 as evidenced by the H-bond between the backbone amide proton of Glu62 and the backbone carboxyl group of Glu59. Such a β-hairpin conformation at the identical position was also identified in three highly Bet v 1 homologous proteins from yellow lupine with the same overall fold [28,29].

The long C-terminal helix α3 of Gly m 4 is bent as in the lupine proteins, contrasting with the straight helix α3 in Bet v 1. Whereas in the yellow lupine protein LIPR10.1A the N-terminal part of the helix is disordered, and in LIPR10.2A a kink of almost 60° is found near Phe142, the whole helix is well ordered in LIPR10.1B as it is in Gly m 4 [28,29]. The similarity of the helix conformation in LIPR10.1B and Gly m 4 is illustrated by the relatively low backbone rmsd of 0.77 Å, resulting from a comparison of helices α3 exclusively. The same comparison for Gly m 4 and Bet v 1, in contrast, yields an rmsd of 1.16 Å.

Helix α3 of Gly m 4 is shifted relative to the β-sheet by almost one turn along the helix axis as compared with Bet v 1. The most important contact points of the helix and the β-sheet are aromatic amino acids with side chains large enough to bridge
H. Berkner and others

is observed for the N-terminal part of β allergens (Figures 5A and 5B).

and Gly m 4 as compared with superposition with the Bet v 1 results in a comparatively low backbone rmsd value of 1.72 Å for lupine proteins LlPR10.1A, LlPR10.1B and LlPR10.2A. This helix relative to the Phe3/Phe5 both pack against Leu132. A corresponding shift of the helix in Gly m 4; backbone rmsd = 0.6 Å. (B) Cartoon of the average structure of Gly m 4. The bent and shifted C-terminal helix sets this structure slightly apart from the strongly related Bet v 1 structure.

Whereas this glutamate is not conserved in Api g 1 and Dau c 1, to be important for IgE binding to Bet v 1 and Pru av 1 [8,31,32].

The second potential cross-reactive epitope consists of residues located in the two short helices α1 and α2 and the C-terminus of the protein including parts of helix α3. These residues also form a coherent surface area of ~1100 Å², large enough to serve as an IgE-binding site. The structural similarity of Gly m 4 and Bet v 1 in this region is not as high as that in the

Links between immunological cross-reactivity and structural properties

Characterizing individual and cross-reactive IgE epitopes of allergens is an important prerequisite for understanding the molecular basis of the clinical phenomenon of pollen-related food allergy and for the design of hypo-allergenic proteins for a safer allergen-specific immunotherapy. A number of studies using techniques like mutational epitope analysis and mimotope selection have been carried out to localize putative epitopes on the surface of Bet v 1 and its homologous tree pollen and food allergens. Surface patches conserved with respect to conformation and charge distribution on homologous allergens are prime candidates for cross-reactive epitopes.

The histamine release experiments revealed a drastically higher allergenic activity of Bet v 1 in comparison with Gly m 4. Consistent with this finding, immunoblot inhibition showed that Bet v 1 had the strongest capacity to inhibit IgE binding to Gly m 4 (i.e. cross-reactivity), probably because Bet v 1 is the sensitizing allergen in pollen-related soy allergy, that is IgE antibodies are primarily produced against the major birch pollen allergen. Mal d 1 also showed significant inhibitory capacity in the immunoblotting experiments, followed by Pru av 1 (Figure 2). No cross-reactivity was observed for Api g 1 and Dau c 1. Thus we searched for surface patches on Gly m 4 consisting of amino acids conserved either only in Bet v 1 or additionally in Mal d 1 and Pru av 1 (Figures 6A and 6B). With this strategy, we could identify three compact conserved surface patches as well as one less coherent patch on Gly m 4 as candidate epitopes (Figure 6C).

One patch consists of the continuous stretch of amino acids Ser10–Ile35, corresponding to the area around the P-loop characteristic for the Bet v 1 protein family. This area spans a surface of ~1000 Å² in Gly m 4 and has already been identified as a potential cross-reactive epitope in Bet v 1 and Pru av 1 by site-directed mutagenesis [31–33]. The crystal structure of an IgG antibody fragment inhibiting IgE binding in complex with Bet v 1 supports these findings [34]. The glutamate preceding the loop has turned out to be especially important for IgE binding, as mutation of this residue strongly influences the IgE binding capacity of Bet v 1 as well as that of Pru av 1. The conformation of this loop in Gly m 4 is virtually identical with the Bet v 1/Mal d 1 loops but contrasts with the Api g 1/Dau c 1 loops, the latter containing a positively charged lysine residue in the position corresponding to the glutamate. Thus the P-loop region is not only potentially responsible for cross-reactivity of Bet v 1 and Pru av 1, but it may also represent a cross-reactive epitope of Gly m 4 and probably of Mal d 1.

The second potential cross-reactive epitope consists of residues also present in Gly m 4, and its side chain points in the same direction as in Bet v 1 and LlPR10.1B (Figure 5C).
P-loop region. Due to the shift of helix α3 along its axis, the distances of the amino acids located in helices α1 and α2 to those in α3 differ slightly from the corresponding distances in Bet v 1, but these differences do not impair the general similarity of the protein surfaces. Helix α1 was identified as part of an IgE epitope of Bet v 1 by mimotope selection [2], and Asn28 located in helix α2 was shown to be crucial for IgE binding to Pru av 1 [35]. These findings strongly support the notion that the conserved surface patch represents another cross-reactive epitope on Gly m 4.

The third conserved surface patch consists of residues located in strands β1, β7, β6, and adjacent loops β5–β6/β7–α3. It covers a coherent area of ~1100 Å². The spatial arrangement of these β-strands is highly similar in Bet v 1 and Gly m 4. This putative cross-reactive epitope is probably the one showing the highest flexibility, as it contains a number of residues located in the extended loop β7–α3. The respective loop in Bet v 1 contains Asp125, which seems to play a role in IgE binding, as its mutation modulates IgE reactivity [3].

These surface areas largely overlap with the three surface patches conserved among Bet v 1 and other tree pollen allergens from the order fagales like Aln g 1, Cor a 1 or Car b 1. These potential cross-reactive epitopes [7] are, probably, not only responsible for the pollen-food allergy syndrome, but also for cross-reactivity among tree pollen.

In addition to these three highly coherent epitopes, a fourth area of conserved residues covering ~1050 Å² is found in Gly m 4. It includes residues located in strands β1, β7 and β6 as well as in loops β4–β5 and β6–β7. Orientation of the β-strands and the conformation of the loops are identical in Gly m 4 and Bet v 1. Two sequence stretches in Bet v 1 identified by mimotope selection both include residues from this area [36], and a six-point mutant of Bet v 1 including mutation of Ser112, which corresponds to the N-terminal serine residue of β7 in Gly m 4, showed drastically reduced IgE binding [3]. S112C single mutants of Bet v 1 and Pru av 1 and a P108G mutant of Bet v 1 also showed reduced IgE reactivity [31,33].

From clinical studies, Gly m 4 [11] as well as Api g 1 [37] causes more severe symptoms than, for example, Pru av 1 [38] or Cor a 1 [39]. As the unfolding temperatures of Bet v 1, Gly m 4 and Pru av 1 are all in the range of 65–75°C as determined by CD spectroscopy (H. Berkner, unpublished work), differential stability is highly likely not to be the major cause of the differences in the severity of symptoms.

**Conclusion**

Cross-reactive epitopes on homologous allergens result from a highly similar tertiary structure in combination with local or extended sequence similarity. We show here that, apart from a few slight conformational differences, the three-dimensional structure of Gly m 4 is virtually identical with that of the major birch pollen allergen Bet v 1. Generally, our results also confirm that the grade of total sequence similarity correlates with the grade of cross-reactivity, because Bet v 1, Mal d 1 and Pru av 1, with sequence similarities of 66–69% to Gly m 4, show significant cross-reactivity, whereas Api g 1 and Dau c 1, with sequence similarity of 58 and 56% to Gly m 4, did not exhibit any cross-reactivity with the soybean allergen. Sequence similarity, however, is not able to predict cross-reactivity in detail: Bet v 1, the protein with the lowest sequence similarity to Gly m 4 in the first group, shows the highest cross-reactivity, whereas the protein with the highest sequence similarity, Pru av 1, even fails to inhibit Gly m 4 in some of the studied patients.
In the context of our findings that many of the structural features of Gly m 4 are particularly similar to the PR10 proteins of yellow lupine, it would be interesting to investigate general allergenicity and cross-reactivity of Gly m 4 with these highly homologous proteins (sequence similarities of 74–84%). In fact, while lupine has been identified as a food allergen source [40]. Immunoblotting as well as an in silico approach [41,42] supports the notion that a PR10 protein of white lupine with sequence similarities of 75–94% to the yellow lupine proteins is responsible for lupine allergy.

Our results also demonstrate the importance of considering patient-specificity of the IgE reaction for the design of hypoallergenic proteins for immunotherapy. The inhibitory potential of Pru av 1 depends on the specific patients tested, suggesting that these patients recognize different epitopes [2]. Further efforts for a detailed characterization of the conformational IgE epitopes of relevant allergens are thus necessary to understand the clinical complexity of the phenomenon at the molecular level and to pave the way for a safe and patient-tailored immunotherapy.

ACKNOWLEDGEMENTS
We thank Nadine Herz and Ramona Heissmann for expert technical assistance and Katrin Lehmann (Sanofi-Aventis, Brüningstr. 50, 65928, Frankfurt, Germany) for her contributions in the early stages of this work.

FUNDING
This work was supported by the Deutsche Forschungsgemeinschaft [grant number RO 617/11-4 (to P R.), by a doctoral
fellowship awarded to P. N. from the Fonds des Verbandes der Chemischen Industrie in co-operation with the Bundesministerium für Bildung und Forschung, and by the Elitenetwork of Bavaria.

REFERENCES

23 Reference deleted