

## 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\*, Philipp Neudecker<sup>†\*</sup>, Diana Mittag<sup>§†</sup>, Barbara K. Ballmer-Weber<sup>‡</sup>, Kristian Schweimer, Stefan Vieths<sup>§§</sup>, Paul Rösch<sup>1</sup>

\*Both authors contributed equally

Department of Biopolymers and Research Center for Bio-Macromolecules,  
Universität Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany  
roesch@unibt.de

<sup>†</sup>present address: Departments of Molecular Genetics, Biochemistry, and Chemistry,  
University of Toronto, 1 King's College Circle, Toronto, ON, M5S 1A8, Canada

<sup>‡</sup>University Hospital Zurich, Department of Dermatology, Allergy Unit, Gloriastr. 31,  
8091 Zurich, Switzerland

<sup>§</sup>present address: The Walter and Eliza Hall Institute of Medical Research, 1G Royal  
Parade, Parkville Vic 3050, Australia

<sup>§§</sup>Department of Allergology, Paul-Ehrlich-Institut, Paul-Ehrlich-Straße 51-59, 63225  
Langen, Germany

<sup>1</sup>To whom correspondence should be addressed

page heading title: Soybean Gly m 4 allergen structure

Key words: allergens; protein structure; NMR; soybean; epitope; histamine release

Abbreviations: IPAP, inphase antiphase; OAS, oral allergy syndrome; PBS, phosphate buffered saline; PR10, class 10 of pathogenesis related proteins; RDC, residual dipolar coupling; rmsd, root mean square deviation; TBS, Tris buffered saline.

Accepted Manuscript

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BSR20080117

## Synopsis

In many cases, patients allergic to birch pollen also show allergic reactions after ingestion of certain fruits or vegetables. This observation is explained on 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 structure 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 as we as, in combination with immunological data, allow to propose surface patches that might represent cross-reactive epitopes.

Accepted Manuscript

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BSR20080117

## 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 about 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 1 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 to 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 on the molecular level.

According to their sequence similarities and related joined characteristics - molecular mass around 17 kDa, acidic isoelectric point around 6 - these proteins are members of the Bet v 1 protein superfamily and the family of class 10 of pathogenesis related (PR) 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 LIPR10.2B serves as a cytokinin reservoir [5] and with a structural study suggesting the norcochlorogenic acid synthase to be a true member of the PR10 family [6]. Clearly, detailed structural studies of the PR10 proteins may contribute to unravel their relations and thus add to the knowledge about functional features of this class of proteins.

The notion that 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] do not only comprise identical arrangements of secondary structure elements, they are virtually identical, explaining the phenomenon of cross-reactivity on a molecular level. Detailed analysis of common features and differences regarding primary and tertiary structure of Bet v 1 homologous allergens can thus help 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 homologs, 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 on 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, for

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 to 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 structural analysis of Gly m 4 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.

Accepted Manuscript

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BSR20080117

## Experimental

### Structure determination of Gly m 4

*Sample preparation:* Recombinant Gly m 4 was expressed, purified, and stored as lyophilised powder as described [15]. NMR samples were prepared by dissolving either  $^{15}\text{N}$ - or  $^{15}\text{N},^{13}\text{C}$ - labelled Gly m 4 in 10 – 50 mM potassium phosphate buffer pH 7.0, 10 %  $\text{D}_2\text{O}$ , and 0.02 % sodium azide. For specific measurement of NOEs involving aromatic protons, unlabelled Gly m 4 was dissolved in 50 mM potassium phosphate buffer pH 7.0 in  $\text{D}_2\text{O}$ . 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: 2D  $[^1\text{H},^1\text{H}]$ -NOESY using unlabelled Gly m 4 in  $\text{D}_2\text{O}$ ,  $^{15}\text{N}$ -NOESYHSQC,  $^{13}\text{C}$ -NOESYHSQC with 120 ms mixing time, and  $^{15}\text{N},^{15}\text{N}$ -HMQC-NOESYHSQC,  $^{13}\text{C},^{15}\text{N}$ -HMQC-NOESYHSQC with 150 ms mixing time. Slowly exchanging amide protons involved in H-bond formation were identified from a  $^{15}\text{N}$ -HSQC-spectrum recorded 1 h after the lyophilised protein was dissolved in  $\text{D}_2\text{O}$ .  $^3\text{JHNH}\alpha$  scalar coupling constants were derived from the HNHA experiment.  $^1\text{D}_{\text{HN}}$  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 experiments.  $\{^1\text{H}\}^{15}\text{N}$  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 to distance restraints with upper limits of 3.0 Å (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 3.3 Å between acceptor and amide nitrogen. Dihedral restraints for  $\Phi$  angles were derived from  $^3\text{JHNH}\alpha$  scalar coupling constants including a  $\Phi$  angle between  $-80^\circ$  and  $-40^\circ$  for coupling constants smaller than 6.0 Hz, while values greater than 8.0 Hz were converted to a  $\Phi$  angle between  $-160^\circ$  and  $-80^\circ$ . Additional dihedral restraints for  $\Phi$  and  $\Psi$  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 residual dipolar couplings (RDCs) with a harmonic potential. RDCs measured for amino acids located in areas flexible on the pico- to nanosecond time scale ( $\{^1\text{H}\}^{15}\text{N}$  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 XPLOR [20], PROCHECK 3.5.4 [21] and MolMol [22] and visualized with Pymol (2006, DeLano Scientific LLC, South San Francisco, California, USA). Surface areas of putative cross-reactive IgE epitopes were calculated using the software nACCESS with a probe radius of 2.5 Å [23]. Coordinates were deposited in the protein data bank (PDB, accession code 2K7H).

## Soybean allergic patients

Sera of six patients (age  $22 \pm 5$ , 5 female, 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 this study. All patients reacted to ingestion of soy with an oral allergy syndrome (OAS) involving symptoms like itching and swelling of the oral cavity. In addition, more systemic symptoms such as urticaria and throat tightness occurred in 2 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. 200  $\mu$ l diluted blood were incubated with 100  $\mu$ l of different dilutions of allergens (0.0001 – 100  $\mu$ g/ml) in PBS for 30 min at 37 °C. After centrifugation for 5 min at 900xg and 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 3 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 1,4-Dithiothreitol (Sigma-Aldrich, Deisenhofen, Germany), loaded onto the gel at a concentration of 1  $\mu$ g/cm, and transferred onto 0.2  $\mu$ m pore-sized nitrocellulose membranes by semi-dry blotting using a Novablot apparatus (Amersham Pharmacia Biotech, Uppsala, Sweden). Remaining binding sites were blocked twice by incubation with 0.3 % Tween 20 in Tris-buffered saline (TBS, 50 mM Tris/HCl, 150 mM NaCl, pH 7.4). The nitrocellulose strips were incubated overnight with 130  $\mu$ l 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  $\mu$ l of serum and additionally 20  $\mu$ g of rGly m 4, rBet v 1, rApi g 1, rDau c 1, rPru av 1 and rMal d 1, respectively.

## 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 of the curves resulting from stimulation with soy isolate and 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 this study and authentic Gly m 4 occurring in soybean extract. Bet v 1 elicited histamine release at 10 to 100-fold lower protein concentrations than Gly m 4 indicating 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) while 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 2/4 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 Glu44/45 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 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 anti-parallel  $\beta$ -sheet and two short V-shaped  $\alpha$ -helices  $\alpha$ 1 and  $\alpha$ 2 that pack together against the long carboxyterminal helix  $\alpha$ 3. The  $\beta$ -sheet consists of strands 2  $\rightarrow$  11 ( $\beta$ 1), 120  $\leftarrow$  111 ( $\beta$ 7), 97  $\rightarrow$  105 ( $\beta$ 6), 87  $\leftarrow$  79 ( $\beta$ 5), 62  $\rightarrow$  74 ( $\beta$ 4), 59  $\leftarrow$  52 ( $\beta$ 3), and 37  $\rightarrow$  44 ( $\beta$ 2). Three of the strands contain  $\beta$ -bulges, namely strand  $\beta$ 2 at residues Phe37/Lys38 and Val43/Glu44, strand  $\beta$ 4 at Glu71, and strand  $\beta$ 5 at Val85. Strand  $\beta$ 6 is interrupted at Leu103. The consequence of these irregularities of the  $\beta$ -strands is curving of the whole  $\beta$ -sheet around helix  $\alpha$ 3 that spans 24 residues from Gln129 to Ala152. Helices  $\alpha$ 1 and  $\alpha$ 2 comprise residues Pro15-Val23 and Ala26-Ala33, respectively (Figure 3B).

*Comparison with Bet v 1 and other homologs:* 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 homologs 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 to that of Bet v 1 and includes the hydrophobic cavity formed by the  $\beta$ -sheet and helix  $\alpha$ 3 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 rmsds originate from differences in the connection of strands  $\beta$ 3 and  $\beta$ 4, in helix  $\alpha$ 3 and the position of helix  $\alpha$ 3 relative to the  $\beta$ -sheet.

Whereas amino acids Phe58 to Lys65 form a flexible loop in Bet v 1, a short type II  $\beta$ -turn, Glu59 to Glu62, connects the extended strands  $\beta$ 3 and  $\beta$ 4 in Gly m 4 as evidenced by the H-bond between the backbone amide proton of Glu62 and the backbone carbonyl group of Glu59. Such  $\beta$ -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  $\alpha$ 3 of Gly m 4 is bent as in the lupine proteins, contrasting the straight helix  $\alpha$ 3 in Bet v 1. While 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  $\alpha$ 3 exclusively. The same comparison for Gly m 4 and Bet v 1, in contrast, yields an rmsd of 1.16 Å.

Helix  $\alpha 3$  of Gly m 4 is shifted relative to the  $\beta$ -sheet by almost one turn along the helix axis as compared to Bet v 1. The most important contact points of the helix and the  $\beta$ -sheet are aromatic amino acids with side chains large enough to bridge these structural elements. For example, Phe143 contacts Ser11, and Phe3/ Phe5 both pack against Leu132. A corresponding shift of the helix relative to the  $\beta$ -sheet is also observed in the three yellow lupine proteins LIPR10.1A, LIPR10.1B, and LIPR10.2A. This results in a comparatively low backbone rmsd value of 1.72 Å for superposition of all secondary structure elements of LIPR10.1B and Gly m 4 as compared to superposition with the Bet v 1 allergens (Figure 5A,B).

The loop connecting  $\alpha 2$  and  $\beta 2$  is shorter in Gly m 4 than it is in Bet v 1, thus adopting a different conformation. Flexibility in terms of high frequency motion on the picosecond to nanosecond timescale as indicated by reduced  $\{^1\text{H}\}^{15}\text{N}$  NOE values is observed for the amino terminal part of  $\beta 7$  near the glycine-rich loop  $\beta 6$ - $\beta 7$  and for the extended loop connecting strand  $\beta 7$  with helix  $\alpha 3$  (Figure 4). The latter was also found to be flexible in Pru av 1 [8]. The Gly-rich loop between  $\beta 2$  and  $\beta 3$  with the consensus sequence Gly-X-Gly-Gly-X-Gly-Thr ( $\gg$ P-loop $\ll$ ) is highly conserved in the Bet v 1 homologous allergens and the lupine proteins (Figure 1). It adopts the same conformation in all these proteins as reflected by a very low rmsd of 0.52 Å for the alignment of all heavy atoms of the P-loop in Bet v 1 and Gly m 4 (Figure 5C). This loop and the neighbouring amino acids, especially the preceding Glu44/45, have been reported to be important for IgE-binding to Bet v 1 and Pru av 1 [8,31,32]. While this glutamate is not conserved in Api g 1 and Dau c 1, it is also present in Gly m 4, and its side chain points in the same direction as in Bet v 1 and LIPR10.1B (Figure 5C).

*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 (Figure 6A,B). 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 Ser39 – Ile55, corresponding to the area around the P-loop characteristic for the Bet v 1 protein family. This area spans a surface of  $\sim 1000 \text{ \AA}^2$  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]. Especially the glutamate preceding the loop has turned out to be 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 to the Bet v 1/ Mal d 1 loops but contrasts the Api g 1/ Dau c 1 loops, the latter containing a positively charged lysine 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 located in the two short helices  $\alpha 1$  and  $\alpha 2$  and the carboxy terminus of the protein including parts of helix  $\alpha 3$ . These residues also form a coherent surface area of  $\sim 1100 \text{ \AA}^2$ , 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 P-loop region. Due to the shift of helix  $\alpha 3$  along its axis, the distances of the amino acids located in helices  $\alpha 1$  and  $\alpha 2$  to those in  $\alpha 3$  differ slightly from the corresponding distances in Bet v 1, but these differences do not impair the general similarity of the protein surfaces. Helix  $\alpha 1$  was identified as part of an IgE-epitope of Bet v 1 by mimotope selection [2], and Asn28 located in helix  $\alpha 2$  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  $\beta 1$ ,  $\beta 7$ ,  $\beta 6$ ,  $\beta 5$  and adjacent loops  $\beta 5$ - $\beta 6$ /  $\beta 7$ - $\alpha 3$ . It covers a coherent area of  $\sim 1100 \text{ \AA}^2$ . The spatial arrangement of these  $\beta$ -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  $\beta 7$ - $\alpha 3$ . 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 likely 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  $\sim 1050 \text{ \AA}^2$  is found in Gly m 4. It includes residues located in strands  $\beta 1$ ,  $\beta 7$ , and  $\beta 6$  as well as in loops  $\beta 4$ - $\beta 5$  and  $\beta 6$ - $\beta 7$ . Orientation of the  $\beta$ -strands and the conformation of the loops is 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 amino terminal serine of  $\beta 7$  in Gly m 4 showed drastically reduced IgE-binding [3]. Ser112Cys single mutants of Bet v 1 and Pru av 1 and a Pro108Gly 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] cause more severe symptoms than, for example, Pru av 1 [38] or Cor a 1 [39]. As the unfolding temperature of Bet v 1, Gly m 4, and Pru av 1 are all in the range of  $65 \text{ }^\circ\text{C}$  to  $75 \text{ }^\circ\text{C}$  as determined by CD spectroscopy (Berkner, unpublished), differential stability is highly likely not the major cause for 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 to 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, while 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, white lupine has been identified as a food allergen source [40]. Immunoblotting as well as an *in silico* approach [41,42] support 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 hypo-allergenic 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.

## Acknowledgement

We would like to thank Nadine Herz and Ramona Heissmann for expert technical assistance and Katrin Lehmann for her contributions in the early stages of this project.

## Funding

This project was supported by a grant from the Deutsche Forschungsgemeinschaft to PR (RO 617/11-4) and a doctoral fellowship to PN from the Fonds des Verbandes der Chemischen Industrie in cooperation with the Bundesministerium für Bildung und Forschung (BMBF). The project was supported by the Elitenetwork of Bavaria.

## References

- 1 Vieths, S., Scheurer, S. and Ballmer-Weber, B. (2002) Current understanding of cross-reactivity of food allergens and pollen. *Ann. N. Y. Acad. Sci.* 964, 47-68
- 2 Mittag, D., Batori, V., Neudecker, P., Wiche, R., Friis, E. P., Ballmer-Weber, B. K., Vieths, S. and Roggen, E. L. (2006) A novel approach for investigation of specific and cross-reactive IgE epitopes on bet v 1 and homologous food allergens in individual patients. *Mol. Immunol.* 43, 268-278. doi:10.1016/j.molimm.2005.02.008
- 3 Ferreira, F., Ebner, C., Kramer, B., Casari, G., Briza, P., Kungl, A. J., Grimm, R., Jahn-Schmid, B., Breiteneder, H., Kraft, D., Breitenbach, M., Rheinberger, H. J. and Scheiner, O. (1998) Modulation of IgE reactivity of allergens by site-directed mutagenesis: Potential use of hypoallergenic variants for immunotherapy. *FASEB J.* 12, 231-242
- 4 Murzin, A. G., Brenner, S. E., Hubbard, T. and Chothia, C. (1995) SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247, 536-540. doi:10.1006/jmbi.1995.0159
- 5 Fernandes, H., Pasternak, O., Bujacz, G., Bujacz, A., Sikorski, M. M. and Jaskolski, M. (2008) Lupinus luteus pathogenesis-related protein as a reservoir for cytokinin. *J. Mol. Biol.* 378, 1040-1051. doi:10.1016/j.jmb.2008.03.027
- 6 Berkner, H., Schweimer, K., Matecko, I. and Rosch, P. (2008) Conformation, catalytic site, and enzymatic mechanism of the PR10 allergen-related enzyme norcochlorogenic acid synthase. *Biochem. J.* 413, 281-290. doi:10.1042/BJ20080306
- 7 Gajhede, M., Osmark, P., Poulsen, F. M., Ipsen, H., Larsen, J. N., Joost van Neerven, R. J., Schou, C., Lowenstein, H. and Spangfort, M. D. (1996) X-ray and NMR structure of bet v 1, the origin of birch pollen allergy. *Nat. Struct. Biol.* 3, 1040-1045
- 8 Neudecker, P., Schweimer, K., Nerkamp, J., Scheurer, S., Vieths, S., Sticht, H. and Rosch, P. (2001) Allergic cross-reactivity made visible: Solution structure of the major cherry allergen pru av 1. *J. Biol. Chem.* 276, 22756-22763. doi:10.1074/jbc.M101657200 [doi]; M101657200 [pii]
- 9 Schirmer, T., Hoffmann-Sommergrube, K., Susani, M., Breiteneder, H. and Markovic-Housley, Z. (2005) Crystal structure of the major celery allergen api g 1: Molecular analysis of cross-reactivity. *J. Mol. Biol.* 351, 1101-1109. doi:10.1016/j.jmb.2005.06.054
- 10 Crowell, D. N., John, M. E., Russell, D. and Amasino, R. M. (1992) Characterization of a stress-induced, developmentally regulated gene family from soybean. *Plant Mol. Biol.* 18, 459-466
- 11 Kleine-Tebbe, J., Vogel, L., Crowell, D. N., Haustein, U. F. and Vieths, S. (2002) Severe oral allergy syndrome and anaphylactic reactions caused by a bet v 1-related PR-10 protein in soybean, SAM22. *J. Allergy Clin. Immunol.* 110, 797-804
- 12 Mittag, D., Vieths, S., Vogel, L., Becker, W. M., Rihs, H. P., Helbling, A., Wuthrich, B. and Ballmer-Weber, B. K. (2004) Soybean allergy in patients allergic to birch pollen: Clinical investigation and molecular characterization of allergens. *J. Allergy Clin. Immunol.* 113, 148-154. doi:10.1016/j.jaci.2003.09.030

- 13 King, T. P., Hoffman, D., Lowenstein, H., Marsh, D. G., Platts-Mills, T. A. and Thomas, W. (1995) Allergen nomenclature. *Allergy*. 50, 765-774
- 14 Ballmer-Weber, B. K. and Vieths, S. (2008) Soy allergy in perspective. *Curr. Opin. Allergy Clin. Immunol.* 8, 270-275. doi:10.1097/ACI.0b013e3282ffb157
- 15 Neudecker, P., Lehmann, K. and Rosch, P. (2003) Sequence-specific  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  resonance assignments of SAM22, an allergenic stress-induced protein from soy bean. *J. Biomol. NMR.* 26, 191-192. doi:5123893 [pii]
- 16 Cavanagh, J., Fairbrother, W. J., III, P. A. G. and Skelton, N. J. (2006) *Protein NMR spectroscopy: Principles and practice.*, Academic Press, San Diego, CA
- 17 Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D. and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed src homology 2 domain studied by  $^{15}\text{N}$  NMR relaxation. *Biochemistry.* 33, 5984-6003
- 18 Johnson, B. A. and Blevins, R. A. (1994) NMRView: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR.* 4, 603-14
- 19 Cornilescu, G., Delaglio, F. and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR.* 13, 289-302
- 20 Tjandra, N., Omichinski, J. G., Gronenborn, A. M., Clore, G. M. and Bax, A. (1997) Use of dipolar  $^1\text{H}$ - $^{15}\text{N}$  and  $^1\text{H}$ - $^{13}\text{C}$  couplings in the structure determination of magnetically oriented macromolecules in solution. *Nat. Struct. Biol.* 4, 732-738
- 21 Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR.* 8, 477-486
- 22 Koradi, R., Billeter, M. and Wuthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graph.* 14, 51-5, 29-32. doi:0263785596000094 [pii]
- 23 Hubbard, S. J. and Thornton, J. M. (1993) NACCESS.
- 24 Ballmer-Weber, B. K., Holzhauser, T., Scibilia, J., Mittag, D., Zisa, G., Ortolani, C., Oesterballe, M., Poulsen, L. K., Vieths, S. and Bindslev-Jensen, C. (2007) Clinical characteristics of soybean allergy in europe: A double-blind, placebo-controlled food challenge study. *J. Allergy Clin. Immunol.* 119, 1489-1496. doi:10.1016/j.jaci.2007.01.049
- 25 Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* 227, 680-685
- 26 Neudecker, P., Lehmann, K., Nerkamp, J., Haase, T., Wangorsch, A., Fotisch, K., Hoffmann, S., Rosch, P., Vieths, S. and Scheurer, S. (2003) Mutational epitope analysis of pru av 1 and api g 1, the major allergens of cherry (*prunus avium*) and celery (*apium graveolens*): Correlating IgE reactivity with three-dimensional structure. *Biochem. J.* 376, 97-107
- 27 Wangorsch, A., Ballmer-Weber, B. K., Rosch, P., Holzhauser, T. and Vieths, S. (2007) Mutational epitope analysis and cross-reactivity of two isoforms of api g 1, the major celery allergen. *Mol. Immunol.* 44, 2518-2527. doi:10.1016/j.molimm.2006.12.023

- 28 Biesiadka, J., Bujacz, G., Sikorski, M. M. and Jaskolski, M. (2002) Crystal structures of two homologous pathogenesis-related proteins from yellow lupine. *J. Mol. Biol.* 319, 1223-1234. doi:10.1016/S0022-2836(02)00385-6
- 29 Pasternak, O., Biesiadka, J., Dolot, R., Handschuh, L., Bujacz, G., Sikorski, M. M. and Jaskolski, M. (2005) Structure of a yellow lupin pathogenesis-related PR-10 protein belonging to a novel subclass. *Acta Crystallogr. D Biol. Crystallogr.* 61, 99-107. doi:10.1107/S0907444904028173
- 30 Markovic-Housley, Z., Degano, M., Lamba, D., von Roepenack-Lahaye, E., Clemens, S., Susani, M., Ferreira, F., Scheiner, O. and Breiteneder, H. (2003) Crystal structure of a hypoallergenic isoform of the major birch pollen allergen bet v 1 and its likely biological function as a plant steroid carrier. *J. Mol. Biol.* 325, 123-133
- 31 Spangfort, M. D., Mirza, O., Ipsen, H., Joost van Neerven, R. J., Gajhede, M. and Larsen, J. N. (2003) Dominating IgE-binding epitope of bet v 1, the major allergen of birch pollen, characterized by X-ray crystallography and site-directed mutagenesis. *J. Immunol.* 171, 3084-3090
- 32 Neudecker, P., Lehmann, K., Nerkamp, J., Haase, T., Wangorsch, A., Fotisch, K., Hoffmann, S., Rosch, P., Vieths, S. and Scheurer, S. (2003) Mutational epitope analysis of pru av 1 and api g 1, the major allergens of cherry (*prunus avium*) and celery (*apium graveolens*): Correlating IgE reactivity with three-dimensional structure. *Biochem. J.* 376, 97-107. doi:10.1042/BJ20031057 [doi]; BJ20031057 [pii]
- 33 Scheurer, S., Son, D. Y., Boehm, M., Karamloo, F., Franke, S., Hoffmann, A., Haustein, D. and Vieths, S. (1999) Cross-reactivity and epitope analysis of pru a 1, the major cherry allergen. *Mol. Immunol.* 36, 155-167
- 34 Mirza, O., Henriksen, A., Ipsen, H., Larsen, J. N., Wissenbach, M., Spangfort, M. D. and Gajhede, M. (2000) Dominant epitopes and allergic cross-reactivity: Complex formation between a fab fragment of a monoclonal murine IgG antibody and the major allergen from birch pollen bet v 1. *J. Immunol.* 165, 331-338
- 35 Wiche, R., Gubesch, M., Konig, H., Fotisch, K., Hoffmann, A., Wangorsch, A., Scheurer, S. and Vieths, S. (2005) Molecular basis of pollen-related food allergy: Identification of a second cross-reactive IgE epitope on pru av 1, the major cherry (*prunus avium*) allergen. *Biochem. J.* 385, 319-327. doi:10.1042/BJ20040842
- 36 Ganglberger, E., Grunberger, K., Sponer, B., Radauer, C., Breiteneder, H., Boltz-Nitulescu, G., Scheiner, O. and Jensen-Jarolim, E. (2000) Allergen mimotopes for 3-dimensional epitope search and induction of antibodies inhibiting human IgE. *FASEB J.* 14, 2177-2184. doi:10.1096/fj.99-1000com
- 37 Luttkopf, D., Ballmer-Weber, B. K., Wuthrich, B. and Vieths, S. (2000) Celery allergens in patients with positive double-blind placebo-controlled food challenge. *J. Allergy Clin. Immunol.* 106, 390-399. doi:10.1067/mai.2000.108711
- 38 Reuter, A., Lidholm, J., Andersson, K., Ostling, J., Lundberg, M., Scheurer, S., Enrique, E., Cistero-Bahima, A., San Miguel-Moncin, M., Ballmer-Weber, B. K. and Vieths, S. (2006) A critical assessment of allergen component-based in vitro diagnosis in cherry allergy across europe. *Clin. Exp. Allergy.* 36, 815-823. doi:10.1111/j.1365-2222.2006.2492.x
- 39 Luttkopf, D., Muller, U., Skov, P. S., Ballmer-Weber, B. K., Wuthrich, B., Skamstrup Hansen, K., Poulsen, L. K., Kastner, M., Haustein, D. and Vieths, S.

(2002) Comparison of four variants of a major allergen in hazelnut (*corylus avellana*) cor a 1.04 with the major hazel pollen allergen cor a 1.01. *Mol. Immunol.* 38, 515-525

40 Peeters, K. A., Nordlee, J. A., Penninks, A. H., Chen, L., Goodman, R. E., Buijnzeel-Koomen, C. A., Hefle, S. L., Taylor, S. L. and Knulst, A. C. (2007) Lupine allergy: Not simply cross-reactivity with peanut or soy. *J. Allergy Clin. Immunol.* 120, 647-653. doi:10.1016/j.jaci.2007.05.032

41 Faeste, C. K., Lovik, M., Wiker, H. G. and Egaas, E. (2004) A case of peanut cross-allergy to lupine flour in a hot dog bread. *Int. Arch. Allergy Immunol.* 135, 36-39. doi:10.1159/000080041

42 Guarneri, F., Guarneri, C. and Benvenga, S. (2005) Identification of potentially cross-reactive peanut-lupine proteins by computer-assisted search for amino acid sequence homology. *Int. Arch. Allergy Immunol.* 138, 273-277. doi:10.1159/000088864

43 Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. and Thompson, J. D. (2003) Multiple sequence alignment with the clustal series of programs. *Nucleic Acids Res.* 31, 3497-3500

Accepted Manuscript

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BSR20080117

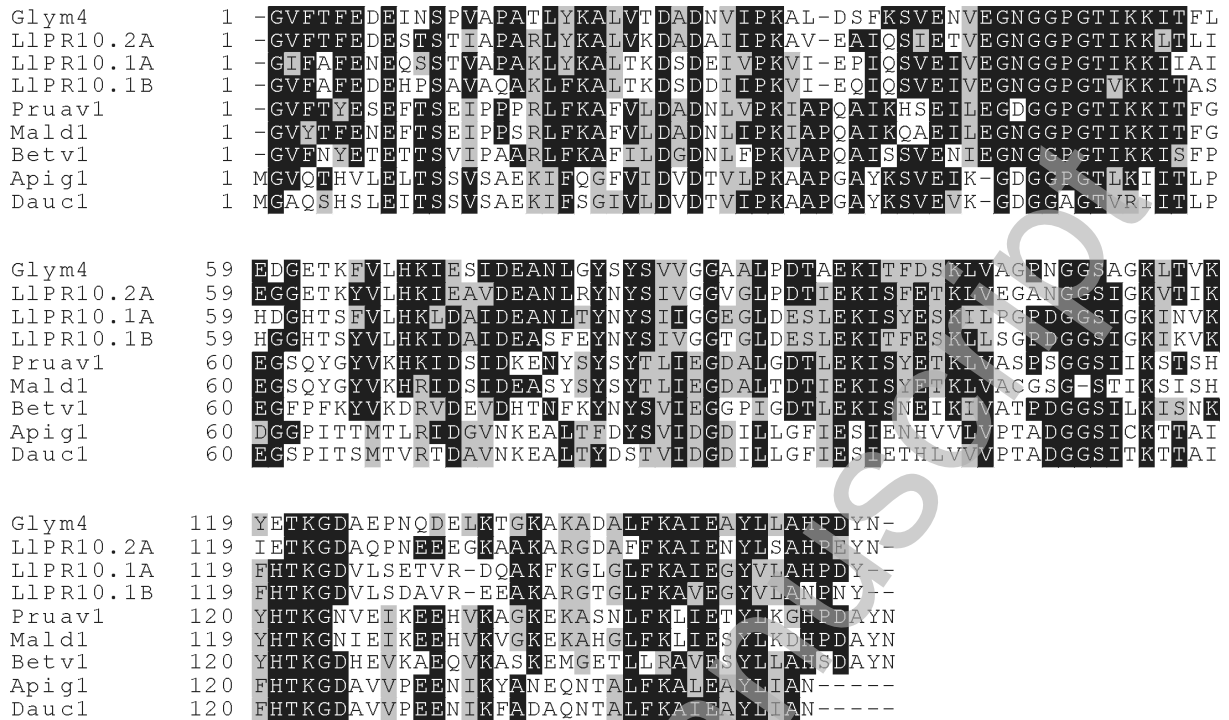
**Table 1:** Structure calculation statistics

**Experimental restraints**

Distance restraints			
NOE distance restraints	Total		2481
	Intraresidual		786
	Sequential		674
	Medium range ( $2 \leq \Delta i \leq 4$ )		331
	Long range ( $\Delta i \geq 5$ )		673
	Ambiguous		17
Hydrogen bonds*			24
Dihedral restraints			
Scalar couplings ( $^3J(\text{HN}, \text{H}\alpha)$ )			64
TALOS			228
Residual dipolar couplings ( $1D_{\text{HN}}$ )			109
<b>Properties of the accepted structure ensemble</b>			
Restraint violation			
Average distance restraint violation		$0.004 \text{ \AA} \pm 0.001 \text{ \AA}$	
Maximum distance restraint violation		$0.24 \text{ \AA}$	
Average dihedral restraint violation		$0.47^\circ \pm 0.05^\circ$	
Maximum dihedral restraint violation		$7.3^\circ$	
Average RDC restraint violation		$0.14 \text{ Hz} \pm 0.02 \text{ Hz}$	
Maximum RDC restraint violation		$1.35 \text{ Hz}$	
Deviation from ideal geometry			
Bond length		$0.00050 \text{ \AA} \pm 0.00005 \text{ \AA}$	
Bond angle		$0.108^\circ \pm 0.007^\circ$	
Coordinate precision			
Overall rmsd			
Backbone			$0.60 \text{ \AA}$
Heavy atoms			$0.88 \text{ \AA}$
Secondary structure rmsd			
aa 2-11; 15-23; 26-33; 37-44; 52-59; 62-74;			
79-87; 97-105; 111-120; 129-152			
Backbone			$0.53 \text{ \AA}$
Heavy atoms			$0.83 \text{ \AA}$
Ramachandran plot statistics			
Most favored regions			92.9 %
Additional allowed regions			7.1 %
Generously allowed regions			0.0 %
Disallowed regions			0.0 %

\*Two distance restraints per hydrogen bond ( $N_{\text{donor}}-O_{\text{acceptor}} \leq 3.3 \text{ \AA}$ ,  $H_{\text{donor}}-O_{\text{acceptor}} \leq 2.3 \text{ \AA}$ )

### Figures

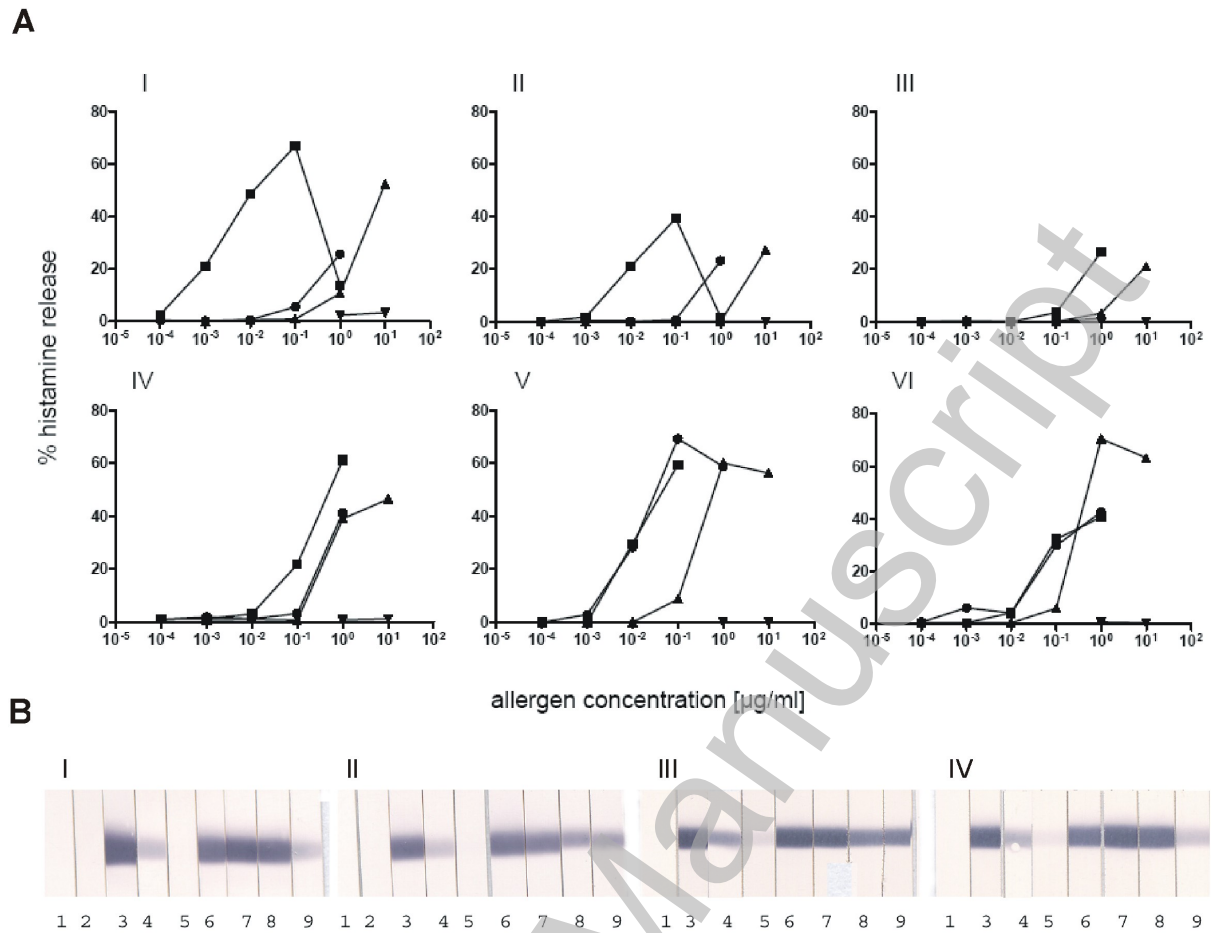


**Figure 1:** Sequence alignment of Gly m 4 with three homologous proteins from yellow lupine with known structure and the Bet v 1 allergens used in immunoblot inhibition experiments; sequences are sorted according to their similarity to Gly m 4; black background = identical amino acid; grey background = conservative exchange; threshold for shading: consensus in more than four sequences; alignment performed with ClustalW [43] and shaded with boxshade 3.21 (<http://www.ch.embnet.org>).

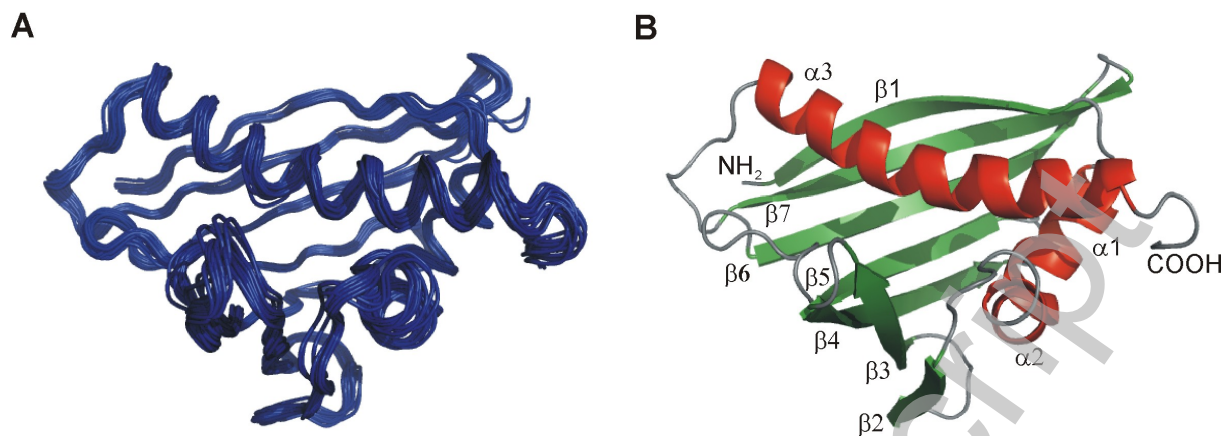
Accepted Article

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BSR20080117





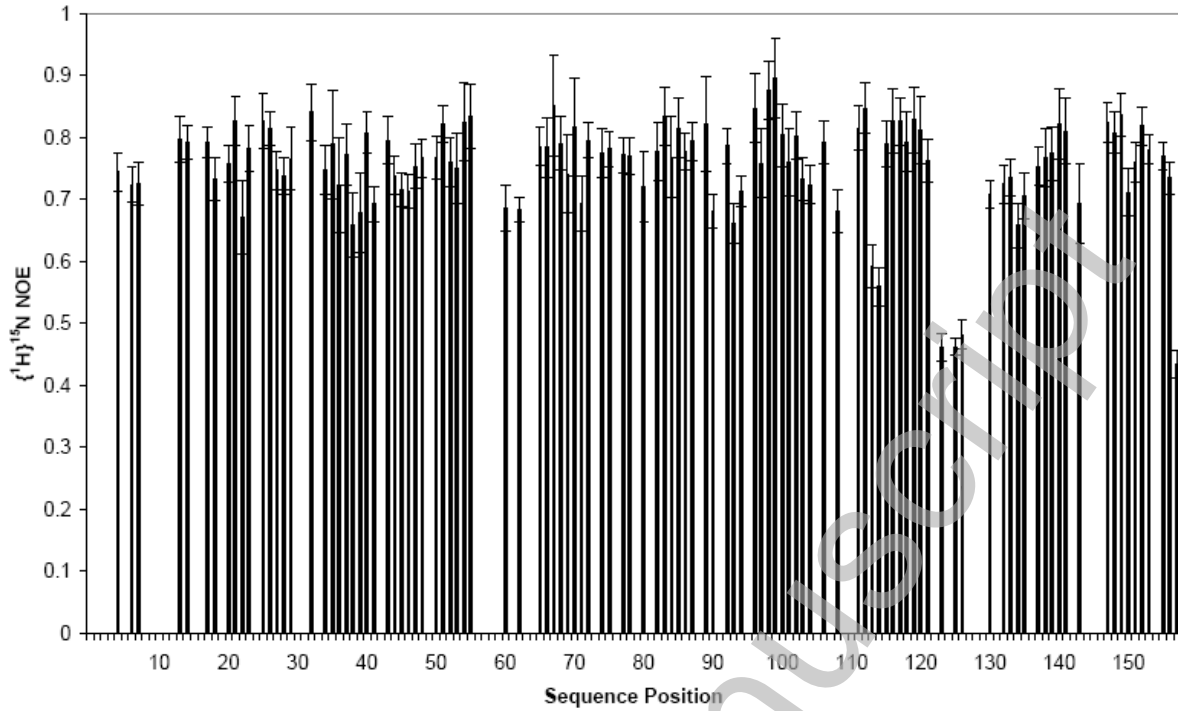
**Figure 2:** (A) Histamine release in six different patients allergic to soy after stimulation with soy isolate (▲), Gly m 4 (●), Bet v 1 (■), and ovalbumin (▼) as a negative control. (B) Inhibition of IgE-binding to Gly m 4 with allergens of the Bet v 1 family in four patients I-IV with patients II and IV corresponding to patients II and IV in (A): Lane 1: no serum, lane 2: serum of non-allergic control patient, lane 3: no inhibitor, lane 4: rGly m 4, lane 5: rBet v 1, lane 6: rApi g 1, lane 7: rDau c 1, lane 8: rPru av 1, lane 9: rMal d 1.



**Figure 3:** (A) Backbone overlay of the 20 lowest energy solution structures of Gly m 4; backbone rmsd = 0.6 Å. (B) Cartoon of the average structure of Gly m 4. The bent and shifted carboxy terminal helix sets this structure slightly apart from the strongly related Bet v 1 structure.

Accepted Manuscript

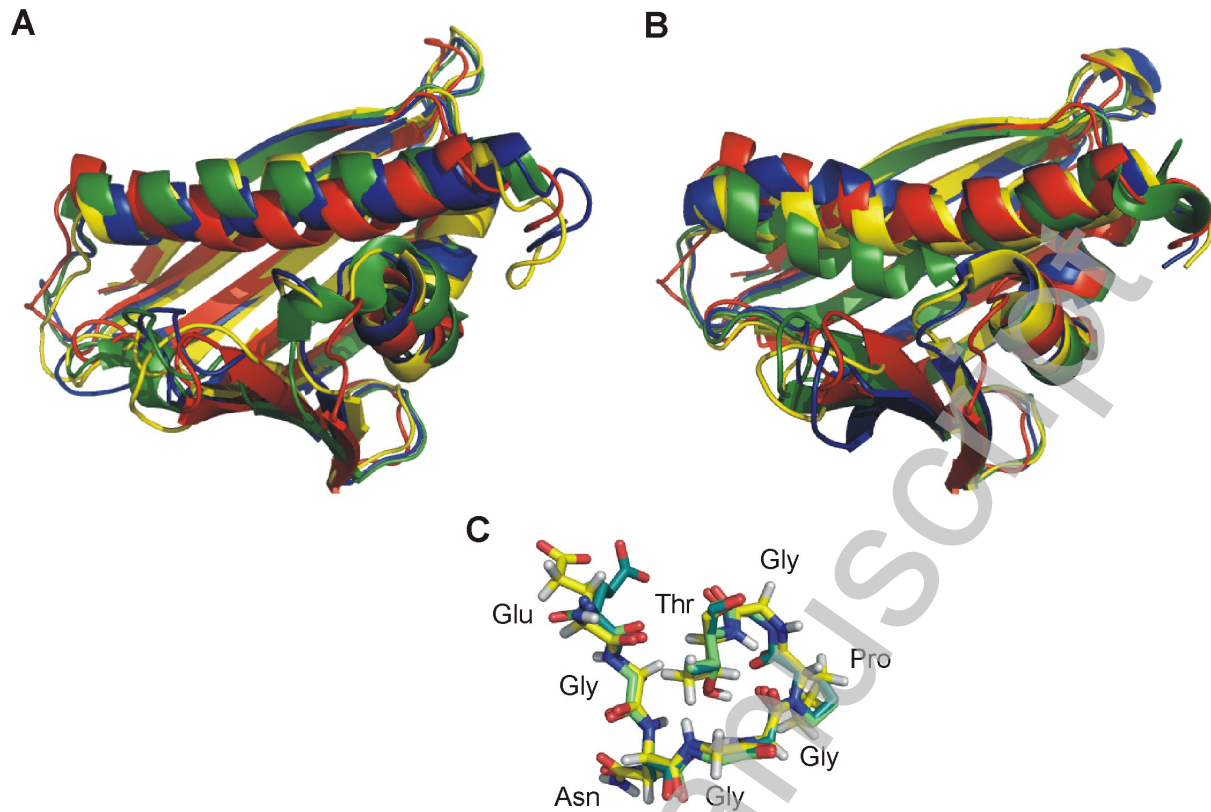
THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BSR20080117



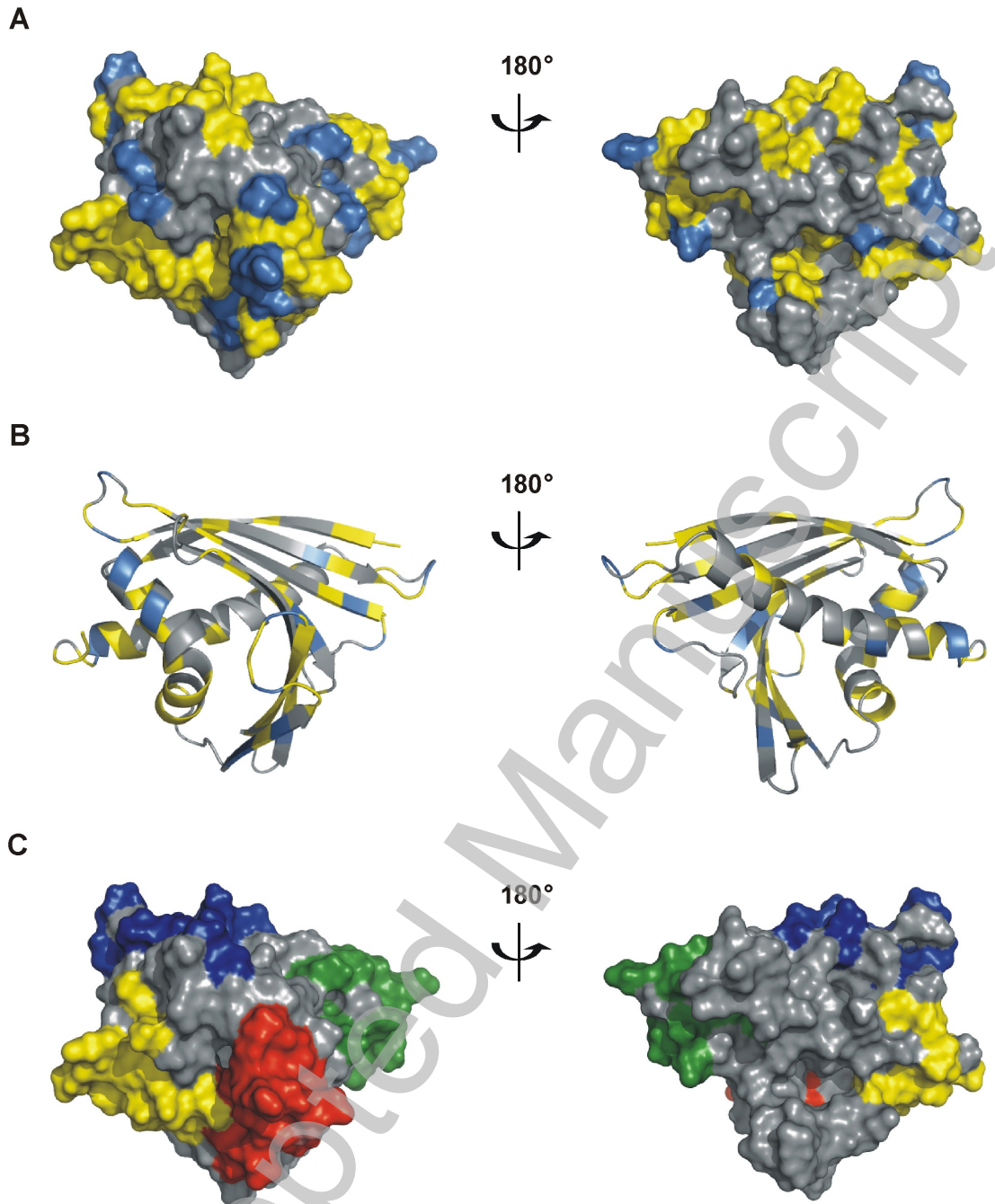
**Figure 4:**  $\{^1\text{H}\}^{15}\text{N}$  heteronuclear NOE values for Gly m 4; missing values could not be determined because of signal overlap; errors were calculated according to [17]. The increased flexibility of the amino terminus of strand  $\beta 7$  and of the connecting loop  $\beta 7 \rightarrow \alpha 3$  is clearly visible.

Accepted Manuscript

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BSR20080117



**Figure 5:** Overlay of the lowest energy solution structure of Gly m 4 (2K7H, red) with (A) the structures of the major birch pollen allergen Bet v 1 (1BV1, blue), the major cherry allergen Pru av 1 (1E09, yellow), and the celery allergen Api g 1 (2BK0, green); (B) the structures of three Bet v 1 homologous proteins from yellow lupine, namely LIPR10.1A (1ICX, blue), LIPR10.1B (1IFV, yellow), LIPR10.2A (1XDF, green); (C) Overlay of the P-loop region of Gly m 4 (yellow), Bet v 1 (bluegrey) and LIPR10.1B (light green); oxygen atoms = red, nitrogen atoms = blue. Whereas the P-loop region of Gly m 4, Bet v 1, and LIPR10.1B are virtually identical, the agreement of the structures of Gly m 4 and the lupine proteins is clearly better than the agreement of the structures of Gly m 4 and the allergens Bet v 1, Pru av 1, and Api g 1, mainly due to the bent carboxy terminal helix.



**Figure 6:** (A) Surface representation of the lowest energy structure of Gly m 4 with residues conserved in Bet v 1 labelled in blue and residues conserved in Bet v 1, Mal d 1 and Pru av 1 labelled in yellow; (B) Cartoon representation of Gly m 4 in the same orientation and with the same labeling as in (A); (C) Surface representation of the lowest energy structure of Gly m 4 in the same orientation as in (A) with putative cross-reactive epitopes marked in red (P-loop), yellow (helices  $\alpha 1$ ,  $\alpha 2$ ,  $\alpha 3$ ), green (strands  $\beta 1$ ,  $\beta 7$ ,  $\beta 6$ ,  $\beta 5$ ) and blue (strands  $\beta 1$ ,  $\beta 7$ ,  $\beta 6$ ).