International Journal of
Molecular Sciences

MDPI

Article

The Structural Flexibility of PR-10 Food Allergens

Sebastian Führer [1], Jana Unterhauser [1], Ricarda Zeindl [1], Reiner Eidelpes [1], Monica L. Fernández-Quintero [2], Klaus R. Liedl [2] and Martin Tollinger [1,*]

[1] Institute of Organic Chemistry and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innrain 80/82, 6020 Innsbruck, Austria; bastif92@web.de (S.F.); jana.unterhauser@uibk.ac.at (J.U.); ricarda.zeindl@uibk.ac.at (R.Z.); reiner.eidelpes@ki.se (R.E.)
[2] Department of General, Inorganic and Theoretical Chemistry and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innrain 80/82, 6020 Innsbruck, Austria; monica.fernandez-quintero@uibk.ac.at (M.L.F.-Q.); klaus.liedl@uibk.ac.at (K.R.L.)
* Correspondence: martin.tollinger@uibk.ac.at; Tel.: +43-512-504-57730

Abstract: PR-10 proteins constitute a major cause of food allergic reactions. Birch-pollen-related food allergies are triggered by the immunologic cross-reactivity of IgE antibodies with structurally homologous PR-10 proteins that are present in birch pollen and various food sources. While the three-dimensional structures of PR-10 food allergens have been characterized in detail, only a few experimental studies have addressed the structural flexibility of these proteins. In this study, we analyze the millisecond-timescale structural flexibility of thirteen PR-10 proteins from prevalent plant food sources by NMR relaxation-dispersion spectroscopy, in a comparative manner. We show that all the allergens in this study have inherently flexible protein backbones in solution, yet the extent of the structural flexibility appears to be strikingly protein-specific (but not food-source-specific). Above-average flexibility is present in the two short helices, α1 and α2, which form a V-shaped support for the long C-terminal helix α3, and shape the internal ligand-binding cavity, which is characteristic for PR-10 proteins. An in-depth analysis of the NMR relaxation-dispersion data for the PR-10 allergen from peanut reveals the presence of at least two subglobal conformational transitions on the millisecond timescale, which may be related to the release of bound low-molecular-weight ligands from the internal cavity.

Keywords: protein flexibility; conformational breathing; ligand binding; NMR spectroscopy; relaxation dispersion

check for updates

Citation: Führer, S.; Unterhauser, J.; Zeindl, R.; Eidelpes, R.; Fernández-Quintero, M.L.; Liedl, K.R.; Tollinger, M. The Structural Flexibility of PR-10 Food Allergens. *Int. J. Mol. Sci.* **2022**, *23*, 8252. https://doi.org/10.3390/ijms23158252

Academic Editors: Gyula Batta and Orsolya Tőke

Received: 8 July 2022
Accepted: 23 July 2022
Published: 26 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Pathogenesis-related (PR) proteins represent a major source of allergic reactions to plants, and they play a major role in the defense mechanisms to external stimuli or pathogenic infections [1]. In the last decades, a vast number of PR proteins have been discovered, studied, and classified into different groups based on their similarity regarding their biological activities, sequences, and regulatory or physicochemical properties. Currently, PR proteins are divided into 17 subfamilies, with distinct biological features [2,3]. Highly prevalent and important is the family of PR proteins in class 10 (PR-10), which were first discovered in parsley [4], and which comprise hundreds of members across the plant kingdom. Their role in the general defense mechanism of plants is evident by the fact that these proteins are upregulated after viral, bacterial, or fungal invasion, or due to abiotic factors, such as cold, aridity, oxidative stress, and UV radiation, or, in general, by phytohormones [5].

The eponymous and most prominent PR-10 allergen is the birch pollen (*Betula verrucosa*) protein Bet v 1, which was first characterized in 1983 [6], and which provokes immediate hypersensitivity reactions of type I, resulting in the development of Bet v 1-specific IgE antibodies. The so-called cross-reactivity to other food and pollen allergens

has been observed in clinical studies [7–9], and it arises from the highly similar amino acid sequences and three-dimensional structures of these proteins. This leads to the presence of analogous and overlapping epitopes in these proteins, which are recognized by IgE antibodies, which trigger an allergic response [10]. Allergic symptoms include the oral allergy syndrome, which is characterized by the itching and swelling of the throat, tongue, and lips [11], along with rare anaphylactic shocks in birch-pollen-sensitized individuals [12,13]. The most prevalent Bet v 1 cross-reactive plant food sources are apple (inducing allergic symptoms in 80% of Bet v 1-sensitized patients), hazelnut (59%), peach (51%), kiwi (48%), peanut (24%), and plum (24%) [14].

PR-10 allergens are of similar molecular sizes (156 to 160 residues) and masses (~17.5 kDa). A striking feature of PR-10 allergens is their highly conserved and characteristic secondary structures, which comprise three α-helices, along with a seven-stranded antiparallel β-sheet (Figure 1). The first strand, β1, is connected to strand β2 via two short helices, α1 and α2. Together with β1 and β2, the remaining five β-strands (β3–β7) complete the β-sheet, which displays a pronounced curvature. The structural composition is closed by a long C-terminal helix (α3), which is located above the β-sheet and the two short helices α1 and α2. Generally, PR-10 allergens display medium to high amino acid-sequence identities among each other, with different levels of conservation along the sequence. While the glycine-rich region (GXGGXG) in the loop connecting strands β2 and β3 (positions 46–51, using the numbering scheme of the birch-pollen allergen Bet v 1) is highly conserved in PR-10 allergens, the most variable region regarding both length and sequence is located at the C-terminus (Supporting Figure S1) [1,5].

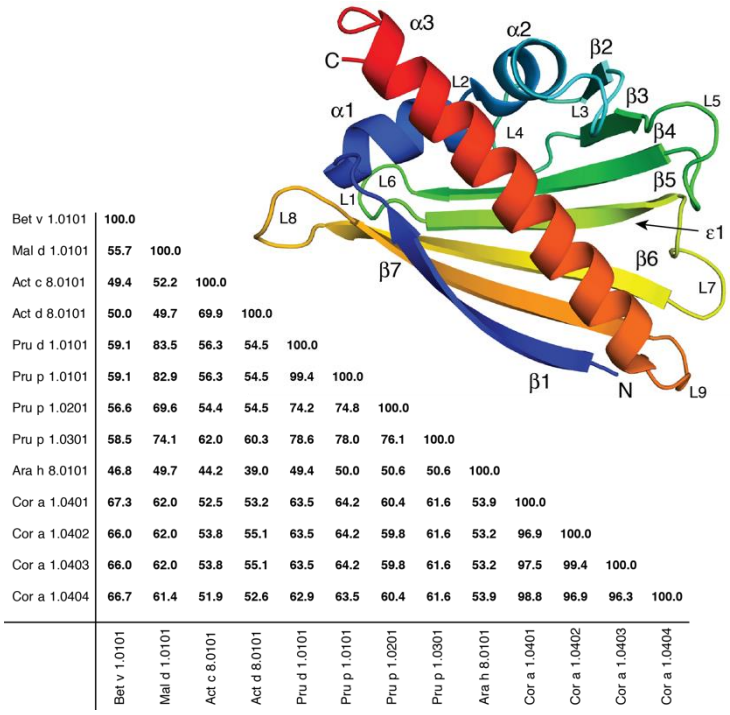


Figure 1. Sequence-identity matrix of thirteen selected PR-10 allergens from birch pollen (*Betula verrucosa*, Bet v 1), apple (*Malus domestica*, Mal d 1), golden kiwi fruit (*Actinidia chinensis*, Act c 8), green kiwi fruit (*Actinidia deliciosa*, Act d 8), plum (*Prunus domestica*, Pru d 1), peach (*Prunus persica*, Pru p 1), peanut (*Arachis hypogaea*, Ara h 8), and hazelnut (*Corylus avellana*, Cor a 1). The structure of Bet v 1 (PDB: 4A88, isoform Bet v 1.0101) is displayed above. Secondary-structure elements and the nine connecting loops (L1–L9) are labeled, and N- and C-termini are indicated. The main entrance (ε1) to the inner cavity of the protein is indicated by an arrow.

The tertiary fold of PR-10 proteins comprises a large internal cavity of a vastly variable size, with volumes up to ~4000 Å^3 [5]. While the cavity is, in large parts,

formed by hydrophobic amino acids, several hydrophilic amino acids with hydrogen-bond donor/acceptor capacities confer a certain degree of amphiphilicity to its surface [15]. Entrances to the cavity are located between helix α3 and loops L3, L5, and L7 (ε1, the largest opening), and in some PR-10 proteins, between helix α3 and strand β1 (ε2) [5,16]. PR-10 proteins promiscuously bind a vast variety of low-molecular-weight ligands to their internal cavities with different affinities [5], including fatty acids, flavonoids, phenolic acids, cytokinins, and steroids [17,18]. Whether and how the ligand binding to PR-10 allergens is related to their immunologic properties are currently under debate [19,20].

The biophysical research on PR-10 allergens in the last decades has focused on ligand binding and structure elucidation by NMR spectroscopy or X-ray crystallography, and on understanding how the structural properties of these proteins are related to the observed allergenic behavior [16]. Only a few studies have addressed the structural flexibility of these proteins, which is a feature that may be related to the affinity of low-molecular-weight ligands binding to the internal cavity, and IgE binding to the outer surfaces of these proteins [21–24]. In this work, we provide a thorough and quantitative comparison of the structural flexibilities of 12 PR-10 food allergens from apple, hazelnut, peach, kiwi, peanut, and plum, along with—and in comparison to—the sensitizing birch-pollen allergen, as observed by solution NMR spectroscopy. The possible implications for low-molecular-weight ligand binding and interactions with antibodies are discussed.

2. Results

In order to generate a comprehensive dataset covering the most important food sources, we chose thirteen different PR-10 allergens (and the isoforms thereof): Pru p 1 from peach (*Prunus persica*), Cor a 1 from hazelnut (*Corylus avellana*), Mal d 1 from apple (*Malus domestica*), Ara h 8 from peanut (*Arachis hypogaea*), Act c 8 from golden kiwi fruit (*Actinidia chinensis*), Act d 8 from green kiwi fruit (*Actinidia deliciosa*), and Pru d 1 from plum (*Prunus domestica*), along with the sensitizing allergen from the pollen of birch trees (*Betula verrucosa*), Bet v 1.

2.1. Surface Properties of the PR-10 Allergens

Figure 2 compares the surface properties of these proteins. The clustering based on the charge profiles and H-bond donor/acceptor capacities of the surfaces, identified and classified by the YRB highlighting scheme of Hagemans et al. [25], shows that these proteins differ significantly regarding their surface properties, which is in agreement with their considerable sequence variability. Cluster 1 comprises the isoforms Cor a 1.0401, Cor a 1.0404, Mal d 1.0101, and Pru p 1.0101, while Cluster 2 includes Ara h 8.0101, Act d 8.0101, Bet v 1.0101, Cor a 1.0402, Cor a 1.0403, Pru d 1.0101, Pru p 1.0201, and Pru p 1.0301. Cluster 3 only contains Act c 8.0101. On the one hand, the three clusters differ particularly regarding the distribution of charged residues, which are mostly present in the first cluster. Cluster 2, on the other hand, contains a higher number of hydrophobic residues and less charged residues on the surface of the helix α3. The golden kiwifruit allergen Act c 8.0101 forms a separate cluster, which can be explained by the high number of surface-exposed cysteine residues (Act c 8.0101 contains seven cysteines, with six of them located on the protein surface). Moreover, all three clusters show high densities of amino acid residues, with the hydrogen-bond capacity on the surface. This comparison emphasizes that these PR-10 allergens are remarkably diverse with regard to the electrostatic properties of the surface that is exposed to the solvent, which is in line with the observation that their immunologic properties, including their cross-reactivity and IgE binding, are diverse [26–28], despite the fact that these allergens display highly similar three-dimensional scaffolds (Supporting Figure S2).

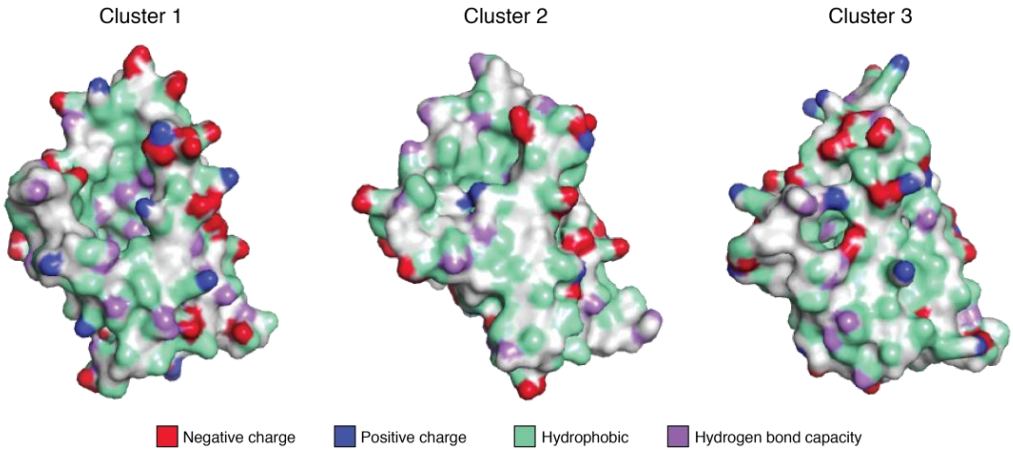


Figure 2. Clustering based on the charge profiles and hydrogen-bond donor/acceptor capacities of the thirteen PR-10 allergens. Cluster 1 comprises the proteins Cor a 1.0401, Cor a 1.0404, Mal d 1.0101, and Pru p 1.0101; Cluster 2 comprises Ara h 8.0101, Act d 8.0101, Bet v 1.0101, Cor a 1.0402, Cor a 1.0403, Pru d 1.0101, Pru p 1.0201, and Pru p 1.0301; Cluster 3 only contains Act c 8.0101. The surfaces are color-coded following the YRB highlighting scheme of Hagemans et al. [25]. The color-coding of positively and negatively charged residues, hydrophobic residues, and residues with hydrogen-bond capacities is specified. All three clusters are shown in the same orientation, with the C-terminal α-helix facing to the front, with its C-terminus pointing down.

2.2. Millisecond-Timescale Flexibility of PR-10 Allergens

We used backbone amide ^{15}N relaxation-dispersion (RD) NMR experiments to probe the structural flexibility of the selected PR-10 allergens by experiment. RD data contain quantitative information about the structural flexibility that occurs on the millisecond timescale on a per-residue basis [29]. As noted previously for PR-10 allergens from birch pollen and hazelnut, these proteins are remarkably flexible, with a significant portion of the backbone amides showing nonflat relaxation-dispersion curves [21,22]. Figure 3A compares the backbone amide ^{15}N relaxation-dispersion amplitudes ($\Delta R_{2,eff}$ values) of the thirteen PR-10 proteins in this study, in detail. Significant differences regarding the percentages of amino acid residues with nonflat RD profiles ($\Delta R_{2,eff}$ exceeding 1 s^{-1}) are evident, ranging from 50% for the plum allergen Pru d 1, to more than 80% in different isoforms of the peach and hazelnut allergens, Pru p 1 and Cor a 1, respectively. Interestingly, however, these values do not correlate with the amino acid sequences in a straightforward manner. For example, peach Pru p 1.0101 and plum Pru d 1.0101, one of the most disparate pairs of allergens regarding the percentage of flexible residues, share 99.4% of their sequence identity and have only one amino acid substituted, at position 117 in the center of strand β7. Cor a 1.0401 and Cor a 1.0404, which are two isoforms of the hazelnut PR-10 allergen, share 98.8% of their sequence identity, but they appear to be significantly different regarding the extent of the structural flexibility in their protein backbones. Indeed, these data reinforce the notion that different isoforms from one particular food source (e.g., hazelnut) can have substantially different backbone flexibilities, even if the sequence identities are high (>96% for the four hazelnut PR-10 isoforms in this study), while, for other food sources (e.g., peach), isoforms with much lower sequence identities (<80%) can have similar backbone flexibilities. Moreover, Figure 3A indicates that the birch pollen Bet v 1.0101, peach Pru p 1.0301, peanut Ara h 8.0101, and golden kiwi Act c 8.0101 are the PR-10 proteins with the highest percentages of residues with $\Delta R_{2,eff}$, exceeding 10 s^{-1}, while apple Mal d 1.0101, hazelnut Cor a 1.0404, and green kiwi Act d 8.0101 display particularly small numbers of amino acids, with $\Delta R_{2,eff} \geq 10$ s^{-1}. To conclude, it is evident from this compilation that, while all PR-10 allergens investigated in this study are structurally flexible, the extent of the structural flexibility is strikingly protein-specific.

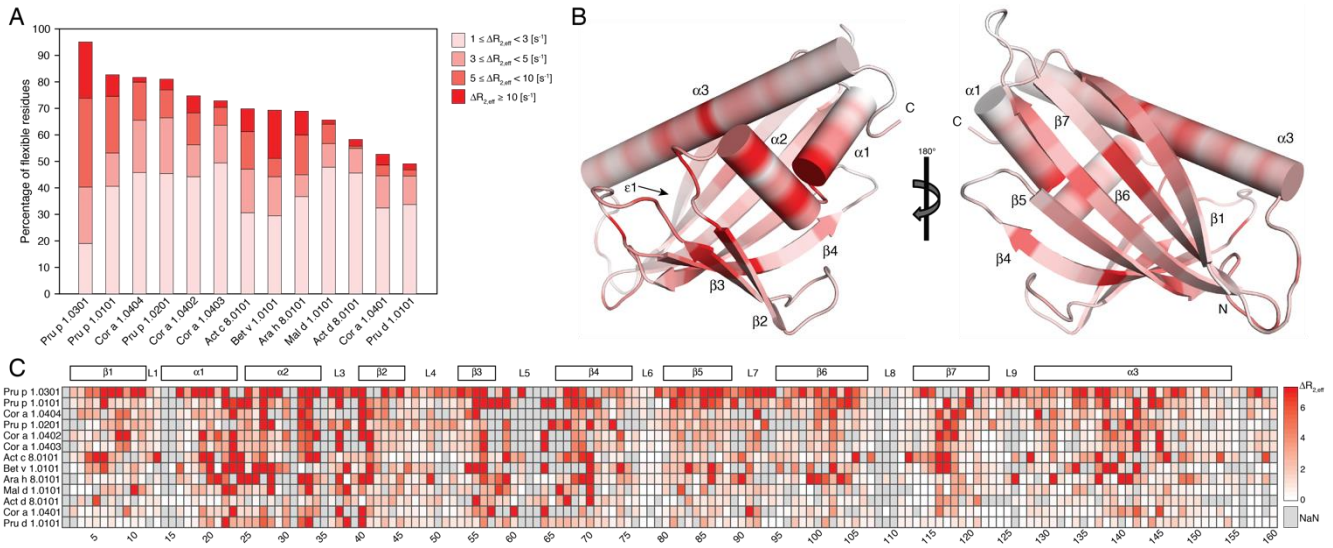


Figure 3. Structural flexibility of thirteen selected PR–10 allergens. (**A**) Bar plot showing the percentages of residues with backbone amide ^{15}N relaxation-dispersion amplitudes ($\Delta R_{2,eff}$) exceeding 1 s^{-1} at 700 MHz. Bars are color-coded according to ranges of $\Delta R_{2,eff}$ values. (**B**) Residue-specific mean values of $\Delta R_{2,eff}$ over all thirteen PR–10 proteins are plotted on the structure of Bet v 1.0101 (PDB: 4A88), color-coded from rigid (white) to flexible (red). (**C**) Heat map comparing backbone amide ^{15}N RD amplitudes vs. residue numbers, colored according to the shown gradient. Proteins are ordered as in (**A**). Grey squares indicate sequence positions for which experimental data are not available (*i*) due to gaps in a specific PR–10 protein sequence, (*ii*) because of the resonance overlap of backbone amides, (*iii*) for prolines, or (*iv*) in cases where resonance assignments are not available.

Figure 3B,C compile the available experimental information regarding the location of the structural flexibility in these proteins. Evidently, the structural flexibility is distributed along the entire backbone of all thirteen proteins, rather than being limited to the loops or turns of the PR-10 fold. Surprisingly high levels of flexibility are present in all secondary-structure elements: α1–α3 and β1–β7. A region with above-average structural flexibility is situated between residues 20 and 40, which corresponds to the two short helices α1 and α2 that form a V-shaped support for the long C-terminal helix α3, which itself also contains flexible amino acid residues (in particular, at its center, between residues 135–148). As a matter of fact, this holds true for all the PR-10 proteins in this study (the relaxation-dispersion amplitudes for individual PR-10 proteins are shown in Supporting Figure S3). Nonflat RD profiles are also found for the β-sheet, and especially at the beginning of strand β2, which packs to helix α2 at the tip of β3, which is directly adjacent and hydrogen-bonded to strand β2, in the center of strand β7. In addition, surface-exposed loops between these secondary-structure elements (e.g., loop L2 connecting helices α1 and α2, and loop L3 connecting helix α2 with strand β3) display considerable millisecond-timescale flexibility. For loop L5, which is located between strands β3 and β4, the NMR resonances are broadened beyond detection in the majority of the PR-10 proteins, which indicates high levels of structural flexibility.

Together, some of these flexible segments (helices α1, α2, strands β2 and β3, and the central part of helix α3) shape the surface of the inner cavity of these proteins, which suggests that the cavity itself is malleable. In addition, a number of amino acid residues that form the largest entrance (ε1) to the cavity (loops L3 and L5, helix α3) display above-average flexibility. This is clearly contrasted by the glycine-rich loop (L4, connecting strands β2 and β3) and loop L6 (between β4 and β5), which display lower percentages of residues with nonflat relaxation-dispersion profiles.

2.3. Millisecond-Timescale Flexibility of Ara h 8

We chose the peanut allergen Ara h 8.0101 for an in-depth analysis of the backbone millisecond flexibility and ligand binding to the inner cavity. For this protein, which is a minor allergen that is present in peanut that shows IgE cross-reactivity with Bet v 1.0101, a high-resolution X-ray structure in a complex with (-)-epicatechin is available [30]. NMR resonance assignments for Ara h 8.0101 were obtained using triple-resonance experiments, as described [31], and the backbone amide ^{1}H-^{15}N-HSQC spectrum of this allergen is shown in Figure 4A. Excluding the N-terminal glycine residue, we were able to assign the backbone amide resonances of 144 (out of 146) nonproline residues of Ara h 8.0101, corresponding to 98.6% completeness. The majority of the resonances are well dispersed in the ^{1}H-^{15}N-HSQC spectrum, and the backbone amide ^{15}N relaxation-dispersion data are accessible for 95.1% of all the backbone amide resonances. A total of 69.2% of the backbone amides in Ara h 8.0101 show nonflat relaxation-dispersion profiles, with RD amplitudes exceeding 1 s^{-1} at 700 MHz, which is the median of the allergens in this study. This particular protein thus constitutes a representative example of PR-10 food allergens.

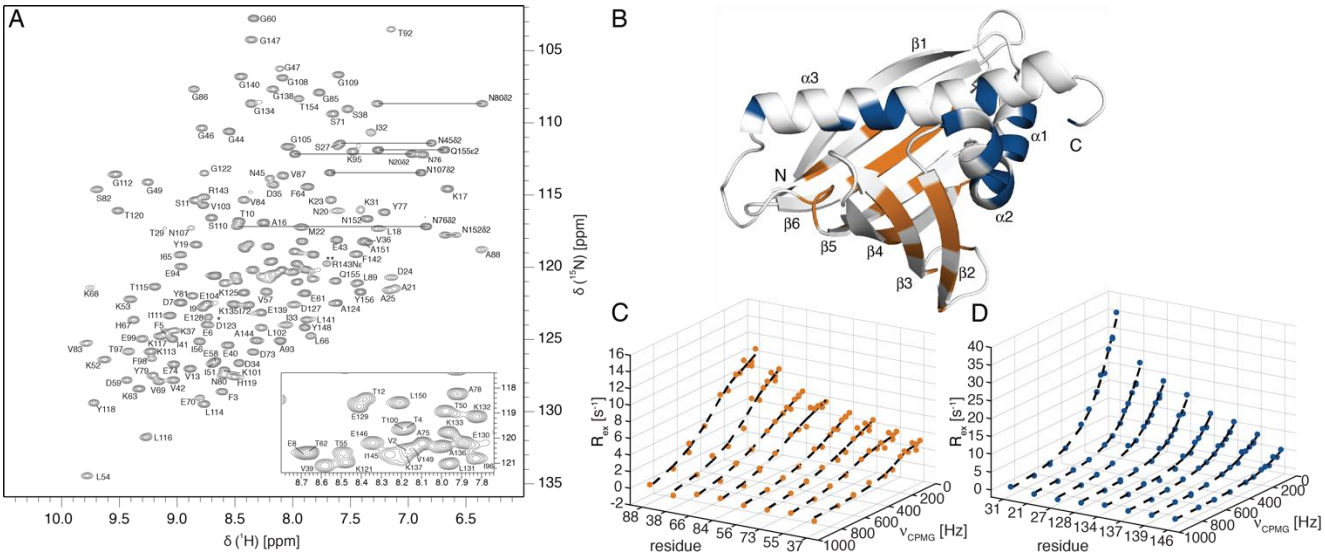


Figure 4. Backbone amide NMR resonance assignments and relaxation-dispersion data for Ara h 8.0101. (**A**) The 700 MHz ^{1}H–^{15}N–HSQC spectrum of Ara h 8.0101 (0.5 mM) in 20 mM sodium phosphate (pH: 6.9), supplemented with 10% D_2O, at 25 °C. Assignments are shown using single–letter codes. Horizontal lines indicate asparagine and glutamine NH_2 side–chain resonances. The resonance labeled by one asterisk indicates a residue below the intensity cutoff, the signal labeled by two asterisks indicates the aliased signal of the single arginine sidechain ε–NH. (**B**) Structure of Ara h 8.0101 (PDB: 4MA6) with backbone amides that were used in global fits of the ^{15}N relaxation-dispersion data in the β-sheet and three α-helices, highlighted in orange and blue, respectively. (**C**,**D**) Backbone amide ^{15}N relaxation-dispersion data for eight representative amino acid residues in the β-sheet (**C**) and in helices α1–α3 (**D**). Exchange contributions (R_{ex}), which are calculated as $R_{2,eff} - R_{2,eff}$ ($\nu_{CPMG} = \infty$), are shown, along with the best fit-lines, which were obtained separately from the global fits of all the residues in (**C**,**D**), respectively.

We analyzed the backbone amide ^{15}N relaxation-dispersion data of Ara h 8.0101, recorded at 600 MHz and 700 MHz, by fitting the equation derived by Baldwin [32] to the experimental data. Initially, we attempted to fit a single global conformational exchange process (transition) between two states to the experimental data for all the residues by assuming a common exchange-rate constant (k_{ex}), along with uniform populations, but allowing for residue-specific values of the chemical-shift difference between the two states ($\Delta\delta_{ex}$), as described in the Methods section. This resulted in fits of insufficient quality and low-convergence performance (data not shown). However, the fitting procedure improved

significantly, when we assumed that two distinct processes were present, by grouping amino acid residues in Ara h 8.0101 according to their positions in the secondary-structure elements (Figure 4B). One subglobal cluster contained all the residues that had nonflat relaxation-dispersion profiles in helices α1, α2 (including the short loop L2 between α1 and α2), and α3 (20 residues in total), while the second subglobal cluster comprised all the other residues, located in the central β-sheet and in loops L1 or L3–L9 (26 residues). For each cluster, we again assumed uniform exchange-rate constants and populations, but residue-specific $\Delta\delta_{ex}$ values. Figure 4C,D shows the exemplary relaxation-dispersion data and fitting curves for these two groups of residues. It is evident from a visual inspection of the data for these amino acids that they are dissimilar regarding the shapes of the RD profiles. The residues in the β-sheet and in the loops (Lys37, Ser38, Thr55, Ile56, Leu66, Asp73, Val84, Ala88 in Figure 4C) have rising profiles with flat plateaus at low CPMG frequencies, whereas the three α-helices (Ala21, Ser27, Lys31, Glu128, Gly134, Lys137, Glu139, Glu146 in Figure 4D) display ascending RD profiles. This suggests that the underlying conformational exchange processes are not identical.

The fitting of the experimental data shows that the two processes are indeed distinct: the exchange rate for the subglobal conformational transition comprising helices α1–α3 is $k_{ex} = 1060 \pm 70\ s^{-1}$, with a minor state population (7.9 ± 1.4%), while a somewhat slower transition with a lower minor state population is present in the β-sheet ($k_{ex} = 810 \pm 80\ s^{-1}$, minor state population: 1.3 ± 0.1%). These figures translate to slightly different timescales for the β-sheet (1.2 ± 0.1 ms) and the three helices (0.9 ± 0.1 ms). Of note, the residue-specific ^{15}N chemical-shift differences that were obtained from the fits of the relaxation-dispersion data ($\Delta\delta_{ex}$) are quite substantial (Table S1). For the backbone amides in the β-sheet, $\Delta\delta_{ex}$ values exceeding 3 ppm (^{15}N) were found for residues in strands β2 (Ser38), β4 (Leu66, His67), β5 (Val84), β6 (Thr97, Phe98), and loop L7 (Ala88). Such large $\Delta\delta_{ex}$ values indicate the transient loosening or even unfolding of the β-sheet. Hydrogen/deuterium exchange experiments showed that the hydrogen bonds in the four hazelnut Cor a 1.040x isoforms are indeed labile, and amide protons are susceptible to exchange with the surrounding water [21]. For helices α1–α3, the $\Delta\delta_{ex}$ values are generally smaller (<2.5 ppm), and they exceed 1.5 ppm only for Ala21, Ser27, and Lys31 in the two helices (α1 and α2) that form the V-shaped support for α3. A comparison of the $\Delta\delta_{ex}$ values with the chemical-shift differences that are expected for the unfolding of the protein backbone shows that these values are not correlated for the β-sheet and the three helices (Figure S4). Rather, the experimental $\Delta\delta_{ex}$ values are systematically smaller than the unfolding shift changes. This implies that the conformational transitions that we observe by relaxation-dispersion NMR do not result from the unfolding of the protein scaffold to a random coil. Nevertheless, the considerable size of the observed chemical-shift differences indicates that the underlying conformational transitions involve the local restructuring and/or loosening of the hydrogen bonds in the secondary structures.

2.4. Ligand Binding to Ara h 8

We probed the interaction of (-)-epicatechin with Ara h 8.0101 in solution in an NMR titration experiment. Hurlburt et al. showed that this naturally occurring flavan-3-ol binds to an amphiphilic pocket inside the cavity of Ara h 8, which is formed by helices α1–α3 and strands β2–β5 (Figure 5) [30]. The addition of (-)-epicatechin to Ara h 8.0101 leads to a gradual change in the chemical shifts of the resonances, in accordance with an intermediate-to-fast binding process on the NMR chemical-shift timescale, as is typically observed for low-affinity binding (Figure 5A). The largest chemical-shift changes of the backbone ^{1}H-^{15}N-resonances ($\Delta\delta_{obs}$) are found for residues in helices α1 (Met22) and α2 (Ala25 and Lys31), loop L3 (Ile33 Asp34, and Val36), strand β4 (Thr55, Ile56, and Val57), loop L7 (Val87), strand β6 (Phe98), and for the central part of the C-terminal helix α3 (Gly138, Leu141, and Ile145). These amino acid residues cluster around the (-)-epicatechin binding site in the crystal structure (Figure 5C). Indeed, in this complex, similar amino acid residues (Asp26, Thr29, Ile56, His67, Tyr81, Lys137, and Leu141) were identified as the primary

interaction sites of (-)-epicatechin via hydrogen bonding and noncovalent interactions [30], which confirms that, in solution, this ligand binds to Ara h 8.0101 in a similar manner as in the crystalline state. In the quantitative analysis of the NMR titration data for backbone amide resonances with chemical-shift perturbations ($\Delta\delta_{obs}$), exceeding 0.1 ppm yields an average complex dissociation constant (K_D) of 0.76 ± 0.37 mM. Interestingly, 53.3% of these backbone amides exhibit relaxation-dispersion amplitudes exceeding 5 s^{-1} in the apo-form of the protein, while only 20.3% of all the other backbone amides in this protein fall into this category. Likewise, 20.0% of the interacting residues have RD amplitudes exceeding 10 s^{-1}, contrasting only 7.6% in the remainder of the backbone. This suggests that the (-)-epicatechin binding site displays above-average flexibility before the ligand is bound.

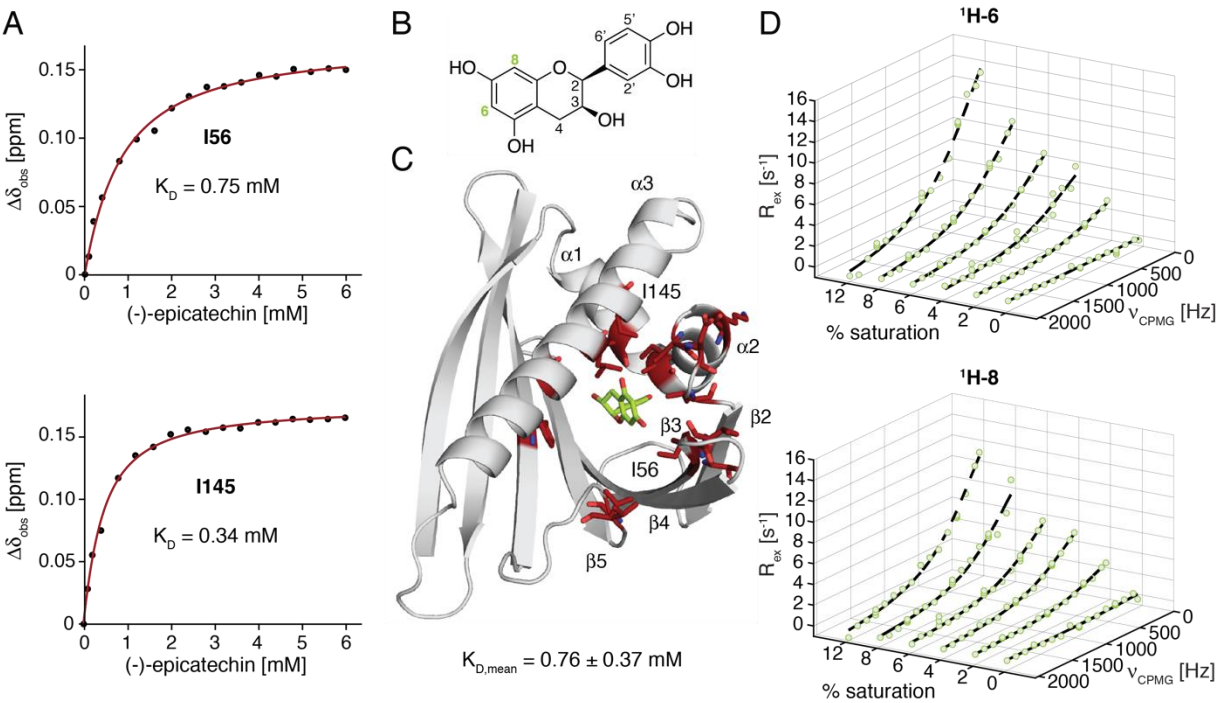


Figure 5. NMR spectroscopic analysis of binding of (–)–epicatechin to Ara h 8.0101. (**A**) Binding curves for two representative amino acid residues (Ile56 in strand β3, and Ile145 in helix α3). Site-specific K_D values were derived from ligand-induced ^{1}H and ^{15}N chemical-shift perturbations ($\Delta\delta_{obs}$) by nonlinear least-squares fitting [33]. (**B**) Structure and numbering scheme of (–)–epicatechin. (**C**) X-ray structure of Ara h 8.0101 in complex with (–)–epicatechin (PDB: 4MA6). Amino acid residues in red show backbone amides with $\Delta\delta_{obs}$ > 0.1 ppm upon (-)-epicatechin binding that were used to determine mean value and standard deviation of K_D. (**D**) Ligand-observed ^{1}H–CPMG relaxation-dispersion profiles, recorded at 600 MHz, for (-)-epicatechin binding to Ara h 8.0101. Data are shown for protons ^{1}H–6/8 (indicated in green in panel (**B**)), and for increasing saturation (0–12%) of the ligand by protein. Solid lines represent the global fit to the data, yielding the off-rate (k_{off}) of the (-)-epicatechin from the complex.

We next employed ligand-observed relaxation-dispersion experiments to investigate the kinetics of the (-)-epicatechin binding to Ara h 8.0101 [34]. The two aromatic protons (^{1}H-6 and ^{1}H-8) of (-)-epicatechin (Figure 5B) were resolved in 600 MHz ^{1}H-NMR spectra, with a scalar $^4J_{HH}$ coupling constant of <2 Hz, which qualified them for ^{1}H-CPMG relaxation-dispersion experiments. Figure 5D shows the experimental ^{1}H-CPMG RD profiles recorded for an increasing saturation (0–12%) of the ligand with Ara h 8. The RD profile for (-)-epicatechin in the absence of Ara h 8 is perfectly flat, which indicates that scalar coupling does not interfere with the experiment. The addition of protein produces nonflat RD profiles, with a clear dependence on the saturation level, which is indicative of the ligand binding and release in the millisecond time regime. The fitting of the data with a

two-site kinetic model yields the rate constants that describe the ligand release for ^{1}H-6 (k_{off} = 1600 ± 300 s^{-1}) and ^{1}H-8 (k_{off} = 1300 ± 300 s^{-1}), which translates to a residence time of (-)-epicatechin in the protein-bound state of just below one millisecond (0.6 ± 0.1 ms and 0.8 ± 0.2 ms, respectively).

3. Discussion

The accumulated NMR spectroscopic data in this study show that, in solution, PR-10 allergens are remarkably flexible. In some of these proteins, more than 80% of all the amino acid residues have nonflat relaxation-dispersion curves, which indicates the backbone flexibility that occurs on the millisecond timescale. Quite generally, the most flexible segments of PR-10 proteins are centered around helices α1 and α2, which form the V-shaped support for the long C-terminal helix α3. Together with the curved β-sheet, these helices shape the internal binding cavity of these proteins, which suggests that the surface of the cavity itself is structurally heterogeneous and is capable of adjusting to different ligands [35]. Conformational variability in the internal cavity has been observed in the crystal structures of various PR-10 proteins, such as the strawberry allergen Fra a 1, where the C-terminal helix α3 bends toward the central β-sheet when the ligand is bound, forming a more compact and closed three-dimensional structure [36]. Likewise, loop L5, at the entrance to the internal cavity of this allergen, appears to be conformationally heterogenous in the crystalline state, but it adopts a more closed conformation upon ligand binding [37]. In solution, most NMR resonances in loop L5 of PR-10 food allergens are broadened beyond detection due to conformational heterogeneity.

It is possible that the inherent flexibility of PR-10 proteins represents a structural reorganization process that is required for ligand entry and/or release. An in-depth analysis of the NMR experimental data of the peanut allergen Ara h 8 suggests that structural flexibility is reminiscent of "conformational breathing". The chemical-shift data indicate that the structural changes that arise from this process are quite substantial and might involve the loosening of hydrogen bonds, even in secondary-structure elements. This is in accordance with our previous observation that structural flexibility in PR-10 proteins enhances the exposure of the protein backbone to the surrounding water [21]. The conformational breathing of the PR-10 scaffold appears to involve at least two subglobal processes, both occurring on the millisecond timescale (1.2 ± 0.1 ms for the β-sheet, and 0.9 ± 0.1 ms for the three α-helices). These values qualitatively agree with the residence time of (-)-epicatechin in the bound state, which was found to be just below one millisecond (0.6–0.8 ms). While these values suggest that ligand release may indeed be mechanistically coupled to the inherent structural flexibility of the PR-10 scaffold, additional studies using a comprehensive set of ligand molecules with diverse affinities will be required to corroborate this observation.

Structural flexibility also plays an important role regarding the interactions between allergens and antibodies [38]. Antibody binding may be accompanied by rigidification and a loss of conformational entropy, involving both the antigen's paratope, as well as the epitope on the surface of a particular allergen. For the four hazelnut Cor a 1.040x isoforms, an inverse correlation between the overall structural flexibility of the allergen and IgE binding has indeed been recently reported [21]. It could be shown that more flexible isoforms of this particular allergen have lower IgE binding potential than more rigid ones. The compilation of the data presented here clearly illustrates that the structural flexibility in PR-10 allergens is not evenly distributed along the protein backbone. The glycine-rich loop, which has been identified as a cross-reactive IgE epitope in some of these allergens, is particularly rigid and highly conserved, and it has a conformation that is compatible with antibody binding [15,35,39]. This is contrasted by the C-terminal helix α3, in which the conformational flexibility is pronounced, and the variation in the primary sequence is high. Additionally, helix α3 has been identified as an important IgE epitope in Bet v 1 [40].

These examples suggest that entropic contributions that arise from epitope rigidification may well be present in some—but probably not in all—cases of antibody/antigen

binding. A combination of the (static) structural data and site-specific information about the structural flexibility are required to fully comprehend these interactions.

4. Materials and Methods

4.1. Recombinant Protein Expression and Purification

The construction of plasmids encoding for the PR-10 proteins, and the subsequent expression and purification, has been described for Act c 8.0101, Act d 8.0101 [41], Bet v 1.0101 [23], Cor a 1.0401, Cor a 1.0402, Cor a 1.0403, Cor a 1.0404 [31], Mal d 1.0101 [42], and Pru p 1.0101 [43]. The protocols for Ara h 8.0101 (accession no. AAQ91847), Pru d 1.0101 (accession no. ABW99634), Pru p 1.0201 (accession no. AJE61290), and Pru p 1.0301 (accession no. AJE61291) were adapted from [43]. Amino acid sequences are provided in the Supporting Figure S1. All NMR samples contained 20 mM sodium phosphate buffer (pH: 6.9), 9% D_2O, and a varying amount of ^{15}N-labeled protein (0.2–0.8 mM). Only Bet v 1.0101 was dissolved in 5 mM sodium phosphate buffer (pH: 8.0) (9% D_2O). The Cor a 1.040x samples, in addition, contained 2 mM dithiothreitol (DTT).

4.2. NMR Relaxation-Dispersion Experiments

Backbone amide ^{15}N relaxation-dispersion experiments for each 0.5 mM protein sample were recorded at 25 °C on a 700 MHz Bruker Avance Neo spectrometer equipped with a Prodigy CryoProbe, using sensitivity-enhanced Carr–Purcell–Meiboom–Gill (CPMG) sequences [44], with the continuous-wave decoupling of ^1H during the CPMG period [45]. Different CPMG field strengths (ν_{CPMG}) of 33.3, 66.7, 100.0, 133.3, 166.7, 200.0, 266.7, 333.3, 466.7, 600.0, 733.3, and 933.3 Hz were applied, with repeat experiments at 66.7 Hz and 600.0 Hz. The length of the CPMG pulse train was set to T_{relax} = 30 ms. Partial peak volumes (intensities in 5 × 5 grids) were analyzed with nmrDraw [46] to determine the effective relaxation rates: $R_{2,eff} = -1/T_{relax} \cdot \ln(I/I_0)$, with I being the intensity at a given ν_{CPMG}, and I_0 being the reference intensity with the T_{relax} set to 0. Uncertainties in relaxation rates (σ) were calculated from the repeat experiments. The obtained relaxation-dispersion profiles were fitted with an in-house MATLAB script by using the equation derived by Baldwin [32] to extract the per-residue relaxation-dispersion amplitudes ($\Delta R_{2,eff}$) as $\Delta R_{2,eff} = R_{2,eff}$ ($\nu_{CPMG} = 0$) − $R_{2,eff}$ ($\nu_{CPMG} = \infty$). Here, $R_{2,eff}$ ($\nu_{CPMG} = \infty$) and $R_{2,eff}$ ($\nu_{CPMG} = 0$) refer to the extrapolated $R_{2,eff}$ values at infinite and zero ν_{CPMG} field strengths, respectively.

For Ara h 8.0101, the 700 MHz relaxation-dispersion data were complemented by experiments recorded at 600 MHz, using a Bruker Avance II+ spectrometer equipped with a Prodigy CryoProbe. The simultaneous fitting of the data at the two magnetic field strengths was performed. The 600 MHz and 700 MHz relaxation-dispersion profiles were fitted to global two-site exchange models, assuming uniform exchange-rate constants (k_{ex}) for the transition between two states (A and B), and uniform populations of states A and B (p_A and p_B, with $p_A = 1 - p_B$) for all residues in the fit, but residue-specific values of $|\Delta\delta_{ex}|$. In fits assuming the presence of two distinct exchange processes, residues with relaxation-dispersion amplitudes ($\Delta R_{2,eff}$) > 1 s^{-1} at 700 MHz were grouped into those belonging to β-strands and loops (residues Lys37, Ser38, Val39, Ile41, Glu43, Gly47, Thr50, Lys53, Thr55, Ile56, Leu66, His67, Asp73, Glu74, Val84, Val87, Ala88, Leu89, Thr97, Phe98, Glu99, Thr100, Lys101, Leu102, Thr115, Tyr118) or α-helices (Ala16, Lys17, Asn20, Ala21, Met22, Lys23, Asp24, Ala25, Ser27, Lys31, Ile33, Asp123, Glu128, Gly134, Lys137, Gly138, Glu139, Ile145, Glu146, Tyr156). Uncertainties in the exchange parameters were estimated via a Monte Carlo approach, in which 100 synthetic datasets were generated using the exchange parameters obtained by the global fits described above, along with the experimental error in the rates (σ). Data fitting was repeated for all datasets. Errors quoted in the paper are standard deviations in fitted exchange parameters that were obtained in this procedure.

Ligand-observed ^1H relaxation-dispersion experiments were performed as described in [34] to determine the binding kinetics of the ligand (-)-epicatechin on the 600 MHz spectrometer at 25 °C. Samples of Ara h 8.0101 at different concentrations (0, 40, 80, 120,

160, 200, and 240 μM) were prepared in 20 mM sodium phosphate buffer (pH 6.9), 9% D_2O, and 2 mM (-)-epicatechin. The length of the CPMG element (T_{relax}) was set to 60 ms, and the numbers of n cycles were set to 4, 8, 16, 24, 32, 40, 48, 56, 64, 72, 80, 88, 96, 104, 112, and 120, which resulted in ν_{CPMG} values of 66.7, 133.3, 266.7, 400.0, 533.3, 666.7, 800.0, 933.3, 1066.7, 1200.0, 1333.3, 1466.7, 1600.0, 1733.3, 1866.7, and 2000.0 Hz, respectively. Additionally, repeat experiments at 400.0 Hz, 933.3 Hz, and 1600.0 Hz were recorded. The length of the transverse relaxation filter (T_{filt}) was set to 60 ms, and the number of transients was adjusted between 48 and 512 for the uncomplexed ligand to the highest ligand saturation.

The TopSpin program (4.1.1) was used for the processing and analysis of the data. Baseline correction was performed between 5.5 ppm and 7.5 ppm (covering the ^{1}H resonances of protons ^{1}H-6/8 in (-)-epicatechin that were used for analysis). In each spectrum, the peaks were picked, and the absolute intensity at $\nu_{CPMG} = 0$ (I_0), and any given ν_{CPMG} (I_{CPMG}), was determined. The conversion into effective relaxation rates and the fitting of the data to extract the off-rate of the ligand molecule (k_{off}) were conducted as described [34].

4.3. NMR Titration Experiments

The titration of Ara h 8.0101 with (-)-epicatechin was performed by the stepwise addition of the (-)-epicatechin (90 mM stock solution in dimethyl sulfoxide (DMSO)) to 450 μL of a 0.2 mM ^{15}N-labeled protein sample at 25 °C. The molar ratios of the Ara h 8.0101:(-)-epicatechin were 1:30, 1:28, 1:26, 1:24, 1:22, 1:20, 1:18, 1:16, 1:14, 1:12, 1:10, 1:8, 1:6, 1:4, 1:2, 1:1, and 2:1. The final concentration of dimethyl sulfoxide was below 6.67% (*v/v*). For each ratio, a ^{1}H-^{15}N-HSQC spectrum was recorded on the 700 MHz spectrometer. To identify amino acid residues that are involved in (-)-epicatechin binding, the combined amide chemical-shift perturbation was calculated as $\Delta\delta_{obs} = [((\Delta\delta_H)^2 + (\Delta\delta_N/5)^2)/2]^{1/2}$, where $\Delta\delta_H$ and $\Delta\delta_N$ are the chemical-shift changes observed for ^{1}H and ^{15}N, respectively [47]. A ^{1}H-^{15}N-HSQC spectrum was recorded by adding the maximum amount of DMSO (30 μL) to a 0.2 mM ^{15}N-Ara h 8.0101 sample to evaluate the effect of DMSO on the observed chemical shifts ($\Delta\delta_{DMSO}$). The observed shift differences ($\Delta\delta_{obs}$) were then corrected for the DMSO, assuming a linear relation between the DMSO concentration and $\Delta\delta_{DMSO}$. A threshold value of 0.1 ppm was used to identify amino acid residues, which are affected by (-)-epicatechin binding. Residues exceeding this threshold were further used for determining the dissociation constant (K_D) by the nonlinear least-squares fitting [33] of the $\Delta\delta_{obs}$, with an in-house written MATLAB script, assuming the binding of one equivalent of (-)-epicatechin per Ara h 8.0101 molecule, in agreement with the X-ray crystal structure [30].

4.4. In Silico Methods

For the structural comparison of the thirteen selected PR-10 allergens, we used NMR solution structures (Mal d 1.0101 (PDB: 5MMU), Pru p 1.0101 (PDB: 6Z98), and Cor a 1.040x (PDB: 6Y3H, 6Y3I, 6Y3K, 6Y3L)), or X-ray structures (Bet v 1.0101 (PDB: 4A88) and Ara h 8.0101 (PDB: 4M9B)) and CS-Rosetta structure models [48–50], based on the available backbone amide and side-chain Cα/Cβ NMR resonance assignments for Act c 8.0101 and Act d 8.0101 [41], and for Pru d 1.0101, Pru p 1.0201, and Pru p 1.0301 (unpublished). To classify and differentiate the PR-10 allergens based on their electrostatic and hydrophobic surface profiles, we generated images of the surfaces of all the PR-10 allergens, and we color-coded them based on the YRB highlighting scheme by using PyMol. The YRB highlighting scheme allows for the capture of both the hydrophobic and electrostatic properties of the respective protein surfaces [25]. This was performed by modifying the YRB script published by Hagemans et al. to not only be able to differentiate between the charged and hydrophobic regions, but also to include the H-bond donor/acceptor capacity in our coloring scheme. The clustering of these images was performed by using the image module implemented in Keras (Chollet, F., and others (2015). Keras. GitHub. Retrieved from https://github.com/fchollet/keras, accessed on 30 October 2021). The images were then preprocessed and converted into an array consisting of the features extracted during this process. We then applied k-means clustering and defined 3 clusters [51].

Supplementary Materials: The supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms23158252/s1.

Author Contributions: Conceptualization, S.F., R.E., J.U., R.Z. and M.T.; validation, S.F., R.E., J.U., R.Z. and M.T.; formal analysis, S.F. and R.E., J.U., R.Z., M.T.; investigation, S.F. and R.E., J.U., R.Z.; resources, M.T. and K.R.L.; writing—original draft preparation, S.F.; writing—review and editing, S.F., R.Z., M.L.F.-Q. and M.T.; visualization, S.F., M.L.F.-Q. and M.T.; supervision, M.T. and K.R.L.; funding acquisition, M.T. and K.R.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Austrian Science Fund FWF (P26849 to M.T., P30737 to K.R.L.), and by the Austrian Research Promotion Agency FFG (West Austrian BioNMR 858017).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mills, E.N.C.; Shewry, P.R. *Plant Food Allergens*; Blackwell Science: Hoboken, NJ, USA, 2004.
2. Van Loon, L.C.; Pierpoint, W.S.; Boller, T.; Conejero, V. Recommendations for naming plant pathogenesis-related proteins. *Plant Mol. Biol. Rep.* **1994**, *12*, 245–264. [CrossRef]
3. van Loon, L.C.; Rep, M.; Pieterse, C.M. Significance of inducible defense-related proteins in infected plants. *Annu. Rev. Phytopathol.* **2006**, *44*, 135–162. [CrossRef] [PubMed]
4. Somssich, I.E.; Schmelzer, E.; Bollmann, J.; Hahlbrock, K. Rapid activation by fungal elicitor of genes encoding "pathogenesis-related" proteins in cultured parsley cells. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 2427–2430. [CrossRef] [PubMed]
5. Fernandes, H.; Michalska, K.; Sikorski, M.; Jaskolski, M. Structural and functional aspects of PR-10 proteins. *FEBS J.* **2013**, *280*, 1169–1199. [CrossRef]
6. Ipsen, H.; Lowenstein, H. Isolation and immunochemical characterization of the major allergen of birch pollen (*Betula verrucosa*). *J. Allergy Clin. Immunol.* **1983**, *72*, 150–159. [CrossRef]
7. Eriksson, N.E.; Formgren, H.; Svenonius, E. Food hypersensitivity in patients with pollen allergy. *Allergy* **1982**, *37*, 437–443. [CrossRef]
8. Halmepuro, L.; Vuontela, K.; Kalimo, K.; Bjorksten, F. Cross-reactivity of IgE antibodies with allergens in birch pollen, fruits and vegetables. *Int. Arch. Allergy Appl. Immunol.* **1984**, *74*, 235–240. [CrossRef]
9. Ebner, C.; Birkner, T.; Valenta, R.; Rumpold, H.; Breitenbach, M.; Scheiner, O.; Kraft, D. Common epitopes of birch pollen and apples—Studies by western and northern blot. *J. Allergy Clin. Immunol.* **1991**, *88*, 588–594. [CrossRef]
10. Ebner, C.; Hirschwehr, R.; Bauer, L.; Breiteneder, H.; Valenta, R.; Ebner, H.; Kraft, D.; Scheiner, O. Identification of allergens in fruits and vegetables: IgE cross-reactivities with the important birch pollen allergens Bet v 1 and Bet v 2 (birch profilin). *J. Allergy Clin. Immunol.* **1995**, *95*, 962–969. [CrossRef]
11. Mari, A.; Ballmer-Weber, B.K.; Vieths, S. The oral allergy syndrome: Improved diagnostic and treatment methods. *Curr. Opin. Allergy Clin. Immunol.* **2005**, *5*, 267–273. [CrossRef]
12. Kleine-Tebbe, J.; Vogel, L.; Crowell, D.N.; Haustein, U.F.; Vieths, S. Severe oral allergy syndrome and anaphylactic reactions caused by a Bet v 1- related PR-10 protein in soybean, SAM22. *J. Allergy Clin. Immunol.* **2002**, *110*, 797–804. [CrossRef] [PubMed]
13. Le, T.M.; van Hoffen, E.; Lebens, A.F.; Bruijnzeel-Koomen, C.A.; Knulst, A.C. Anaphylactic versus mild reactions to hazelnut and apple in a birch-endemic area: Different sensitization profiles? *Int. Arch. Allergy Immunol.* **2013**, *160*, 56–62. [CrossRef] [PubMed]
14. Geroldinger-Simic, M.; Zelniker, T.; Aberer, W.; Ebner, C.; Egger, C.; Greiderer, A.; Prem, N.; Lidholm, J.; Ballmer-Weber, B.K.; Vieths, S.; et al. Birch pollen-related food allergy: Clinical aspects and the role of allergen-specific IgE and IgG4 antibodies. *J. Allergy Clin. Immunol.* **2011**, *127*, 616–622. [CrossRef] [PubMed]
15. Eidelpes, R.; Hofer, F.; Röck, M.; Führer, S.; Kamenik, A.S.; Liedl, K.R.; Tollinger, M. Structure and zeatin binding of the peach allergen *Pru p 1*. *J. Agric. Food Chem.* **2021**, *69*, 8120–8129. [CrossRef]
16. Morris, J.S.; Caldo, K.M.P.; Liang, S.; Facchini, P.J. PR10/Bet v1-like proteins as novel contributors to plant biochemical diversity. *ChemBioChem* **2021**, *22*, 264–287. [CrossRef]
17. Aglas, L.; Soh, W.T.; Kraiem, A.; Wenger, M.; Brandstetter, H.; Ferreira, F. Ligand binding of PR-10 proteins with a particular focus on the Bet v 1 allergen family. *Curr. Allergy Asthma Rep.* **2020**, *20*, 25. [CrossRef]
18. Mogensen, J.E.; Wimmer, R.; Larsen, J.N.; Spangfort, M.D.; Otzen, D.E. The major birch allergen, Bet v 1, shows affinity for a broad spectrum of physiological ligands. *J. Biol. Chem.* **2002**, *277*, 23684–23692. [CrossRef]

19. Chruszcz, M.; Chew, F.T.; Hoffmann-Sommergruber, K.; Hurlburt, B.K.; Mueller, G.A.; Pomes, A.; Rouvinen, J.; Villalba, M.; Wohrl, B.M.; Breiteneder, H. Allergens and their associated small molecule ligands-their dual role in sensitization. *Allergy* **2021**, *76*, 2367–2382. [CrossRef]
20. Soh, W.T.; Aglas, L.; Mueller, G.A.; Gilles, S.; Weiss, R.; Scheiblhofer, S.; Huber, S.; Scheidt, T.; Thompson, P.M.; Briza, P.; et al. Multiple roles of Bet v 1 ligands in allergen stabilization and modulation of endosomal protease activity. *Allergy* **2019**, *74*, 2382–2393. [CrossRef]
21. Führer, S.; Kamenik, A.S.; Zeindl, R.; Nothegger, B.; Hofer, F.; Reider, N.; Liedl, K.R.; Tollinger, M. Inverse relation between structural flexibility and IgE reactivity of Cor a 1 hazelnut allergens. *Sci. Rep.* **2021**, *11*, 4173. [CrossRef] [PubMed]
22. Grutsch, S.; Fuchs, J.E.; Ahammer, L.; Kamenik, A.S.; Liedl, K.R.; Tollinger, M. Conformational flexibility differentiates naturally occurring Bet v 1 isoforms. *Int. J. Mol. Sci.* **2017**, *18*, 1192. [CrossRef] [PubMed]
23. Grutsch, S.; Fuchs, J.E.; Freier, R.; Kofler, S.; Bibi, M.; Asam, C.; Wallner, M.; Ferreira, F.; Brandstetter, H.; Liedl, K.R.; et al. Ligand binding modulates the structural dynamics and compactness of the major birch pollen allergen. *Biophys. J.* **2014**, *107*, 2972–2981. [CrossRef] [PubMed]
24. Moraes, A.H.; Asam, C.; Almeida, F.C.L.; Wallner, M.; Ferreira, F.; Valente, A.P. Structural basis for cross-reactivity and conformation fluctuation of the major beech pollen allergen Fag s 1. *Sci. Rep.* **2018**, *8*, 10512. [CrossRef]
25. Hagemans, D.; van Belzen, I.A.; Moran Luengo, T.; Rudiger, S.G. A script to highlight hydrophobicity and charge on protein surfaces. *Front. Mol. Biosci.* **2015**, *2*, 56. [CrossRef] [PubMed]
26. Holm, J.; Ferreras, M.; Ipsen, H.; Wurtzen, P.A.; Gajhede, M.; Larsen, J.N.; Lund, K.; Spangfort, M.D. Epitope grafting, re-creating a conformational Bet v 1 antibody epitope on the surface of the homologous apple allergen Mal d 1. *J. Biol. Chem.* **2011**, *286*, 17569–17578. [CrossRef]
27. Roulias, A.; Pichler, U.; Hauser, M.; Himly, M.; Hofer, H.; Lackner, P.; Ebner, C.; Briza, P.; Bohle, B.; Egger, M.; et al. Differences in the intrinsic immunogenicity and allergenicity of Bet v 1 and related food allergens revealed by site-directed mutagenesis. *Allergy* **2014**, *69*, 208–215. [CrossRef]
28. Schmalz, S.; Mayr, V.; Shosherova, A.; Gepp, B.; Ackerbauer, D.; Sturm, G.; Bohle, B.; Breiteneder, H.; Radauer, C. Isotype-specific binding patterns of serum antibodies to multiple conformational epitopes of Bet v 1. *J. Allergy Clin. Immunol.* **2022**, *149*, 1786–1794.e12. [CrossRef] [PubMed]
29. Mittermaier, A.K.; Kay, L.E. Observing biological dynamics at atomic resolution using NMR. *Trends Biochem. Sci.* **2009**, *34*, 601–611. [CrossRef]
30. Hurlburt, B.K.; Offermann, L.R.; McBride, J.K.; Majorek, K.A.; Maleki, S.J.; Chruszcz, M. Structure and function of the peanut panallergen Ara h 8. *J. Biol. Chem.* **2013**, *288*, 36890–36901. [CrossRef]
31. Führer, S.; Zeindl, R.; Tollinger, M. NMR resonance assignments of the four isoforms of the hazelnut allergen Cor a 1.04. *Biomol. NMR Assign.* **2020**, *14*, 45–49. [CrossRef]
32. Baldwin, A.J. An exact solution for R2,eff in CPMG experiments in the case of two site chemical exchange. *J. Magn. Reson.* **2014**, *244*, 114–124. [CrossRef] [PubMed]
33. Fielding, L.; Rutherford, S.; Fletcher, D. Determination of protein-ligand binding affinity by NMR: Observations from serum albumin model systems. *Magn. Reson. Chem.* **2005**, *43*, 463–470. [CrossRef] [PubMed]
34. Moschen, T.; Grutsch, S.; Juen, M.A.; Wunderlich, C.H.; Kreutz, C.; Tollinger, M. Measurement of ligand-target residence times by (1)H relaxation dispersion NMR spectroscopy. *J. Med. Chem.* **2016**, *59*, 10788–10793. [CrossRef] [PubMed]
35. Kofler, S.; Asam, C.; Eckhard, U.; Wallner, M.; Ferreira, F.; Brandstetter, H. Crystallographically mapped ligand binding differs in high and low IgE binding isoforms of birch pollen allergen bet v 1. *J. Mol. Biol.* **2012**, *422*, 109–123. [CrossRef] [PubMed]
36. Orozco-Navarrete, B.; Kaczmarska, Z.; Dupeux, F.; Garrido-Arandia, M.; Pott, D.; Perales, A.D.; Casanal, A.; Marquez, J.A.; Valpuesta, V.; Merchante, C. Structural bases for the allergenicity of Fra a 1.02 in strawberry fruits. *J. Agric. Food Chem.* **2020**, *68*, 10951–10961. [CrossRef]
37. Casanal, A.; Zander, U.; Munoz, C.; Dupeux, F.; Luque, I.; Botella, M.A.; Schwab, W.; Valpuesta, V.; Marquez, J.A. The strawberry pathogenesis-related 10 (PR-10) Fra a proteins control flavonoid biosynthesis by binding to metabolic intermediates. *J. Biol. Chem.* **2013**, *288*, 35322–35332. [CrossRef]
38. Pomes, A.; Mueller, G.A.; Chruszcz, M. Structural aspects of the allergen-antibody interaction. *Front. Immunol.* **2020**, *11*, 2067. [CrossRef]
39. Ahammer, L.; Grutsch, S.; Kamenik, A.S.; Liedl, K.R.; Tollinger, M. Structure of the major apple allergen Mal d 1. *J. Agric. Food Chem.* **2017**, *65*, 1606–1612. [CrossRef] [PubMed]
40. Focke, M.; Linhart, B.; Hartl, A.; Wiedermann, U.; Sperr, W.R.; Valent, P.; Thalhamer, J.; Kraft, D.; Valenta, R. Non-anaphylactic surface-exposed peptides of the major birch pollen allergen, Bet v 1, for preventive vaccination. *Clin. Exp. Allergy* **2004**, *34*, 1525–1533. [CrossRef]
41. Zeindl, R.; Tollinger, M. NMR resonance assignments of the PR-10 allergens Act c 8 and Act d 8 from golden and green kiwifruit. *Biomol. NMR Assign.* **2021**, *15*, 367–371. [CrossRef]
42. Ahammer, L.; Grutsch, S.; Tollinger, M. NMR resonance assignments of the major apple allergen Mal d 1. *Biomol. NMR Assign.* **2016**, *10*, 287–290. [CrossRef] [PubMed]
43. Führer, S.; Trimmel, S.; Breuker, K.; Tollinger, M. NMR resonance assignments of the pathogenesis-related peach allergen Pru p 1.0101. *Biomol. NMR Assign.* **2019**, *13*, 127–130. [CrossRef] [PubMed]

44. Tollinger, M.; Skrynnikov, N.R.; Mulder, F.A.; Forman-Kay, J.D.; Kay, L.E. Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc.* **2001**, *123*, 11341–11352. [CrossRef] [PubMed]
45. Hansen, D.F.; Vallurupalli, P.; Kay, L.E. An improved 15N relaxation dispersion experiment for the measurement of millisecond time-scale dynamics in proteins. *J. Phys. Chem. B* **2008**, *112*, 5898–5904. [CrossRef]
46. Delaglio, F.; Grzesiek, S.; Vuister, G.W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **1995**, *6*, 277–293. [CrossRef]
47. Williamson, M.P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16. [CrossRef]
48. Shen, Y.; Bryan, P.N.; He, Y.; Orban, J.; Baker, D.; Bax, A. De novo structure generation using chemical shifts for proteins with high-sequence identity but different folds. *Protein Sci.* **2010**, *19*, 349–356. [CrossRef]
49. Shen, Y.; Lange, O.; Delaglio, F.; Rossi, P.; Aramini, J.M.; Liu, G.; Eletsky, A.; Wu, Y.; Singarapu, K.K.; Lemak, A.; et al. Consistent blind protein structure generation from NMR chemical shift data. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4685–4690. [CrossRef]
50. Shen, Y.; Vernon, R.; Baker, D.; Bax, A. De novo protein structure generation from incomplete chemical shift assignments. *J. Biomol. NMR* **2009**, *43*, 63–78. [CrossRef]
51. Jin, X.; Han, J. K-Means Clustering. In *Encyclopedia of Machine Learning*; Sammut, C., Webb, G.I., Eds.; Springer: Boston, MA, USA, 2010; pp. 563–564.

Supporting Information

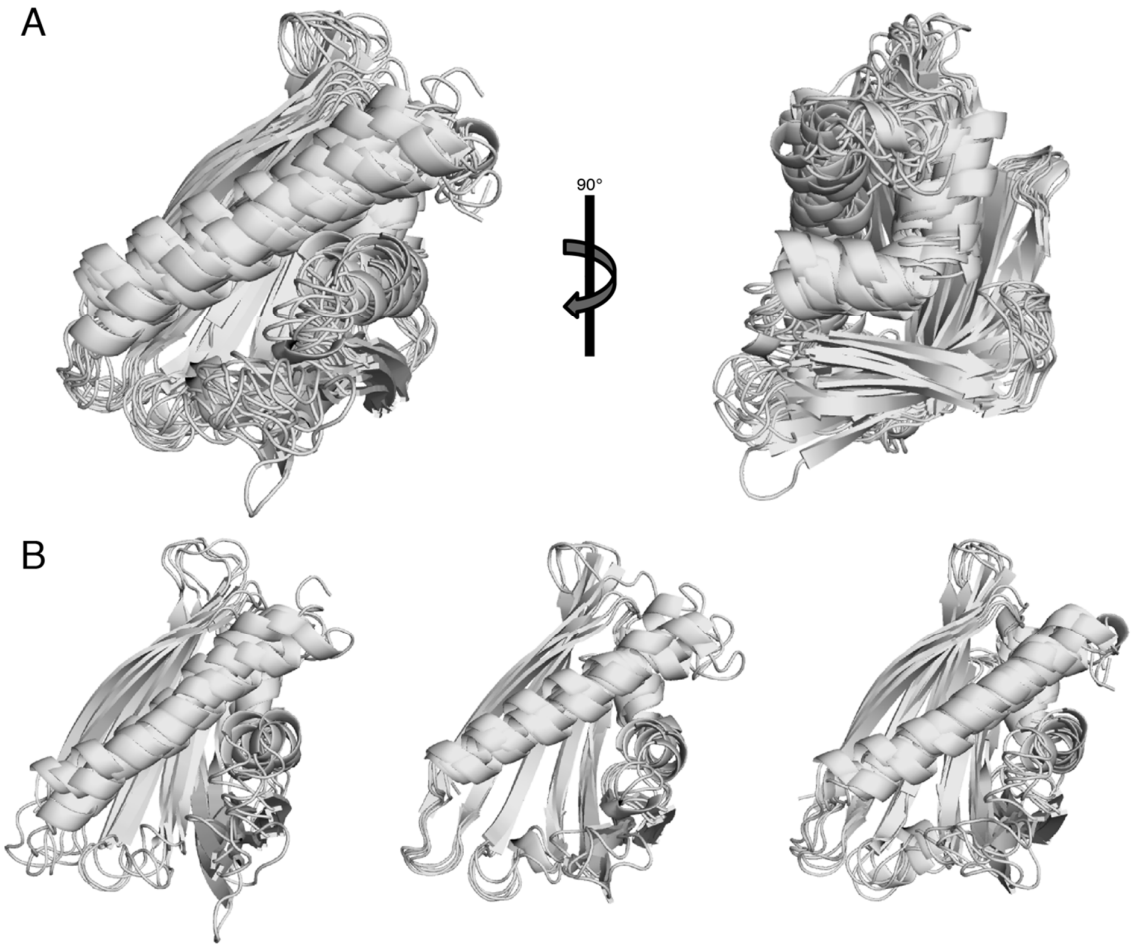
The structural flexibility of PR-10 food allergens

Sebastian Führer, Jana Unterhauser, Ricarda Zeindl, Reiner Eidelpes, Monica L. Fernández-Quintero, Klaus R. Liedl and Martin Tollinger

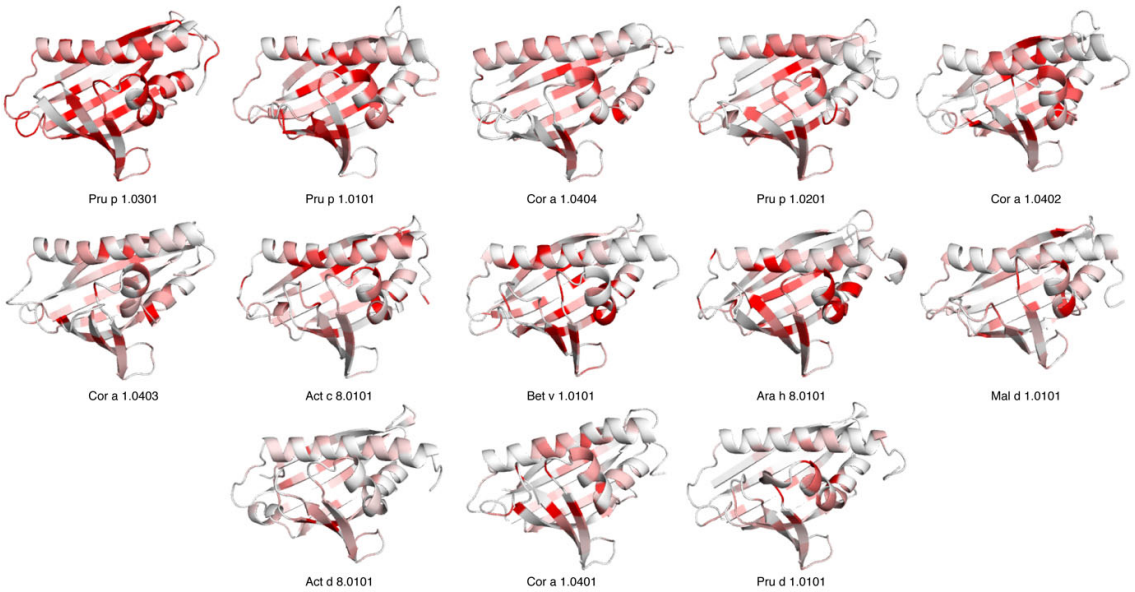
```
Bet v 1.0101 GVFNYETETT SVIPAARLFK AFILDGDNLF PKVAPQAISS VENIEGNGGP GTIKKISFPE GFPFKYVKDR VDEVDHTNFK 80
Mal d 1.0101 GVYTFENEFT SEIPPSRLFK AFVLDADNLI PKIAPQAIKQ AEILEGNGGP GTIKKITFGE GSQYGYVKHR IDSIDEASYS 80
Act c 8.0101 GVVTYDMEIP SKVPPVKLYK AFILDGDTLV PKVLPHAIKC VKILEGDGCA GTIKEVTFGE GSHHKCVKQR VDAIDKDNLT 80
Act d 8.0101 GAITYDMEIP SSISAEKMFK AFVLDGDTII PKALPHAITG VQTLEGDGGV GTIKLTTFGE GSVHKSVKHR IDGLDKENFT 80
Pru d 1.0101 GVFTYESEFT SEIPPPRLFK AFVLDADNLV PKIAPQAIKH SEILEGDGGP GTIKKITFGE GSQYGYVKHK IDSIDKENHS 80
Pru p 1.0101 GVFTYESEFT SEIPPPRLFK AFVLDADNLV PKIAPQAIKH SEILEGDGGP GTIKKITFGE GSQYGYVKHK IDSIDKENHS 80
Pru p 1.0201 GVFTYSDEST SVIPPPRLFK ALVLEADTLI PKIAPQSVKS AEIVEGDGGV GTIKKISFGE GSHYSYVKHR IDGLDKDNFV 80
Pru p 1.0301 GVFTYETEFT SVIPPEKLFK AFILDADNLI PKIAPTAVKD TEILEGDGGV GTIKKVTFGE GSQYGYVKHR IDGIDKDNLT 80
Ara h 8.0101 GVFTFEDEIT STVPPAKLYN AM-KDADSIT PKIID-DVKS VEIVEGNGGP GTIKKLTIVE DGETKFILHK VESIDEANYA 78
Cor a 1.0401 GVFCYEDEAT SVIPPARLFK SFVLDADNLI PKVAPQHFTS AENLEGNGGP GTIKKITFAE GNEFKYMKHK VEEIDHANFK 80
Cor a 1.0402 GVFSYEDEAT SVIPPARLFK SFVLDADNLI PKVAPQHFTG AENLEGNGGP GTIKKITFAE GSEFKYMKHK VEEIDHANFK 80
Cor a 1.0403 GVFCYEDEAT SVIPPARLFK SFVLDADNLI PKVAPQHFTG AENLEGNGGP GTIKKITFAE GSEFKYMKHK VEEIDHANFK 80
Cor a 1.0404 GVFSYEDEAT SVIPPARLFK SFVLDADNLI PKVAPQHFTS AENLEGNGGP GTIKKITFAE GNEFKYMKHK VEEIDHANFK 80
             *.  :. *   * :   :::: ::  :.*.:  **     ..   : :**:*   ****  :: * .     : .: :: :*. .

Bet v 1.0101 YNYSVIEGGP IGDTLEKISN EIKIVATPD- GGSILKISNK YHTKGDHEVK AEQVKASKEM GETLLRAVES YLLAHSDAYN 159
Mal d 1.0101 YSYTLIEGDA LTDTIEKISY ETKLVACGS- -GSTIKSISH YHTKGNIEIK EEHVKVGKEK AHGLFKLIES YLKDHPDAYN 158
Act c 8.0101 YSYTIIEGDV LAEKFESISY HIKIVACPD- GGSICKNRSI YTTKGDCKVS EEEIKLGKEK AAEIFKALEA YLLANPDYC- 158
Act d 8.0101 YSYSIIEGGA L-DVFESISY HIKIVATPD- GGCICKNRSI YTPKCDAQVS EEEIKAGKER ASGIFKKVEA YLLANPDC-- 156
Pru d 1.0101 YSYTLIEGDA LGDNLEKISY ETKLVASPS- GGSIIKSISH YHTKGDVEIK EEHVKAGKEK ASNLFKLIET YLKGHPDAYN 159
Pru p 1.0101 YSYTLIEGDA LGDNLEKISY ETKLVASPS- GGSIIKSTSH YHTKGDVEIK EEHVKAGKEK ASNLFKLIET YLKGHPDAYN 159
Pru p 1.0201 YNYTLVEGDA LSDKVEKITY EIKLVASAD- GGSIIKSTSN YHTKGDVEIK EEDVKAGKEK ATGLFKLIEN YLVANPDAYN 159
Pru p 1.0301 YSYTLIEGDA LSDVIEKIVY DIKLVASPN- GGSIVKTISH YHTKGDVEIK EEQVKAGKEK AAGLFKLVEG YLLANPDAYN 159
Ara h 8.0101 YNYSVVGGVA LPPTAEKITF ETKLVEGPN- GGSIGKLTLK YHTKGDAKPD EEELKKGKAK GEGLFRAIEG YVLANPTQY- 156
Cor a 1.0401 YCYSIIEGGP LGHTLEKISY EIKMAAAPHG GGSILKITSK YHTKGNASIN EEEIKAGKEK AAGLFKAVEA YLLAHPDAYC 160
Cor a 1.0402 YCYSIIEGGP LGHTLEKISY EIKMAAAPHG GGSILKITSK YHTKGNASIS EEEIKAGKEK AAGLFKAVEA YLLAHPDTYC 160
Cor a 1.0403 YCYSIIEGGP LGHTLEKISY EIKMAAAPHG GGSILKITSK YHTKGNASIS EEEIKAGKEK AAGLFKAVEA YLLAHPDTYC 160
Cor a 1.0404 YCYSIIEGGP LGHTLEKIPY EIKMAAAPHG GGSILKITSK YHTKGNASIN EEEIKAGKEK AAGLFKAVEA YLLAHPDAYC 160
             * *::: *   :    *.*   . *:.       *.  *     *  * : . .  *.:* .*   .  ::: :*  *:  :
```

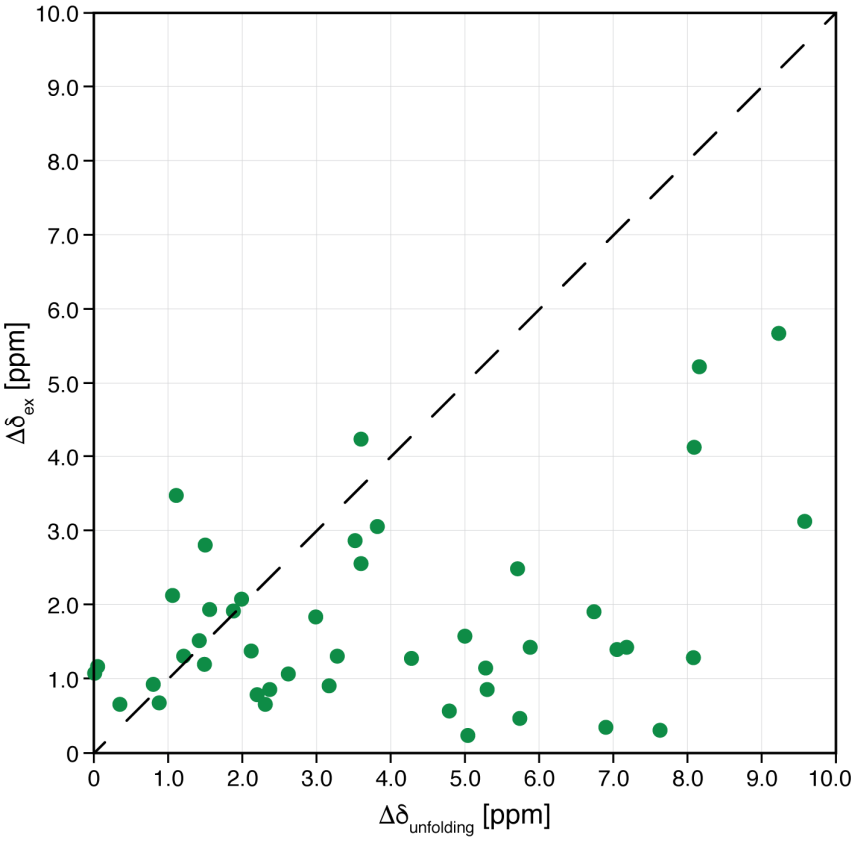
Supporting Figure S1. Sequence alignment of PR-10 allergens. Identical amino acid residues are labeled with asterisks, conserved residues with colons and semiconserved residues with dots. Positions of amino acid residues that are not present in a specific allergen are indicated by a dash.



Supporting Figure S2. Structural comparison of the thirteen PR-10 allergens in this study. (**A**) Overlay of all structures. (**B**) Overlay of Cor a 1.0101, Cor a 1.0402, Cor a 1.0403, Cor a 1.0404 (left) compared to Pru p 1.0101, Pru p 1.0201, Pru p 1.0301, Pru d 1.0101 (middle) and Bet v 1.0101, Mal d 1.0101, Ara h 8.0101, Act c 8.0101, Act d 8.0101 (right).



Supporting Figure S3. Structural flexibility of the thirteen PR-10 allergens. Relaxation dispersion amplitudes ($\Delta R_{2,eff}$ values) exceeding 1 s^{-1} at 700 MHz are plotted on the protein backbones, using a color gradient from rigid (white) to more flexible (red). The same color gradient as in Figure 3 was used for all proteins.



Supporting Figure S4. Comparison of ^{15}N chemical shift differences, $\Delta\delta_{ex}$, obtained from the backbone amide ^{15}N relaxation dispersion data of Ara h 8.0101 with ^{15}N secondary chemical shifts, $\Delta\delta_{unfolding}$, as expected for complete unfolding of the protein backbone. Values of $\Delta\delta_{unfolding}$ were calculated as the difference between observed ^{15}N chemical shifts and predicted ^{15}N chemical shifts for a random coil (Tamiola K, Acar B, Mulder FAA, *J Am Chem Soc* 2010 (132), 18000-18003. doi: 10.1021/ja105656t). The dashed, drawn with a slope of 1.0, line indicates the ^{15}N chemical shift differences, $\Delta\delta_{ex}$, that are expected for complete unfolding of the backbone to a random coil.

Helices		Sheet	
Ala16	0.30 ± 0.16	Lys37	1.51 ± 0.34
Lys 17	0.46 ± 0.06	Ser38	5.22 ± 0.51
Asn20	0.85 ± 0.07	Val39	2.80 ± 0.25
Ala21	1.83 ± 0.14	Ile41	1.30 ± 0.23
Met22	0.78 ± 0.06	Glu43	1.42 ± 0.27
Lys23	1.39 ± 0.11	Gly47	2.55 ± 0.29
Asp24	0.92 ± 0.07	Thr50	0.23 ± 0.10
Ala25	1.06 ± 0.07	Lys53	1.16 ± 0.25
Ser27	1.57 ± 0.13	Thr55	1.42 ± 0.14
Lys31	2.48 ± 0.21	Ile56	2.07 ± 0.22
Ile33	0.65 ± 0.07	Leu66	3.47 ± 0.24
Asp123	0.90 ± 0.14	His67	3.05 ± 0.16
Glu128	1.37 ± 0.10	Asp73	1.93 ± 0.17
Gly134	1.19 ± 0.08	Glu74	1.14 ± 0.12
Lys137	1.07 ± 0.08	Val84	3.12 ± 1.08
Gly138	0.65 ± 0.19	Val87	1.90 ± 0.26
Glu139	0.85 ± 0.12	Ala88	5.67 ± 0.81
Ile145	0.67 ± 0.08	Leu89	1.30 ± 0.17
Glu146	0.56 ± 0.07	Thr97	4.12 ± 0.90
Tyr156	0.25 ± 0.16	Phe98	4.23 ± 1.04
		Glu99	1.91 ± 0.16
		Thr100	1.27 ± 0.22
		Lys101	2.86 ± 0.30
		Leu102	2.21 ± 0.18
		Thr115	0.34 ± 0.18
		Tyr118	1.28 ± 0.36

Supporting Table S1. Residue-specific ^{15}N chemical shift differences, $\Delta\delta_{ex}$, between major and minor states, obtained from sub-global fits of the backbone amide ^{15}N relaxation dispersion data for amino acid residues in the three α-helices and loop L2 (left) and in the β-sheet and all other loops (right). Data are shown for Ara h 8.0101. For amino acid residues highlighted in blue and orange the experimental relaxation dispersion data are shown in Figure 4.