scientific reports

OPEN

Inverse relation between structural flexibility and IgE reactivity of Cor a 1 hazelnut allergens

Sebastian Führer[1], Anna S. Kamenik[2], Ricarda Zeindl[1], Bettina Nothegger[3], Florian Hofer[2], Norbert Reider[3], Klaus R. Liedl[2] & Martin Tollinger[1]✉

A major proportion of allergic reactions to hazelnuts (*Corylus avellana*) are caused by immunologic cross-reactivity of IgE antibodies to pathogenesis-related class 10 (PR-10) proteins. Intriguingly, the four known isoforms of the hazelnut PR-10 allergen Cor a 1, denoted as Cor a 1.0401–Cor a 1.0404, share sequence identities exceeding 97% but possess different immunologic properties. In this work we describe the NMR solution structures of these proteins and provide an in-depth study of their biophysical properties. Despite sharing highly similar three-dimensional structures, the four isoforms exhibit remarkable differences regarding structural flexibility, hydrogen bonding and thermal stability. Our experimental data reveal an inverse relation between structural flexibility and IgE-binding in ELISA experiments, with the most flexible isoform having the lowest IgE-binding potential, while the isoform with the most rigid backbone scaffold displays the highest immunologic reactivity. These results point towards a significant entropic contribution to the process of antibody binding.

Nut allergies in Europe are predominantly related to hazelnuts[1,2] and walnuts[3]. Immunologic reactions to these food sources are triggered by several specific proteins, with a significant proportion of individuals being affected by class 10 of pathogenesis-related proteins (PR-10)[1]. These allergic reactions result in particular from an initial sensitization to the major birch (*Betula verrucosa*) pollen allergen Bet v 1, a PR-10 protein. Thereafter, immunologic cross-reactivity is developed due to the strong similarity between Bet v 1 and homologous nut allergens[4,5]. Indeed, up to 70% of all birch pollen allergic individuals are affected by such birch pollen-related food allergies (BPRFA), with hazelnuts representing one of the most prevalent triggers[6–9]. In affected patients, consumption of hazelnuts provokes a variety of clinical symptoms, including itching, scratching, and swelling of the mouth and throat[10] as well as in rare cases severe anaphylactic shocks[11,12]. Contrary to other food sources, industrial processing at high temperatures does not prevent IgE reactivity in hazelnut products[13]. Additionally, hazelnut allergens only decompose at much higher temperatures[14,15] than, for example, apple allergens[16].

PR-10 proteins, which are expressed in plants upon environmental or pathogenic stimuli, have a molecular weight of ca. 17.5 kDa and comprise about 160 amino acid residues. These proteins exhibit a canonical fold consisting of a seven stranded antiparallel β-sheet (β1—β7) and three α-helices (α1, α2, α3). The two short, consecutive helices α1 and α2 interrupt the β-sheet between strands β1 and β2 while the long C-terminal helix α3 is located above the β-sheet, creating a large and fairly hydrophobic cavity in the protein interior[17]. The PR-10 proteins of the common hazel (*Corylus avellana*) can be grouped into the isoallergens Cor a 1.01 (hazel pollen)[18], Cor a 1.02 and Cor a 1.03 (hazel leaf)[19], and Cor a 1.04 (hazelnut)[20]. Interestingly, hazelnut Cor a 1.04 allergens are closely related to the birch pollen allergen Bet v 1, with amino acid sequence identities of about 83%[5,21], while hazelnut and hazel pollen allergens share sequence identities of only about 63%. Likewise, the IgE-epitopes of the hazelnut allergens appear to show higher similarity to those of birch pollen than hazel pollen[22], and the C-terminal α-helix of hazelnut Cor a 1.04 allergens contain a main T cell epitope that is not present in birch or hazel pollen allergens[23].

Among each other, the four isoforms of the hazelnut (Cor a 1.0401–Cor a 1.0404) share sequence identities of at least 97%. Nevertheless, these four allergens display strikingly different IgE-binding properties in enzyme allergosorbent tests, with a proposed immunologic ranking Cor a 1.0401 > 02 > 03 > 04[22]. The structural basis for the different IgE-binding properties has remained elusive so far. In this work we present the nuclear magnetic

[1]Institute of Organic Chemistry and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innrain 80/82, 6020 Innsbruck, Austria. [2]Institute of General, Inorganic and Theoretical Chemistry and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innrain 80/82, 6020 Innsbruck, Austria. [3]Department of Dermatology, Venerology and Allergology, Medical University of Innsbruck, Anichstraße 35, 6020 Innsbruck, Austria. ✉email: martin.tollinger@uibk.ac.at

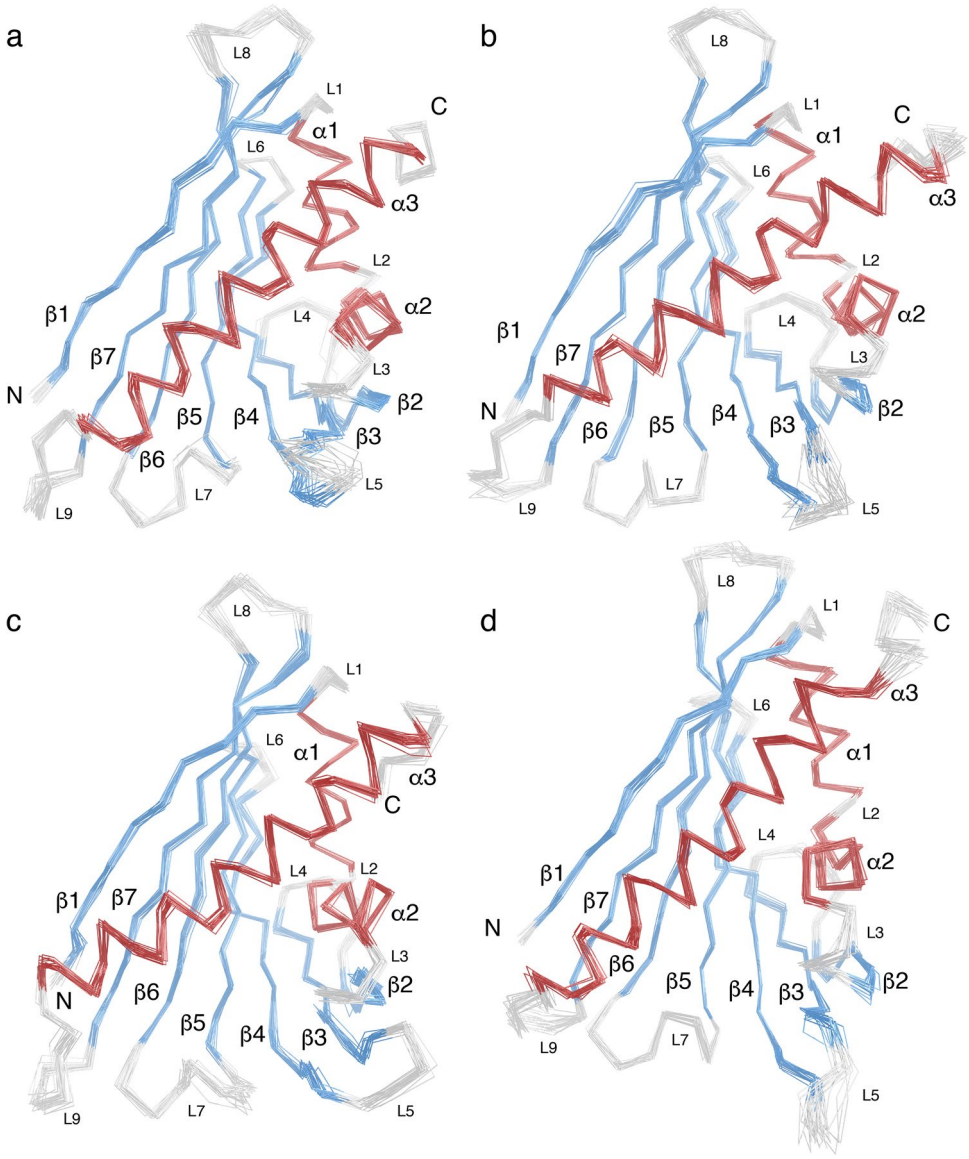


Figure 1. Backbone overlay of the 20 NMR solution structures with the lowest energy of the four hazelnut allergens Cor a 1.0401 (**a**), Cor a 1.0402 (**b**), Cor a 1.0403 (**c**), and Cor a 1.0404 (**d**), with the PDB codes 6Y3H, 6Y3I, 6Y3K, and 6Y3L, respectively. The secondary structure elements are defined as β1 (Val2–Ser11), α1 (Pro15–Leu24), α2 (Ala26–Ala34), β2 (Thr39–Glu45), β3 (Gly51–Ala59), β4 (Phe64–Asp75), β5 (Phe79–Glu87), β6 (Glu96–Ala106), β7 (Gly112–Thr123), and α3 (Glu131–His155). The α-helices are highlighted in red, β-strands in blue, and loops in light gray. Loops L1–L9 as well as N- and C-termini are indicated.

resonance (NMR) solution structures of all four Cor a 1.04 isoforms along with in-depth experimental data regarding the structural flexibility, thermal and temporal stability and IgE-binding properties of these proteins. Our data reveal a clear inverse correlation between structural flexibility and IgE-binding, with the most flexible isoform showing the lowest potential to bind specific IgE.

Results and discussion

The four Cor a 1.04 isoforms have similar structures. To investigate the four hazelnut allergens Cor a 1.0401, Cor a 1.0402, Cor a 1.0403 and Cor a 1.0404 in a comparative manner, we determined their NMR solution structures under identical experimental conditions (Fig. 1). As expected, all four proteins consist of a seven stranded antiparallel β-sheet (β1–β7) and three α-helices (α1, α2, α3), which adopt the canonical PR-10 fold with helices α1 and α2 arranged in a V-shaped manner above the curved β-sheet, acting as support for the long C-terminal helix α3. The NMR structural ensembles of all four isoforms (PDB codes 6Y3H, 6Y3I, 6Y3K, 6Y3L) display high conformational homogeneity and well-defined secondary structure elements. Root-mean-square deviation (RMSD) values of the 20 lowest energy structures are 0.5–0.7 Å for heavy atoms in all cases and 0.4 Å for backbone atoms (Table 1). The four isoforms have very similar three-dimensional structures, with backbone

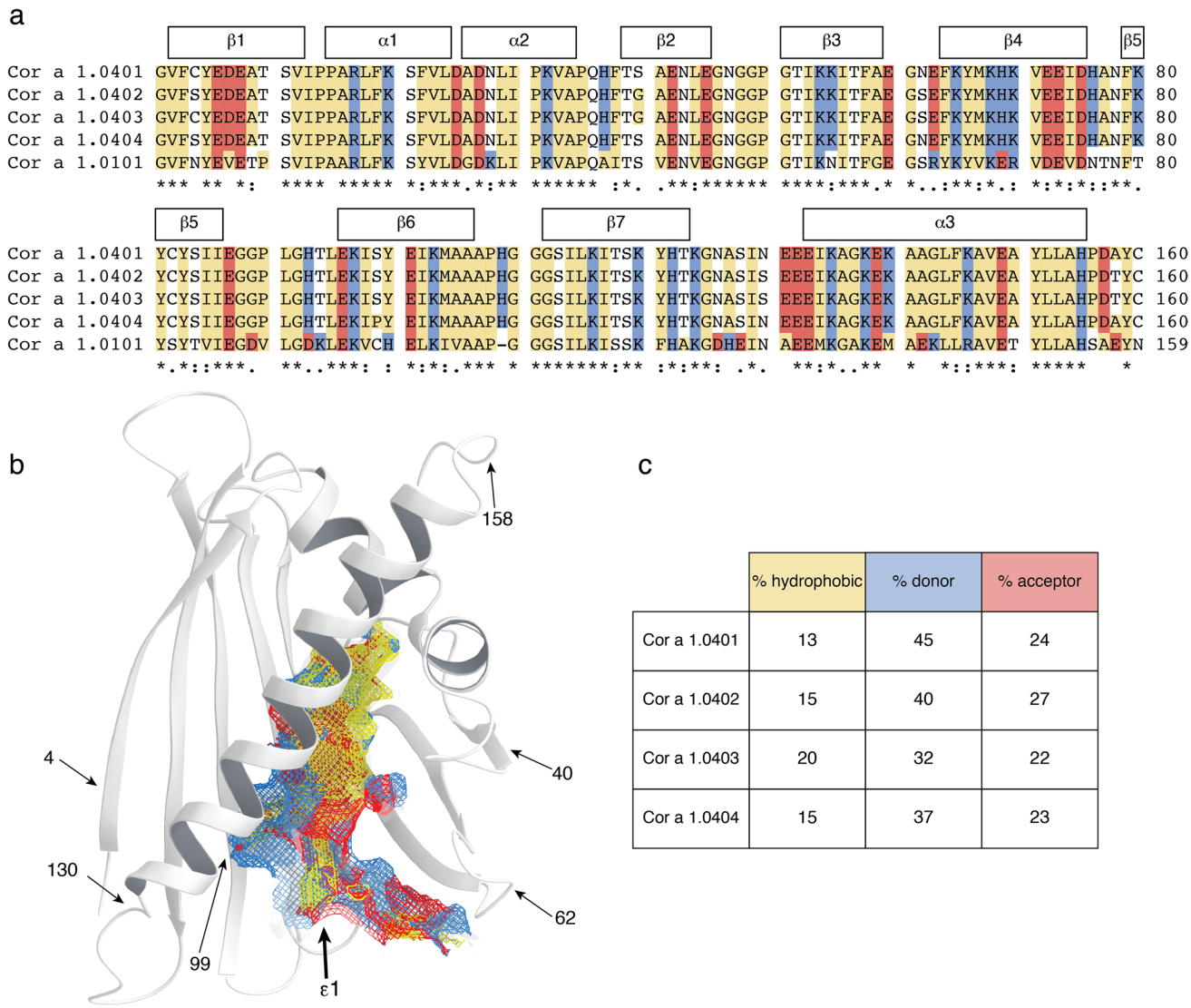
	Cor a 1.0401	Cor a 1.0402	Cor a 1.0403	Cor a 1.0404
PDB ID	6Y3H	6Y3I	6Y3K	6Y3L
Experimental restraints				
Total no. of NOE-based distance restraints[a]	5027	4062	4396	2662
Intraresidue [i = j]	1577	1232	1630	1175
Sequential [|i − j| = 1]	1200	1133	1068	669
Medium range [1 < |i − j| < 5]	900	771	701	327
Long range [|i − j| ≥ 5]	1350	926	997	491
Dihedral angle restraints	267	260	257	255
Hydrogen bond restraints	142	148	140	149
Total no. of restraints	5436	4470	4793	3066
Total no. of restraints per residue	34.0	27.9	30.0	19.2
Long range restraints per residue	8.4	5.8	6.2	3.1
Restraint violations[b]				
Average distance violation	0.032 ± 0.024 Å	0.015 ± 0.014 Å	0.017 ± 0.010 Å	0.020 ± 0.017 Å
Maximal distance violation[c]	0.17 Å	0.092 Å	0.045 Å	0.15 Å
Average dihedral angel violation	0.39 ± 0.90°	0.19 ± 0.17°	0.46 ± 1.00°	0.16 ± 0.15°
Maximal dihedral angle violation[c]	14.10°	0.77°	15.7°	0.73°
RMSD values[d]				
Backbone atoms	0.4 Å	0.4 Å	0.4 Å	0.4 Å
Heavy atoms	0.7 Å	0.5 Å	0.5 Å	0.5 Å
Bond lengths	0.013 Å	0.013 Å	0.013 Å	0.013 Å
Bond angles	2.4°	2.0°	2.1°	2.1°
Ramachandran plot statistics[e]				
Most favored regions	88.3%	96.2%	93.3%	96.2%
Allowed regions	11.7%	3.0%	6.7%	3.3%
Disallowed regions	0.0%	0.7%	0.0%	0.5%
All-atom clashscore[f]	3	0	0	0

Table 1. Statistics and summary of structure determination and refinement. [a]Numbers are given for all residues (Gly1–Cys160). [b]Calculated for all residues as sum over r^{-6}. [c]Largest violation among the 20 lowest energy structures. [d]Pairwise RMSD was calculated among the 20 lowest energy structures. [e]Obtained with the protein structure validation software (PSVS) suite[61]. [f]All-atom clashscore, defined as number of overlaps (≥ 0.4 Å) per 1000 atoms.

RMSD values between them ranging from 2.3 to 3.8 Å, and our Cor a 1.0401 structure compares well to the Cor a 1.0401 structure that has been reported before, with a backbone amide RMSD of 4.4 Å[24]. An overlay of the lowest energy NMR structures of the four Cor a 1.04 isoforms shows that slight variations are present regarding the curvature of the β-sheet and the C-terminal helix α3. Very similar structural variations are also present between other PR-10 allergens, e.g. birch pollen and different food sources (Supplementary Fig. S1 online). Consistently, backbone RMSD values between the four Cor a 1.04 isoforms and the sensitizing allergen from birch pollen, Bet v 1, are all below 4.4 Å. In this regard, the hazelnut isoforms are comparable to other PR-10 food allergens[17].

A sequence comparison of the four Cor a 1.04 isoforms is shown in Fig. 2a. Among each other, between one and six amino acid residues (at positions 4, 40, 62, 99, 130 and 158) are different, with Cor a 1.0401/Cor a 1.0404 (two residues) and Cor a 1.0402/Cor a 1.0403 (one residue) being the most similar isoform pairs. The six variable residues in the four isoforms are distributed over the protein scaffold and not in spatial proximity to each other (Fig. 2b). Intriguingly, Cor a 1.0404 contains a proline residue at position 99 in the center of the β-sheet. The NMR chemical shift data clearly show that Pro99, like all other prolines, possesses trans configuration[25], in accordance with its location in an antiparallel β-sheet. Strand β6 (residues Glu96-Ala106) containing Pro99 and the adjacent strand β7 (Gly112-Thr123) display only slightly reduced β-strand propensities in Cor a 1.0404 when compared to the other three isoforms[26].

One particular and conserved feature of PR-10 proteins is the large internal cavity formed by the secondary structure elements. The four Cor a 1.04 isoforms display cavity volumes of up to 1540 Å, which is within the range of other PR-10 allergens[17]. In Cor a 1.0401 the surface of the internal cavity (Fig. 2b) is formed by the hydrophobic residues Phe22 (α1), Ile30 (α2), Phe38 (L3), Ile56, Phe58, Ala59 (β3), Phe64, Tyr66, Met67 (β4), Tyr83, Ile85 (β5), Gly88, Gly89, Pro90, Gly92 (L7), Ile98, Tyr100 (β6), Tyr121 (β7), and Ala136, Gly137, Leu144 (α3). An extended hydrophobic surface patch is located at the inner end of the cavity, where the three α-helices meet. In addition, a number of charged amino acid residues convey an amphiphilic character to the cavity, including residues His37 (L3), Glu60 (L5), Lys65, His69 (β4), Glu87 (β5), His93, Thr94 (L7), Asn130 (L9), and Glu132, Glu133, Lys140 (α3). This amphiphilic area extends to an entrance of the cavity (ε1), which is located between helix α3 and strands β3 and β4, as well as loops L5 and L6. The amphiphilic nature of the cavity has also



	% hydrophobic	% donor	% acceptor
Cor a 1.0401	13	45	24
Cor a 1.0402	15	40	27
Cor a 1.0403	20	32	22
Cor a 1.0404	15	37	23

Figure 2. Amphiphilicity of the Cor a 1.04 internal cavity. (**a**) Sequence alignment of the hazelnut (Cor a 1.0401, Cor a 1.0402, Cor a 1.0403, Cor a 1.0404) and hazel pollen (Cor a 1.0101) allergens obtained with Clustal Omega[75]. Amino acid residues are labeled with asterisks (identical), colons (conserved) and dots (semiconserved). Secondary structure elements for the Cor a 1.04 isoforms are indicated on top. Hydrophobic, positively charged (donor), and negatively charged (acceptor) amino acid residues are indicated in yellow, blue and red, respectively. (**b**) Cor a 1.0401 structure (6Y3H) with the amphiphilic internal cavity colored as in (**a**), generated using Schrödinger's Maestro Software Suite[64]. The positions of the six variable amino acid residues in the four Cor a 1.04 isoforms (4, 40, 62, 99, 130 and 158) and the cavity entrance ε1 are indicated. (**c**) Percentage of hydrophobic, proton donor, and proton acceptor interaction potential on the surface of the cavities in the four Cor a 1.04 isoforms.

been reported for other food and pollen allergens[27–29]. Of note, the exact composition of the internal cavity in the four Cor a 1.04 isoforms is fairly diverse, in particular regarding the contributions from hydrophobic and charged residues (Fig. 2c). The percentage of hydrophobic surface area varies between 13 and 20%, while the contribution of positively charged residues to the cavity surface ranges from 32 to 45%, indicating a remarkable degree of variability between these proteins.

Plant food allergens from the PR-10 family typically bind natural flavonoids and other plant derivatives in the internal cavity[17]. While natural ligands of the four hazelnut Cor a 1 allergens are not known to date, a flavonoid compound bound to hazel pollen PR-10 allergens was described recently. Structural data for hazel pollen PR-10 allergens are not available to date, but sequence comparison of nut and pollen allergens (63% sequence identity, Fig. 2a) suggests major differences regarding the internal surface composition. These include the exchange of charged and hydrophobic residues (H37A, G89D, T94K, Y100H, K140M) and charge inversion (H69E, H93D) of residues in the cavity. Nevertheless, the specific flavonoid bound to hazel pollen PR-10 allergens was shown to also bind to the hazelnut isoform Cor a 1.0401 in vitro[24], implying that potential ligands bind to hazelnut and hazel pollen PR-10 proteins with relatively low specificity, in agreement with the current literature[17,30]. Moreover,

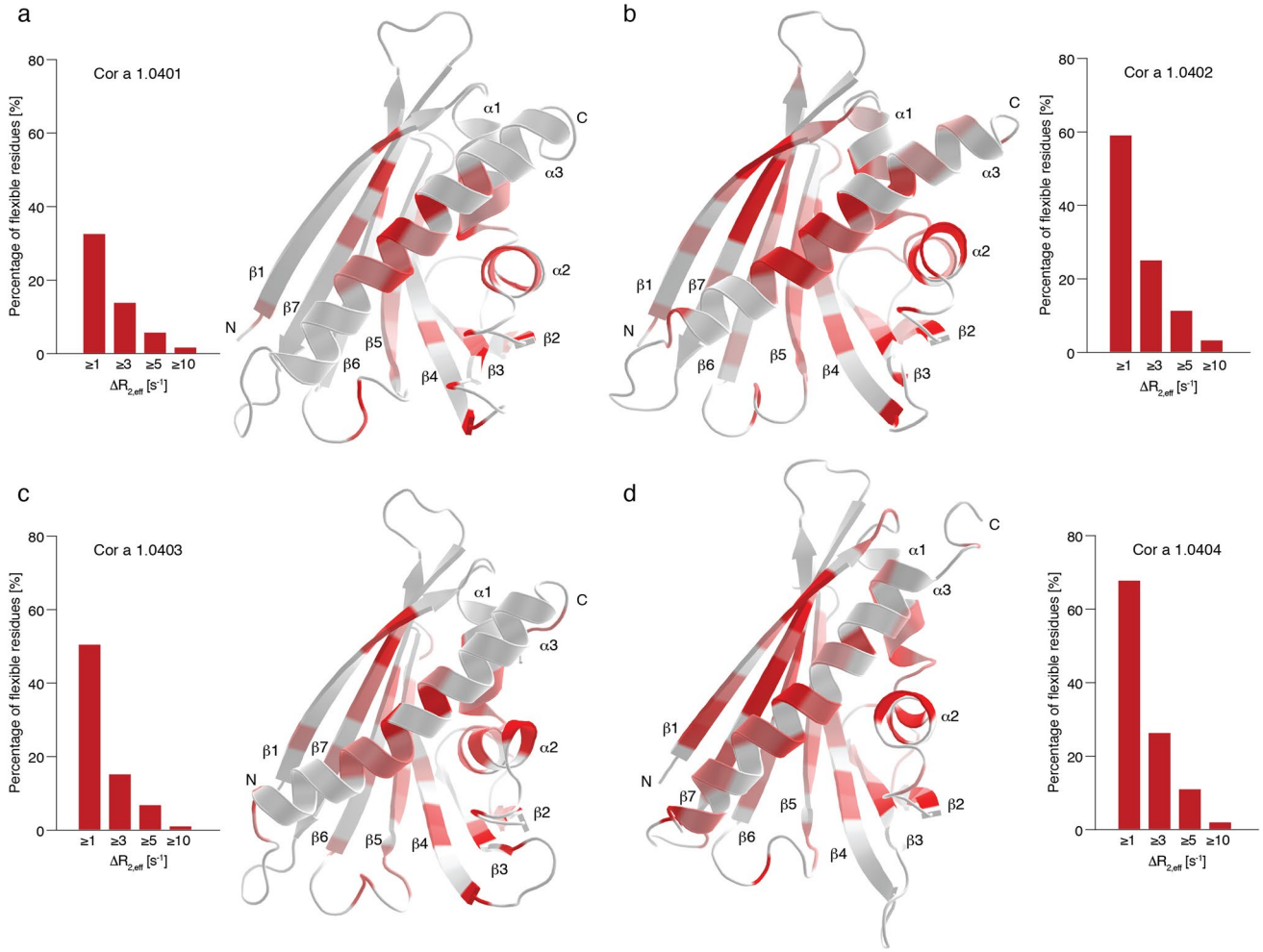


Figure 3. Structural flexibility of the four hazelnut isoforms Cor a 1.0401 (**a**), Cor a 1.0402 (**b**), Cor a 1.0403 (**c**), and Cor a 1.0404 (**d**). The bar plots show the percentage of flexible residues with relaxation dispersion amplitudes ($\Delta R_{2,eff}$ values) exceeding 1, 3, 5, and 10 s^{-1} at 600 MHz. $\Delta R_{2,eff}$ values plotted on the protein backbone with a color gradient from highly flexible (red) to rigid (white), using identical thresholds for all four proteins. Secondary structure elements and the N- and C-termini are indicated.

the diverse nature of the internal cavities of the hazelnut Cor a 1.04 isoforms described above indicates binding of diverse ligands.

The four Cor a 1.04 structures have different flexibilities. The NMR solution structural ensembles of the four Cor a 1.04 isoforms (Fig. 1) suggest that only the C-termini and some loop regions of the four proteins are conformationally heterogeneous, in particular loop L5 connecting strands β3 and β4, which forms part of entrance ε1 to the internal cavity. A similar observation has previously been reported for other PR-10 food allergens, e.g. Mal d 1[27], and it has been suggested that this segment of PR-10 proteins might function as a flexible gate keeper to the protein's interior[31]. For a more in-depth analysis of conformational heterogeneity in the four hazelnut Cor a 1.04 allergens we performed relaxation dispersion (RD) NMR experiments, which provide site-specific information about the presence of different conformers[32]. In RD-NMR experiments, transitions between different conformers occurring on the micro-to-millisecond time scale cause non-flat dispersion profiles, while conformational homogeneity results in flat dispersions. Experimental RD data for representative amino acid residues (Val23, Ser40/Gly40, Tyr100, and Thr118) in the four Cor a 1.04 isoforms are shown in Supplementary Fig. S2 online, and the corresponding RD amplitudes ($\Delta R_{2,eff}$ values) are color-coded on the protein structures in Fig. 3. It is evident from the relaxation dispersion data that a significant portion of the protein backbone is flexible. Non-flat relaxation dispersion profiles are found for numerous residues in all four isoforms not only in loop regions, but also in all secondary structure elements. The most flexible regions include the short helices α1 and α2, parts of the β-sheet (β1, β5, and β7), and the middle of the C-terminal helix α3.

Interestingly, significant differences between the four isoforms are evident. The amplitudes of the RD-profiles exceed 1 s^{-1} for 33%, 59%, 50%, and 68% of all residues in Cor a 1.0401, Cor a 1.0402, Cor a 1.0403, and Cor a 1.0404, respectively (Fig. 3). Likewise, RD amplitudes exceed 3 s^{-1} in 14%, 25%, 15% and 26% of all residues in these proteins, suggesting the "flexibility ranking" Cor a 1.0404 > Cor a 1.0402 > Cor a 1.0403 > Cor a 1.0401.

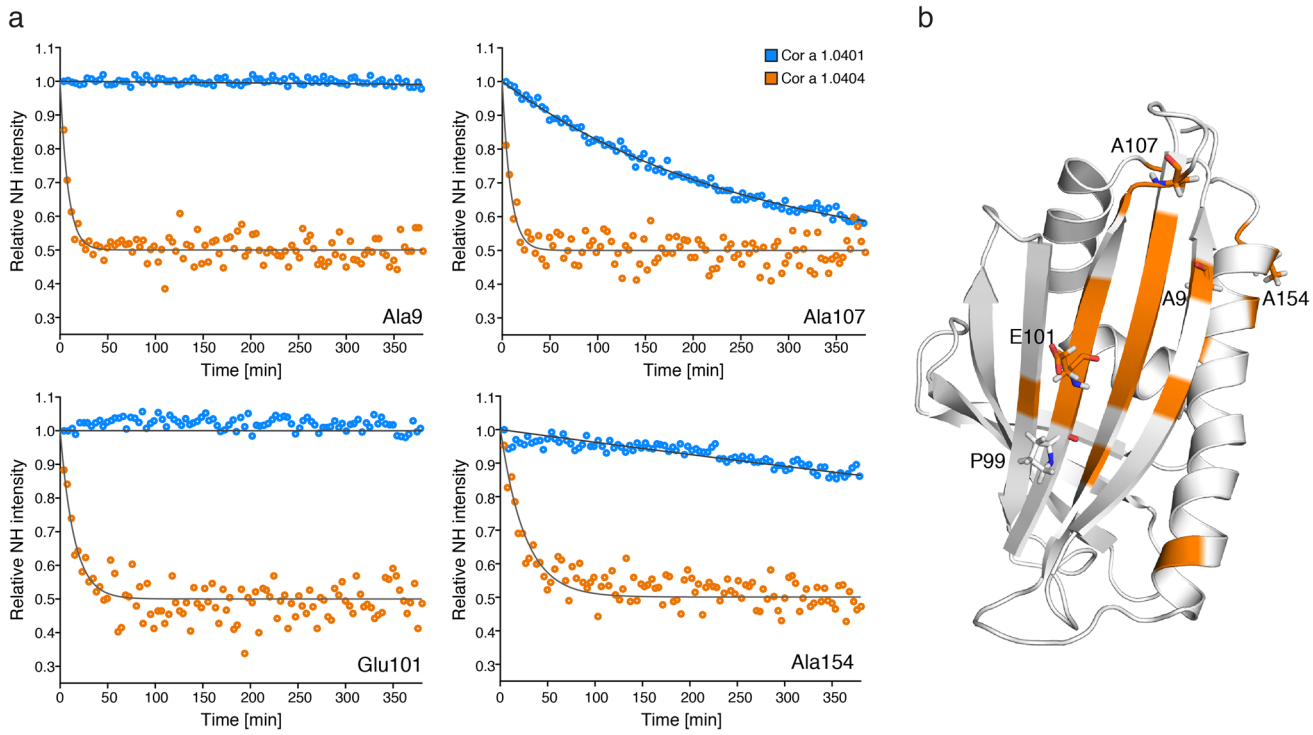


Figure 4. Hydrogen–deuterium exchange in Cor a 1.04 isoforms. (**a**) Time-dependent SOFAST ^{1}H–^{15}N-HMQC backbone amide intensities (circles), recorded at 10 °C, are shown for representative residues with drastically faster H/D exchange in Cor a 1.0404 (orange) than in Cor a 1.0401 (blue), along with best-fit exponential curves. (**b**) Amino acid residues for which H/D exchange is at least four times faster are highlighted in orange on the structure of Cor a 1.0404 (6Y3L). The four exemplary amino acid residues (Ala9, Glu101, Ala107, and Ala154) and Pro99 are displayed as sticks.

Detailed comparison of the data clearly shows that the different flexibilities among the isoforms are not limited to the direct vicinity of the six variable amino acid residues. Indeed, for Cor a 1.0402 and especially for Cor a 1.0404 additional flexibility (when compared to the least flexible isoform Cor a 1.0401) is distributed almost over the entire protein scaffold. Moreover, it is evident that flexible residues are particularly grouped around position 99 in strand β6, which is occupied by proline in Cor a 1.0404 and serine in all other isoforms. This includes residues in the direct proximity of Pro99 in strand β6, such as Ile98 and Tyr100, as well as residues in the adjacent strands β5 and β7. Increased flexibility in β7 of Cor a 1.0404 is evident from both relaxation dispersion (Thr118 and Ser119) and NMR order parameter data (Ile114, Leu115, and Lys120, see Supplementary Fig. S3 online).

In isoforms Cor a 1.0401–03, the backbone amide of Ser99 forms a hydrogen bond to Lys120 in β7. It is likely that the lack of such a hydrogen bond in Cor a 1.0404 causes structural lability of the β-sheet of this particular isoform. Nevertheless, it is clear from the NMR relaxation data that even parts of the protein, which are distal from position 99 (e.g., helix α3), display the highest degree of flexibility in the isoform Cor a 1.0404.

Cor a 1.0404 is structurally labile. The most (Cor a 1.0404) and the least (Cor a 1.0401) structurally flexible isoforms, have been reported as those with the lowest and the highest IgE-binding potential, respectively[22]. Structural flexibility of proteins can result in transient exposure of backbone amides to solvent water, rendering them susceptible to exchange with surrounding water molecules. To further investigate structural flexibility in these two Cor a 1.04 isoforms, the solvent exposure of backbone amides was probed by use of NMR hydrogen–deuterium (H/D) exchange measurements that detect the replacement of the backbone amide protons with deuterium upon addition of D_2O.

Figure 4 shows experimental hydrogen–deuterium exchange data for four exemplary amino acid residues Ala9, Glu101, Ala107, and Ala154. Because H/D exchange of most backbone amide protons in Cor a 1.0404 is remarkably fast, the experimental data for both isoforms were recorded at low temperature (10 °C) using rapid NMR data collection (SOFAST) techniques. ^{1}H–^{15}N correlation spectra were obtained in ca. 4 min, which enabled us to record reliable H/D exchange data, even for rapidly exchanging backbone amides, revealing significantly accelerated exchange in Cor a 1.0404 compared to Cor a 1.0401. In detail, Ala9, Glu101 and Ala154, which are located in secondary structure elements and hydrogen bonded, have very slow exchange rates in Cor a 1.0401, while Ala107, which is surface exposed in loop L8, displays a moderate exchange rate. In Cor a 1.0404, all four amino acid residues display drastically faster H/D exchange. The positions of these residues are indicated on the Cor a 1.0404 structure in Fig. 4, along with all other amino acid residues whose H/D exchange rate is at least four times faster in Cor a 1.0404 than in Cor a 1.0401 under identical conditions. Acceleration is not limited to the immediate sequential neighborhood of the two variable amino acid residues in the two isoforms (positions

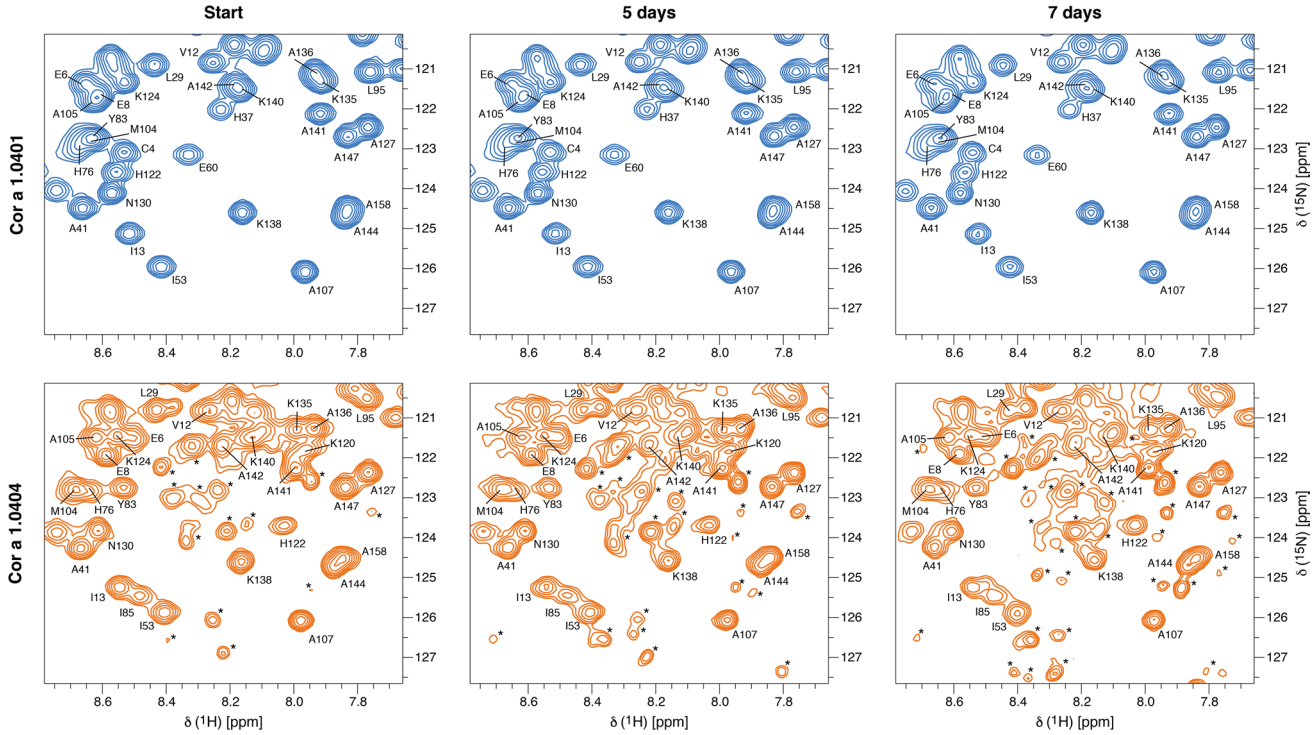


Figure 5. Time-dependent stability of Cor a 1.04 isoforms. Sections of ^{1}H–^{15}N-HSQC spectra of Cor a 1.0401 (top) and Cor a 1.0404 (bottom) directly after purification, as well as 5 and 7 days later. Assignments[26] are indicated by single letter codes and signals labeled with an asterisk indicate aggregation and degradation.

4 and 99). Indeed, a number of residues that show increased exchange rates in Cor a 1.0404 are located in strand β7, which is adjacent to strand β6 containing Pro99.

The fact that hydrogen–deuterium exchange of the two proteins Cor a 1.0401 and Cor a 1.0404, despite sharing similar structural scaffolds, exhibit significantly different exchange rates, indicates that hydrogen bonds in the Cor a 1.0404 isoform are weaker. This structural lability could be elicited by the missing hydrogen bond between Pro99 and Lys120 in the center of the β-sheet. Of note, the only other difference between the two isoforms is position 4 at the N-terminus, which is occupied by a cysteine in Cor a 1.0401 and a serine in Cor a 1.0404.

To investigate the impact of the two variable amino acid residues at positions 4 and 99 in Cor a 1.0401 and Cor a 1.0404 in detail, we performed time-dependent NMR degradation assays (Fig. 5). After 5 and 7 days at room temperature, Cor a 1.0404 shows degradation and aggregation peaks, while Cor a 1.0401 spectra remain unchanged during this time period. NMR resonance assignments of a partly degraded sample of Cor a 1.0404, identified the presence of small unstructured peptides Tyr5-Ile13 and Pro31-Ala34, indicating proteolytic cleavage of this particular isoform. Three mutant forms, C4S Cor a 1.0401, P99A Cor a 1.0404, and P99T Cor a 1.0404 were also investigated regarding degradation. While Cys4Ser probes the natural variability at position 4, the two proline mutants were chosen to examine the influence of Pro99. The NMR spectra of these variants show that their overall fold is not affected by the mutation and all three proteins were stable during the course of the assay (Supplementary Fig. S4 online). This implies that Pro99 in Cor a 1.0404 is the reason for the loss of stability.

Finally, the effect of temperature on Cor a 1.0401 was probed by acquiring ^{1}H–^{15}N-HSQC spectra at different temperatures in the range of 10 °C to 70 °C (Supplementary Fig. S5 online). This protein remains folded even at 70 °C and precipitates at 80 °C. Heating of Cor a 1.0401 is fully reversible until 70 °C without intensity loss after cooling the sample again to 25 °C. This indicates high thermal stability of the Cor a 1.0401 structure in solution, in agreement with the observed IgE reactivity in roasted hazelnuts[33]. In contrast, for the isoform Cor a 1.0404 ^{1}H–^{15}N-HSQC spectra display substantially reduced intensities at temperatures above 40 °C, most likely due to accelerated chemical exchange of backbone amides with solvent.

Taken together, our NMR data indicate general structural lability of the Cor a 1.0404 isoform that is very likely caused by the presence of Pro99 in the center of the β-sheet and a weakened hydrogen bonding network.

Cor a 1.04 isoforms show different IgE-binding potentials. A previous study demonstrated binding of specific IgE from patients' sera (Switzerland and Denmark) to the four hazelnut Cor a 1.04 isoforms[22]. Measurably different IgE-binding properties were reported, with Cor a 1.0404 being the isoform with the lowest IgE-binding potential, while Cor a 1.0401 displayed the highest potential.

We verified the immunologic activity of our recombinantly produced Cor a 1.04 proteins using blood sera from twenty-two Austrian patients (Tyrol) included in a pilot study[34]. The patients show birch-pollen related hazelnut allergy (specific IgE to Cor a 1 > 0.35 kU/mL and Bet v 1 > 0.35 KU/mL), positive skin-prick test to hazelnut extract and oral allergy symptoms to hazelnuts (see demographic data in Table 2). ELISA experiments

Characteristics	Values
Sex, n (f/m)	22 (16/6)
Median (range)	
Age [y]	36 (22–68)
SPT [mm]*	5.5 (3.0–9.0) n = 13
ImmunoCAP	
IgE total [kU/L]	124.0 (9.4–948.0)
Cor a 1.04 specific IgE [kU/L]	8.02 (0.87–52.10)
Bet v 1 specific IgE [kU/L]	14.40 (2.96–91.40)
ELISA	
Cor a 1.0401 specific IgE [kU/L]	2.02 (1.31–15.15)
Cor a 1.0402 specific IgE [kU/L]	1.21 (0.82–4.98)
Cor a 1.0403 specific IgE [kU/L]	1.57 (0.95–13.08)
Cor a 1.0404 specific IgE [kU/L]	0.88 (0.48–1.41)

Table 2. Demographic and immunologic characteristics of 22 patients with birch pollen-related hazelnut allergy. *Skin-Prick test (SPT) was determined as positive if the wheal was > 3 mm.

reveal that Cor a 1.0401 indeed has the highest IgE-binding potential, while Cor a 1.0404 has the lowest. The isoforms Cor a 1.0402 and Cor a 1.0403 have moderate to high potentials, with the latter featuring slightly higher values, as evident from Fig. 6a. In 91% of all patients Cor a 1.0401 shows the highest binding of specific IgE, while Cor a 1.0404 barely shows binding potential. Besides, Cor a 1.0403 has a higher IgE-binding potential than Cor a 1.0402 in 91% of all patients. Sera IgE reactivities toward the first three isoforms appear to be patient-specific, with considerable variations between them, while IgE reactivities toward Cor a 1.0404 are consistently low for all sera. Box plots of the patient-derived data confirm that specific IgE values for isoforms Cor a 1.0401–03 display relatively broad distributions within the patient group, whereas Cor a 1.0404 values are systematically low (Fig. 6b).

In immunosorbent assays, different blotting efficiencies of proteins on hydrophobic surfaces of microplates and different propensities to desorb off the surface during the washing procedure can affect the amount of antigen that is available for IgE-binding and detection. These effects limit the use of immunosorbent assays per se for comparing IgE-binding of different isoforms. Nevertheless, the ranking of the immunologic activity of the four hazelnut Cor a 1 isoforms observed in basophil release assays complies with enzyme allergosorbent tests, reported by Lüttkopf et al.[22]. Likewise, agreement of basophil release data with allergosorbent data was shown in a study comparing the different hazelnut allergens Cor a 1, Cor a 8, Cor a 9, and Cor a 11 by Blanc et al.[35]. This supports the notion that the effect seen in our ELISA experiments indeed originates from different IgE-binding of the four isoforms.

A statistical analysis of the ELISA data and ImmunoCAP data are shown in Fig. 6c, indicating that the IgE-binding potential of Cor a 1.0401 shows no significant difference with the ImmunoCAP data and Cor a 1.0403. Additionally, no difference is observed for Cor a 1.0402 with Cor a 1.0403 and Cor a 1.0404. In summary, the patient derived ELISA data clearly suggest the immunologic ranking Cor a 1.0401 > 03 > 02 > 04.

As mentioned above, between Cor a 1.0401 and Cor a 1.0404, only two amino acid residues are different (Cys4/Ser4 and Ser99/Pro99), with their side-chains located on the protein surface, and our mutational studies suggest that proline at position 99 of Cor a 1.0404 causes the structural lability of this particular isoform. Likewise, Pro99 is presumably the main reason for the dramatically reduced IgE-binding potential of Cor a 1.0404 compared to Cor a 1.0401, as suggested by Lüttkopf et al.[22]. This is corroborated by our IgE-binding data for the isoforms Cor a 1.0402 and Cor a 1.0403, which differ only at position 4 (Ser4/Cys4) and display comparable IgE-binding, with Cor a 1.0403 having only a slightly higher binding potential than Cor a 1.0402. These data suggest that replacement of a serine by a cysteine residue at position 4 has a minor effect on IgE-binding, even though these residues are surface exposed, as previously noted in enzyme allergosorbent tests[22].

In principle, IgE-binding can also be modulated by the formation of dimers or higher oligomers[36]. Cor a 1.04 isoforms contain between two and three cysteine residues (at positions 4, 82 and 160) that are surface exposed and accessible for cysteine-mediated dimerization. Therefore, we probed the oligomerization state of Cor a 1.0401 (three cysteines) and Cor a 1.0404 (Cys82 and Cys160) under the buffer conditions that we employed for our ELISA and NMR experiments using pulsed-field gradient experiments. Cor a 1.0401 has a hydrodynamic radius of 18.6 ± 0.1 Å without and 18.7 ± 0.4 Å with a reducing agent (DTT) being present, respectively, which is consistent with a monomeric protein in both cases and comparable to other monomeric PR-10 proteins[27,37]. This is supported by the retention times observed in size exclusion chromatography, which are characteristic for a monomeric PR-10 protein. Likewise, the hydrodynamic radius of Cor a 1.0404 is 18.6 ± 0.5 Å, again indicating a monomeric protein. Cor a 1.04 isoforms thus show no tendency for dimerization or oligomerization under the experimental conditions that were used for IgE-binding measurements.

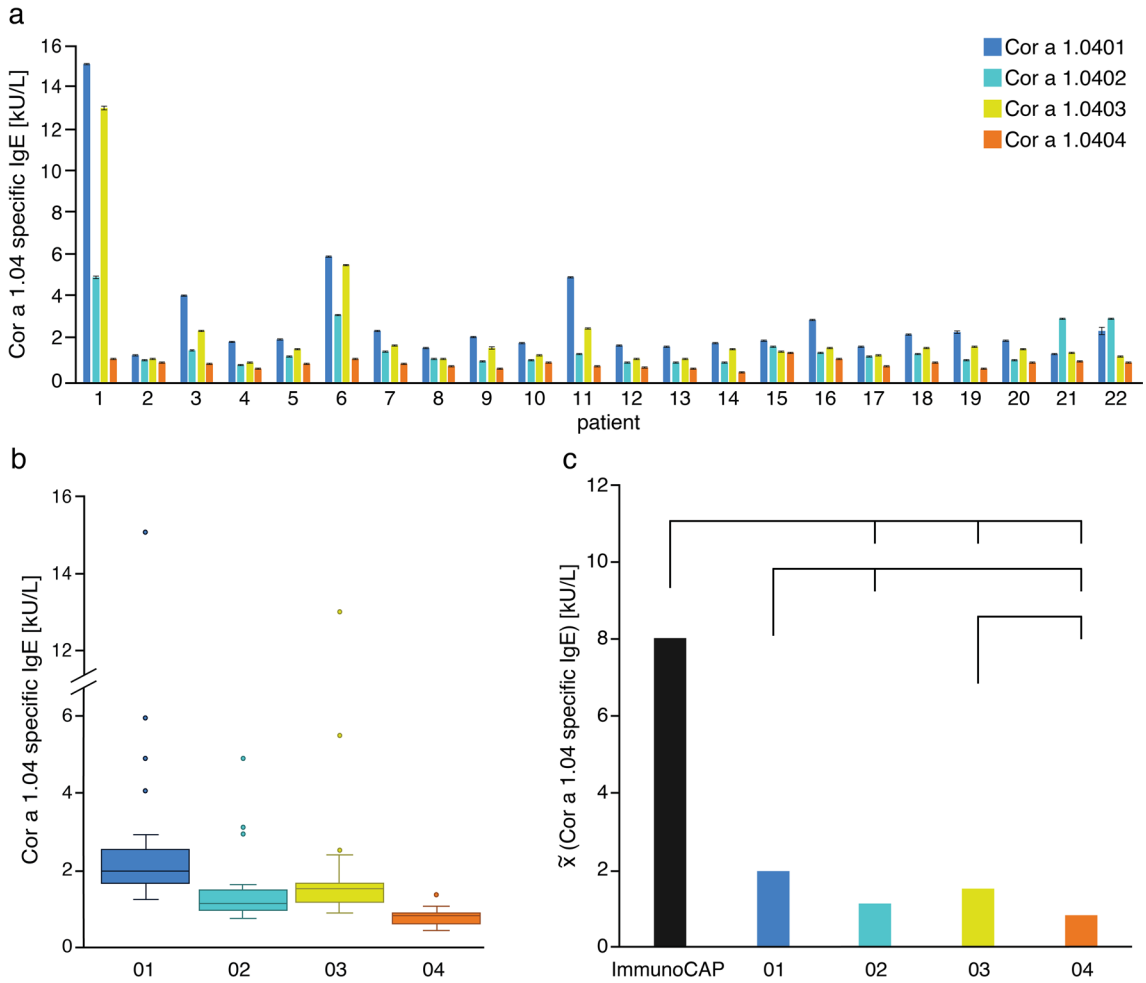


Figure 6. Specific IgE-binding to the four hazelnut isoforms Cor a 1.0401 (blue), Cor a 1.0402 (light blue), Cor a 1.0403 (yellow), and Cor a 1.0404 (orange). (**a**) Bar plots and standard deviations of IgE-binding potentials obtained by ELISA for each patient (n = 22), using identical amounts of recombinantly produced allergen. (**b**) Box plots of the IgE-binding potentials of the four isoforms. The boxes are restricted by the 25th and 75th percentile, median values are given by a horizontal line inside the box and whiskers cover the minimal and maximal values inside the interquartile range (1.5 IQR). Patients with high total IgE levels and therefore high specific IgE are indicated by dots. (**c**) ELISA data (median values) and ImmunoCAP data (black). Differences between the IgE-binding potentials are given by the lines above with a p value < 0.05 as statistically significant. The exact pairwise p values are 0.006 (01–02), 0.453 (01–03), 1.89×10^{-8} (01–04), 0.565 (01-ImmunoCAP), 1.000 (02–03), 0.100 (02–04), 9.33×10^{-7} (02-ImmunoCAP), 0.001 (03–04), 0.001 (03-ImmunoCAP), and 2.44×10^{-14} (04-ImmunoCAP). All three panels suggest the immunologic ranking Cor a 1.0401 > 03 > 02 > 04.

Conclusions

Our experimental NMR and immunologic data suggest an inverse relationship between IgE-binding observed in ELISA experiments and the structural flexibility of the four known isoforms of the Cor a 1 hazelnut allergens, even though their three-dimensional structures are highly similar. Cor a 1.0401 is the isoform with the highest potential to bind IgE, while it is the protein in this group with the most rigid backbone scaffold. Cor a 1.0404, on the other hand, is the isoform that has the lowest IgE-binding potential, yet this protein is strikingly flexible and conformationally heterogeneous in solution and displays a weak hydrogen bonding network. Cor a 1.0402 and Cor a 1.0403 are intermediate between these isoforms regarding their IgE-binding potential and their structural flexibility.

Enhanced levels of structural flexibility have been observed for various allergens[38], including those from the PR-10 family of proteins[39,40]. Only for the major birch pollen allergen, Bet v 1, isoform-specific flexibility data have been reported, revealing Bet v 1.0101 to be fairly rigid and having substantially higher IgE-binding potential than the structurally flexible isoform Bet v 1.0102[39,41]. As with Cor a 1, structural flexibility in Bet v 1 is distributed across the entire PR-10 scaffold, including secondary structure elements and loops. These observations parallel our experimental data for the hazelnut allergen described here.

Structural flexibility has been recognized as a critical component of antigen–antibody binding[42]. The reduction of antibody structural flexibility was proposed as means to lower the entropic cost of complex formation in order to design antibodies with increased affinities[43]. Along these lines, it has been shown by experiment that

complex formation leads to antibody rigidification[43], and reduced antibody flexibility has been proposed as general feature of antigen–antibody binding[44].

Experimental data regarding the rigidification of allergens in antibody complexes are not available to date. Considering the pronounced flexibility of PR-10 allergens in solution, it is possible that complex formation is indeed accompanied by rigidification of the protein scaffold, resulting in a loss in conformational entropy and causing a lower binding affinity for more flexible isoforms. According to the concept of conformational selection, it is also conceivable that in the more rigid isoforms the binding-competent conformation is highly populated in solution. The more flexible isoforms on the other hand may fluctuate between several conformational states, resulting in a lower population of the binding-competent conformation. This lack of structural pre-organization could contribute to lower IgE-binding. The inverse relation between structural flexibility and IgE-binding observed in our Cor a 1 study suggests that entropic contributions indeed play an appreciable role in complex formation between these allergens and antibodies.

Methods

Additional methods are described in the supplementary methods section (Supplementary Information).

Plasmid generation and protein expression. The construction of plasmids encoding the hazelnut allergens Cor a 1.0401 (accession no. AAD48405), Cor a 1.0402 (accession no. AAG40329), Cor a 1.0403 (accession no. AAG40330), and Cor a 1.0404 (accession no. AAG40331) and the subsequent expression has been described previously[26]. Plasmid constructs encoding the mutant forms C4S Cor a 1.0401, P99A Cor a 1.0404, and P99T Cor a 1.0404 were created by site-directed mutagenesis using Phusion DNA polymerase (New England Biolabs, Frankfurt am Main, Germany) for C4S Cor a 1.0401 and Platinum SuperFi DNA polymerase (Thermo Fisher Scientific, Vienna, Austria) for P99A Cor a 1.0404, and P99T Cor a 1.0404 according to the protocol of the manufacturers. The mutant forms P99A Cor a 1.0404 and P99T Cor a 1.0404 were generated using C4S Cor a 1.0401 as template. The used oligonucleotides were 5′-ccatgggcgtgttcctcctacgaagatgagg-3′ and 5′-cctcatcttcgtaggagaacacgcccatgg-3′ for C4S Cor a 1.0401, 5′-ccctggaaaaaatcgcctacgagattaaaatggc-3′ and 5′-gccattttaatctcgtaggcgatttttttccaggg-3′ for P99A Cor a 1.0404, and 5′-ccctggaaaaaatcacctacgagattaaaatggc-3′ and 5′-gccattttaatctcgtaggtgatttttttccaggg-3′ for P99T Cor a 1.0404 (underlined nucleotides were mutated). Mutagenesis was verified by DNA-sequencing (Supplementary Fig. S6 online). Expression of the three mutant proteins was performed as for the wild-type hazelnut isoforms but without using ISOGRO-^{15}N powder for the expression[26]. Unlabeled proteins were expressed in M9 minimal medium without isotopically labeled reagents.

Protein purification and preparation. The protein purification protocol for the four wild-type hazelnut isoforms has been described previously[26]. The same protocol was used for the three mutant forms. All NMR samples contained 20 mM sodium phosphate buffer (pH 6.9), 2 mM DTT, 9% D_2O and 0.5 mM ^{15}N labeled or ^{15}N/^{13}C labeled protein.

NMR structure determination. The resonance assignments for the four isoforms with all corresponding NMR experiments have been reported previously and are available at the Biological Magnetic Resonance Data Bank (https://bmrb.wisc.edu) under the accession numbers 27965, 27961, 27967, and 28016 for Cor a 1.0401, Cor a 1.0402, Cor a 1.0403, and Cor a 1.0404, respectively[26]. For structure determination, three-dimensional ^{1}H–^{15}N-NOESY-HSQC, ^{1}H–^{13}C-NOESY-HSQC, and aromatic ^{1}H-^{13}C-NOESY-HSQC experiments with a mixing time of 150 ms each were carried out at 25 °C on a 500 MHz Agilent DirectDrive 2 spectrometer equipped with a room temperature probe. Processing of the NMR data was performed with NMRPipe[45], and spectra were visualized with nmrDraw[45] and analyzed using CcpNmr[46].

For the manual assignment of the NOE cross-peaks an initial CS-Rosetta[47–49] structure model for each isoform was used and the obtained cross-peaks were converted into distance restraints, based on their intensities, ranging from very strong to very weak with upper limits of 2.8 Å, 3.0 Å, 5.0 Å, 6.0 Å, and 6.5 Å. In an iterative manner, the NOE data were complemented and validated by generating three-dimensional structural bundles with the program XPLOR-NIH (version 2.52)[50,51]. Hydrogen bond restraints were derived from the CS-Rosetta structure model for the secondary structure elements, if no water cross-peak was detected in the proton TOCSY spectra. Dihedral angle restrains (Φ and Ψ) were derived from the program TALOS+[52].

For the final structure determination, 500 structures were calculated in 3000 steps at an initial temperature of 7000 K, followed by 10,000 cooling steps, using an extended starting structure. The 20 structures with the lowest energy were further refined in explicit solvent. For this purpose, each structure was surrounded by a cubic box of TIP3P[53] waters with at least 10 Å from each atom to the box boundaries. The LEaP module of AMBER 18[54] using the Amber force field 14SB[55] parameterized the system. After solvent relaxation[56], simulated annealing calculations with a Langevin thermostat[57] (collision frequency: 2 ps^{-1}) and a Berenden barostat[58] (relaxation time: 2 ps) were performed. Hydrogen bonds were constrained with the SHAKE algorithm[59] and a van der Waals cutoff of 10 Å was used along with the particle-mesh Ewald method[60] for long range electrostatics. A simulated annealing scheme of 50 ns with a time step of 1 fs for each structure was performed using the NOE distance restraints. The protein structure validation software (PSVS) suite[61] was used to validate the refined structures. Internal cavity volumes and surface hydrophobicity were determined using SiteMap[62,63] as implemented in the Schrödinger Maestro Software Suite[64]. RMSD values between the different PR-10 proteins were obtained using the program SuperPose[65].

Temperature dependency and NMR diffusion experiments. The temperature sensitivity of Cor a 1.0401 and Cor a 1.0404 was probed by recording ^{1}H–^{15}N-HSQC spectra on a 700 MHz Bruker Avance Neo

spectrometer equipped with a Prodigy CryoProbe at different temperatures (10, 15, 20, 25, 30, 35, 40, 50, 60, 70 °C). Chemical shifts were referenced at all temperatures using an external standard sample containing 1% (w/v) 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) in 450 µL 20 mM sodium phosphate buffer (pH 6.9), 2 mM DTT, supplemented with 10% D_2O.

For the determination of the oligomerization state of Cor a 1.0401 and Cor a 1.0404 stimulated echo pulsed field gradient experiments were used[66]. The employed gradient field strengths were 2.0, 4.5, 7.0, 9.5, 12.0, 14.5, 17.0, 19.5, 22.0, and 24.5 G/cm with a constant diffusion time of 160 ms. Hydrodynamic radii were determined for all dispersed peaks in the ^{1}H–^{15}N-HSQC spectrum by using an in-house MATLAB fitting program and the Stokes–Einstein equation, as described in detail for Bet v 1[37]. Average values and standard deviations of the hydrodynamic radii were calculated from the 20 residues with the lowest RMSD values.

NMR hydrogen–deuterium exchange experiments. Hydrogen–deuterium exchange of the amide protons of Cor a 1.0401 and Cor a 1.0404 was assessed by SOFAST ^{1}H–^{15}N-HMQC spectra[67] acquired at 700 MHz using a Prodigy CryoProbe at 10 °C and a rapid injection device, as described by Schanda et al.[68]. NMR samples were prepared in 20 mM sodium phosphate buffer (pH 6.9) with 2 mM DTT, and hydrogen–deuterium exchange was initiated by addition of an equal volume of the same buffer in D_2O to a final protein concentration of 0.3 mM. A series of SOFAST ^{1}H–^{15}N-HMQC spectra were recorded immediately after sample mixing, each lasting for 4.12 min (Cor a 1.0401) or 3.8 min (Cor a 1.0404). Using an in-house MATLAB fitting script, the hydrogen–deuterium exchange rates were obtained from exponential decay fits.

NMR relaxation experiments. Backbone amide ^{15}N relaxation dispersion experiments were recorded on a 500 MHz spectrometer, a 600 MHz Bruker Avance II + spectrometer equipped with a Prodigy CryoProbe, and a 700 MHz spectrometer, using sensitivity enhanced Carr-Purcell-Meiboom-Gill (CPMG) sequences[69] with ^{1}H continuous-wave decoupling during the CPMG period[70]. Spectra were recorded at different CPMG field strengths $\nu_{CPMG} = (2\tau_{CPMG})^{-1}$, with τ_{CPMG} being the time between two consecutive 180° pulses in the CPMG pulse train. The ν_{CPMG} field strengths were 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, with repeat experiments at 66.7 Hz and 600.0 Hz, and the length of the CPMG pulse train was set to $T_{relax} = 30$ ms in all experiments. Effective relaxation rates $R_{2,eff} = -1/T_{relax} \times \ln(I/I_0)$ were determined from partial peak volumes (intensities in 5 × 5 grids) of the resonances in the individual spectra (I), along with a reference intensity (I_0) with T_{relax} set to 0. The so-obtained relaxation dispersion profiles were analyzed by fitting the exact CPMG expression for two-site exchange using an in-house-written MATLAB script[71]. Per-residue relaxation dispersion amplitudes ($\Delta R_{2,eff}$) were calculated from the 600 MHz data as $\Delta R_{2,eff} = R_{2,eff}(\nu_{CPMG} = 0) - R_{2,eff}(\nu_{CPMG} = \infty)$, where $R_{2,eff}(\nu_{CPMG} = \infty)$ and $R_{2,eff}(\nu_{CPMG} = 0)$ are the extrapolated $R_{2,eff}$ values at infinite and zero CPMG field strengths, respectively.

IgE-binding experiments. A skin-prick test (SPT) with hazelnut extract (ALK-Abelló, Linz, Austria), histamine as a positive control and diluent as a negative control was performed on the flexor surface of the forearm according to Heinzerling et al.[72,73]. Serum levels of total IgE and IgE specific for Bet v 1 and Cor a 1 were determined by ImmunoCAP (Phadia 250, Thermo Fisher Scientific, Uppsala, Sweden) according to the manufacturer's specifications. The demographical and immunologic data (Table 2) are expressed as median values with corresponding ranges. Patients' blood sera were provided within the AppleCare Study (funded by the European Regional Development Fund Interreg V-A Italy-Austria 2014–2020).

The IgE-binding potential and the biological activity of Cor a 1.0401, Cor a 1.0402, Cor a 1.0403, and Cor a 1.0404 was assessed by using an indirect enzyme-linked immunosorbent assay (ELISA). Unlabeled freshly recombinantly produced proteins and the mAb 107 of the Human IgE ELISA development kit HRP (Mabtech, Nacka Strand, Sweden), which was used for standardization, were diluted in 0.1 M $NaHCO_3$ pH 9.7 to a final concentration of 1 µg/mL. In each well of a 96 Well Corning Costar Assay Plate (Merck, Darmstadt, Germany) either 100 µL mAb 107 (for standardization) or 100 µL protein (for specific IgE detection) were placed and incubated at 4 °C overnight. The plate was washed with 150 µL PBS/0.1% Tween 20 the next day and subsequently saturated with 150 µL PBS/3% BSA per well. Incubation was conducted for 90 min under moderate shaking and afterwards the plate was washed as before. The patients' blood sera were diluted 1:2 (v/v) with PBS and the human IgE standard from the kit was diluted to 100, 50, 20, 10, 5, 2, and 0 ng/mL with PBS. Either 50 µL of diluted serum or diluted standard were added to the wells containing protein or mAb 107, respectively and the plate was incubated at 4 °C overnight. Another washing step was conducted as before and the Anti-Human IgE (ε-chain specific) detecting antibody (Merck, Darmstadt, Germany) was diluted 1:1000 (v/v) with PBS/1% BSA and 50 µL of it were added to each well. Incubation was conducted for 60 min and afterwards the plate was washed as before and 100 µL SureBlue TMB Microwell Peroxidase Substrate (SeraCare, Milford, MA, USA) were added to each well and incubated for 20 min in the dark. The reaction was stopped by the addition of 100 µL TMB Stop Solution (SeraCare, Milford, MA, USA). Between each incubation step the plate was sealed with adhesive plate sealers to avoid evaporation and contamination. The absorption at 650 nm was measured at a microplate reader. Each protein and standard was measured as triplicate and a linear logarithmic function was derived from the standards. Values for Cor a 1.04 specific IgE [kU/L] was determined using this linear function.

The Shapiro–Wilk test was used to evaluate normal distribution. Correlations between specific Cor a 1.04 IgE values with the IgE values determined by ImmunoCAP were assessed by the Spearman rank test. Differences between the groups were tested with the Friedman two-way ANOVA and Post hoc tests. All statistical analyses were done with the program SPSS 25[74].

Data availability

All data generated and analyzed are included in this article and its Supplementary Information file.

Received: 3 December 2020; Accepted: 5 February 2021
Published online: 18 February 2021

References

1. Datema, M. R. *et al.* Component-resolved diagnosis and beyond: Multivariable regression models to predict severity of hazelnut allergy. *Allergy* **73**, 549–559. https://doi.org/10.1111/all.13328 (2018).
2. Datema, M. R. *et al.* Hazelnut allergy across Europe dissected molecularly: A EuroPrevall outpatient clinic survey. *J. Allergy Clin. Immunol.* **136**, 382–391. https://doi.org/10.1016/j.jaci.2014.12.1949 (2015).
3. Lyons, S. A. *et al.* Walnut allergy across Europe: Distribution of allergen sensitization patterns and prediction of severity. *J. Allergy Clin. Immunol. Pract.* https://doi.org/10.1016/j.jaip.2020.08.051 (2020).
4. Ebner, C. *et al.* 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.* **95**, 962–969. https://doi.org/10.1016/s0091-6749(95)70096-x (1995).
5. Vieths, S., Scheurer, S. & Ballmer-Weber, B. Current understanding of cross-reactivity of food allergens and pollen. *Ann. N. Y. Acad. Sci.* **964**, 47–68. https://doi.org/10.1111/j.1749-6632.2002.tb04132.x (2002).
6. Eriksson, N. E., Formgren, H. & Svenonius, E. Food hypersensitivity in patients with pollen allergy. *Allergy* **37**, 437–443. https://doi.org/10.1111/j.1398-9995.1982.tb02323.x (1982).
7. Geroldinger-Simic, M. *et al.* Oral exposure to Mal d 1 affects the immune response in patients with birch pollen allergy. *J. Allergy Clin. Immunol.* **131**, 94–102. https://doi.org/10.1016/j.jaci.2012.06.039 (2013).
8. Geroldinger-Simic, M. *et al.* Birch pollen-related food allergy: Clinical aspects and the role of allergen-specific IgE and IgG4 antibodies. *J. Allergy Clin. Immunol.* **127**, 616–622. https://doi.org/10.1016/j.jaci.2010.10.027 (2011).
9. Hansen, K. S. *et al.* Component-resolved in vitro diagnosis of hazelnut allergy in Europe. *J. Allergy Clin. Immunol.* **123**, 1134–1141. https://doi.org/10.1016/j.jaci.2009.02.005 (2009).
10. Mari, A., Ballmer-Weber, B. K. & Vieths, S. The oral allergy syndrome: Improved diagnostic and treatment methods. *Curr. Opin. Allergy Clin. Immunol.* **5**, 267–273. https://doi.org/10.1097/01.all.0000168793.27948.b0 (2005).
11. 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.* **110**, 797–804. https://doi.org/10.1067/mai.2002.128946 (2002).
12. 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.* **160**, 56–62. https://doi.org/10.1159/000339244 (2013).
13. Wigotzki, M., Steinhart, H. & Paschke, A. Determination of the allergenicity of various hazelnut products by immunoblotting and enzyme allergosorbent test inhibition. *J. Chromatogr. B Biomed. Sci. Appl.* **756**, 239–248. https://doi.org/10.1016/s0378-4347(01)00084-6 (2001).
14. Müller, U. *et al.* Allergens in raw and roasted hazelnuts (*Corylus avellana*) and their cross-reactivity to pollen. *Eur. Food. Res. Technol.* **212**, 2–12. https://doi.org/10.1007/s002170000245 (2000).
15. Verhoeckx, K. C. M. *et al.* Food processing and allergenicity. *Food. Chem. Toxicol.* **80**, 223–240. https://doi.org/10.1016/j.fct.2015.03.005 (2015).
16. Somkuti, J., Houska, M. & Smeller, L. Pressure and temperature stability of the main apple allergen Mal d1. *Eur. Biophys. J.* **40**, 143–151. https://doi.org/10.1007/s00249-010-0633-8 (2011).
17. Fernandes, H., Michalska, K., Sikorski, M. & Jaskolski, M. Structural and functional aspects of PR-10 proteins. *FEBS J.* **280**, 1169–1199. https://doi.org/10.1111/febs.12114 (2013).
18. Breiteneder, H. *et al.* Four recombinant isoforms of Cor a I, the major allergen of hazel pollen, show different IgE-binding properties. *Eur. J. Biochem.* **212**, 355–362. https://doi.org/10.1111/j.1432-1033.1993.tb17669.x (1993).
19. Hoffmann-Sommergruber, K. *et al.* Genomic characterization of members of the Bet v 1 family: Genes coding for allergens and pathogenesis-related proteins share intron positions. *Gene* **197**, 91–100. https://doi.org/10.1016/s0378-1119(97)00246-1 (1997).
20. Hirschwehr, R. *et al.* Identification of common allergenic structures in hazel pollen and hazelnuts: a possible explanation for sensitivity to hazelnuts in patients allergic to tree pollen. *J. Allergy Clin. Immunol.* **90**, 927–936. https://doi.org/10.1016/0091-6749(92)90465-e (1992).
21. Costa, J., Mafra, I., Carrapatoso, I. & Oliveira, M. B. Hazelnut allergens: Molecular characterization, detection, and clinical relevance. *Crit. Rev. Food Sci. Nutr.* **56**, 2579–2605. https://doi.org/10.1080/10408398.2013.826173 (2016).
22. Lüttkopf, D. *et al.* 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. https://doi.org/10.1016/s0161-5890(01)00087-6 (2002).
23. Bohle, B. *et al.* Characterization of the T cell response to the major hazelnut allergen, Cor a 1.04: evidence for a relevant T cell epitope not cross-reactive with homologous pollen allergens. *Clin. Exp. Allergy* **35**, 1392–1399. https://doi.org/10.1111/j.1365-2222.2005.02332.x (2005).
24. Jacob, T. *et al.* Identification of a natural ligand of the hazel allergen Cor a 1. *Sci. Rep.* **9**, 8714. https://doi.org/10.1038/s41598-019-44999-2 (2019).
25. Shen, Y. & Bax, A. Prediction of Xaa-Pro peptide bond conformation from sequence and chemical shifts. *J. Biomol. NMR* **46**, 199–204. https://doi.org/10.1007/s10858-009-9395-y (2010).
26. Führer, S., Zeindl, R. & Tollinger, M. NMR resonance assignments of the four isoforms of the hazelnut allergen Cor a 1.04. *Biomol. NMR Assign.* **14**, 45–49. https://doi.org/10.1007/s12104-019-09918-6 (2020).
27. 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.* **65**, 1606–1612. https://doi.org/10.1021/acs.jafc.6b05752 (2017).
28. Orozco-Navarrete, B. *et al.* Structural bases for the allergenicity of Fra a 1.02 in strawberry fruits. *J. Agric. Food Chem.* **68**, 10951–10961. https://doi.org/10.1021/acs.jafc.9b05714 (2020).
29. Kofler, S. *et al.* Crystallographically mapped ligand binding differs in high and low IgE binding isoforms of birch pollen allergen bet v 1. *J. Mol. Biol.* **422**, 109–123. https://doi.org/10.1016/j.jmb.2012.05.016 (2012).
30. 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.* **277**, 23684–23692. https://doi.org/10.1074/jbc.M202065200 (2002).
31. Neudecker, P. *et al.* Allergic cross-reactivity made visible: solution structure of the major cherry allergen Pru av 1. *J. Biol. Chem.* **276**, 22756–22763. https://doi.org/10.1074/jbc.M101657200 (2001).
32. Mittermaier, A. & Kay, L. E. New tools provide new insights in NMR studies of protein dynamics. *Science* **312**, 224–228. https://doi.org/10.1126/science.1124964 (2006).
33. Wigotzki, M., Steinhart, H. & Paschke, A. Influence of varieties, storage and heat treatment on IgE-binding proteins in hazelnuts (*Corylus avellana*). *Food Agric. Immunol.* **12**, 217–229. https://doi.org/10.1080/09540100050140759 (2000).

34. Nothegger, B. *et al.* Allergen-specific immunotherapy with apples: selected cultivars could be a promising tool for birch pollen allergy. *J. Eur. Acad. Dermatol. Venereol.* https://doi.org/10.1111/jdv.16201 (2020).
35. Blanc, F. *et al.* Further studies on the biological activity of hazelnut allergens. *Clin. Transl. Allergy* **5**, 26. https://doi.org/10.1186/s13601-015-0066-7 (2015).
36. Kofler, S. *et al.* Stabilization of the dimeric birch pollen allergen Bet v 1 impacts its immunological properties. *J. Biol. Chem.* **289**, 540–551. https://doi.org/10.1074/jbc.M113.518795 (2014).
37. Grutsch, S. *et al.* Ligand binding modulates the structural dynamics and compactness of the major birch pollen allergen. *Biophys. J.* **107**, 2972–2981. https://doi.org/10.1016/j.bpj.2014.10.062 (2014).
38. Neudecker, P. *et al.* Solution structure, dynamics, and hydrodynamics of the calcium-bound cross-reactive birch pollen allergen Bet v 4 reveal a canonical monomeric two EF-hand assembly with a regulatory function. *J. Mol. Biol.* **336**, 1141–1157. https://doi.org/10.1016/j.jmb.2003.12.070 (2004).
39. Grutsch, S. *et al.* Conformational flexibility differentiates naturally occurring Bet v 1 isoforms. *Int. J. Mol. Sci.* https://doi.org/10.3390/ijms18061192 (2017).
40. Moraes, A. H. *et al.* Structural basis for cross-reactivity and conformation fluctuation of the major beech pollen allergen Fag s 1. *Sci. Rep.* **8**, 10512. https://doi.org/10.1038/s41598-018-28358-1 (2018).
41. Wagner, S. *et al.* Naturally occurring hypoallergenic Bet v 1 isoforms fail to induce IgE responses in individuals with birch pollen allergy. *J. Allergy Clin. Immunol.* **121**, 246–252. https://doi.org/10.1016/j.jaci.2007.08.006 (2008).
42. Fernandez-Quintero, M. L. *et al.* Conformational selection of allergen-antibody complexes-surface plasticity of paratopes and epitopes. *Protein Eng. Des. Sel* **32**, 513–523. https://doi.org/10.1093/protein/gzaa014 (2019).
43. Yanaka, S., Moriwaki, Y., Tsumoto, K. & Sugase, K. Elucidation of potential sites for antibody engineering by fluctuation editing. *Sci. Rep.* **7**, 9597. https://doi.org/10.1038/s41598-017-10246-9 (2017).
44. Williams, D. C. Jr., Rule, G. S., Poljak, R. J. & Benjamin, D. C. Reduction in the amide hydrogen exchange rates of an anti-lysozyme Fv fragment due to formation of the Fv-lysozyme complex. *J. Mol. Biol.* **270**, 751–762. https://doi.org/10.1006/jmbi.1997.1122 (1997).
45. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293. https://doi.org/10.1007/BF00197809 (1995).
46. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696. https://doi.org/10.1002/prot.20449 (2005).
47. Shen, Y. *et al.* De novo structure generation using chemical shifts for proteins with high-sequence identity but different folds. *Protein Sci.* **19**, 349–356. https://doi.org/10.1002/pro.303 (2010).
48. Shen, Y. *et al.* Consistent blind protein structure generation from NMR chemical shift data. *Proc. Natl. Acad. Sci. USA* **105**, 4685–4690. https://doi.org/10.1073/pnas.0800256105 (2008).
49. Shen, Y., Vernon, R., Baker, D. & Bax, A. D. novo protein structure generation from incomplete chemical shift assignments. *J. Biomol. NMR* **43**, 63–78. https://doi.org/10.1007/s10858-008-9288-5 (2009).
50. Schwieters, C. D., Kuszewski, J. J. & Clore, G. M. Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc.* **48**, 47–62. https://doi.org/10.1016/j.pnmrs.2005.10.001 (2006).
51. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73. https://doi.org/10.1016/s1090-7807(02)00014-9 (2003).
52. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213–223. https://doi.org/10.1007/s10858-009-9333-z (2009).
53. Jorgensen, W. L., Chandrasekhar, J. & Madura, J. D. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935. https://doi.org/10.1063/1.445869 (1983).
54. Case, D. A. *et al. AMBER 2018* (University of California, San Francisco, 2018).
55. Maier, J. A. *et al.* ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255 (2015).
56. Wallnoefer, H. G., Liedl, K. R. & Fox, T. A challenging system: free energy prediction for factor Xa. *J. Comput. Chem.* **32**, 1743–1752. https://doi.org/10.1002/jcc.21758 (2011).
57. Adelman, S. A. & Doll, J. D. Generalized Langevin equation approach for atom/solid-surface scattering: general formulation for classical scattering off harmonic solids. *J. Chem. Phys.* **64**, 2375–2388. https://doi.org/10.1063/1.432526 (1976).
58. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690. https://doi.org/10.1063/1.448118 (1984).
59. Ciccotti, G. & Ryckaert, J. P. Molecular dynamics simulation of rigid molecules. *Comput. Phys. Rep.* **4**, 345–392. https://doi.org/10.1016/0167-7977(86)90022-5 (1986).
60. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092. https://doi.org/10.1063/1.464397 (1993).
61. Bhattacharya, A., Tejero, R. & Montelione, G. T. Evaluating protein structures determined by structural genomics consortia. *Proteins* **66**, 778–795. https://doi.org/10.1002/prot.21165 (2007).
62. Halgren, T. New method for fast and accurate binding-site identification and analysis. *Chem. Biol. Drug Des* **69**, 146–148. https://doi.org/10.1111/j.1747-0285.2007.00483.x (2007).
63. Halgren, T. A. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model* **49**, 377–389. https://doi.org/10.1021/ci800324m (2009).
64. Release 2020-3: Maestro, Schrödinger, LLC, New York, NY, 2020 (2020).
65. Maiti, R., Van Domselaar, G. H., Zhang, H. & Wishart, D. S. SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Res.* **32**, W590-594. https://doi.org/10.1093/nar/gkh477 (2004).
66. Choy, W. Y. *et al.* Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. *J. Mol. Biol.* **316**, 101–112. https://doi.org/10.1006/jmbi.2001.5328 (2002).
67. Schanda, P., Forge, V. & Brutscher, B. Protein folding and unfolding studied at atomic resolution by fast two-dimensional NMR spectroscopy. *Proc. Natl. Acad. Sci. USA* **104**, 11257–11262. https://doi.org/10.1073/pnas.0702069104 (2007).
68. Schanda, P., Brutscher, B., Konrat, R. & Tollinger, M. Folding of the KIX domain: characterization of the equilibrium analog of a folding intermediate using 15N/13C relaxation dispersion and fast 1H/2H amide exchange NMR spectroscopy. *J. Mol. Biol.* **380**, 726–741. https://doi.org/10.1016/j.jmb.2008.05.040 (2008).
69. 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.* **123**, 11341–11352. https://doi.org/10.1021/ja011300z (2001).
70. 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* **112**, 5898–5904. https://doi.org/10.1021/jp074793o (2008).
71. Baldwin, A. J. An exact solution for R2, eff in CPMG experiments in the case of two site chemical exchange. *J. Magn. Reson.* **244**, 114–124. https://doi.org/10.1016/j.jmr.2014.02.023 (2014).
72. Heinzerling, L. *et al.* The skin prick test—European standards. *Clin. Transl. Allergy* **3**, 3. https://doi.org/10.1186/2045-7022-3-3 (2013).
73. Heinzerling, L. M. *et al.* GA(2)LEN skin test study I: GA(2)LEN harmonization of skin prick testing: novel sensitization patterns for inhalant allergens in Europe. *Allergy* **64**, 1498–1506. https://doi.org/10.1111/j.1398-9995.2009.02093.x (2009).

74. IBMCorp, ,. *Released 2017. IBM SPSS Statistics for WIndows, Version 25.0* (IBM Corp, Armonk, 2017).
75. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539. https://doi.org/10.1038/msb.2011.75 (2011).

Acknowledgements

We thank Dr. Katrin Breuker, Dr. Thomas Müller, Christina Meisenbichler, and Michael Palasser for mass spectrometry experiments and Jana Unterhauser for assistance with ELISA experiments. This work was supported by the Austrian Science Fund FWF (P26849 to M.T., P30737 to K.R.L.), and the Austrian Research Promotion Agency FFG (West Austrian BioNMR 858017 to M.T.).

Author contributions

S.F. and M.T. conceived and designed the research. S.F. collected and analyzed all NMR and ELISA data, with assistance by R.Z. for Cor a 1.0402. Structure refinement was done by A.S.K., with assistance by F.H., and B.N. collected the ImmunoCap data. S.F. and A.S.K. prepared the figures. M.T., K.R.L. and N.R. provided equipment for NMR spectroscopy, structure computations and immunologic experiments, respectively. M.T. and S.F. wrote the manuscript, which was reviewed and discussed by all other authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-83705-z.

Correspondence and requests for materials should be addressed to M.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021

Inverse relation between structural flexibility and IgE reactivity of Cor a 1 hazelnut allergens

Sebastian Führer, Anna S. Kamenik, Ricarda Zeindl, Bettina Nothegger, Florian Hofer, Norbert Reider, Klaus R. Liedl, Martin Tollinger*

* Corresponding author: martin.tollinger@uibk.ac.at

Supplementary Methods

Mass spectrometry experiments

For the determination of the accurate mass of the hazelnut Cor a 1.04 isoforms a 7 Tesla Fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometer (Apex Ultra 70, Bruker Daltonics), equipped with an electrospray ionization (ESI) source was used. Unlabeled Cor a 1.0401, Cor a 1.0402, Cor a 1.0403, and Cor a 1.0404 protein samples were desalted five times with ca. 500 µL 100 mM ammonium acetate pH 6.8 and subsequently five times with ca. 500 µL H_2O. Afterwards, the samples were diluted with H_2O/CH_3OH (1:1, v/v) supplemented with 1 % acetic acid to ca. 1 µM. Data analysis was done with Bruker Compass DataAnalysis and Bruker Daltonics Bio Tools or mMass[1] (version 3.0.0).

CS-Rosetta model generation

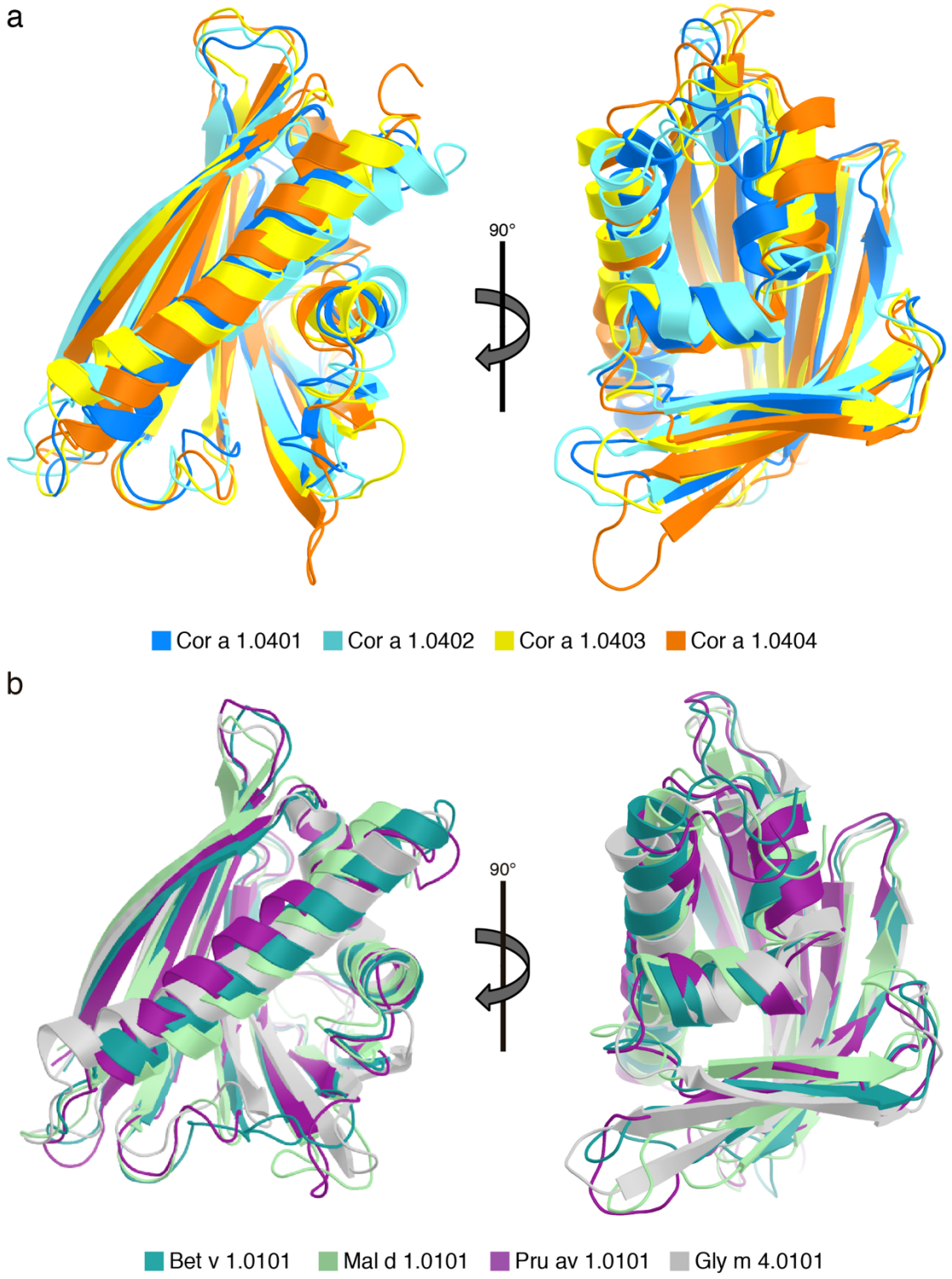
For each isoform, 3000 structure models were generated by the program CS-Rosetta[2-5], using the NMR resonance assignments[6]. The Cα RMSD values for the 10 lowest energy structures were all below 3.5, 2.9, 3.1, and 3.3 Å for isoforms Cor a 1.0401, Cor a 1.0402, Cor a 1.0403, and Cor a 1.0404, respectively, indicating good convergence of the models. The lowest energy model for each isoform was used for initial identification and manual assignment of NOE cross-peaks.

NMR relaxation experiments

Experiments for measuring longitudinal and rotating-frame relaxation rates (R_1 and $R_{1\rho}$) and the $^{15}N\{^1H\}$ steady-state NOE (nuclear Overhauser effect)[7] for Cor a 1.0401 and Cor a 1.0404 were performed at 500 MHz. Relaxation periods for measuring R_1 rates were 11.1, 55.5, 111.0, 222.0, 333.0, 444.0, 555.0, and 666.0 ms with repeat experiments at 55.5 ms and 333.0 ms. For measuring $R_{1\rho}$ rates, relaxation periods were 10.0, 20.0, 30.0, 40.0, 50.0, 60.0, 70.0, and 90.0 ms, with repeats at 20.0 ms and 50.0 ms. For $^{15}N\{^1H\}$ steady-state NOE measurements two spectra were recorded, one with an interscan delay (d1) of 3 s and proton saturation for 1.25 s and the second experiment with d1 set to 4.25 s without proton saturation. The relaxation rates R_1 and $R_{1\rho}$ were determined from exponential fits using an in-house MATLAB fitting script. Transverse relaxation rates R_2 were derived from R_1 and $R_{1\rho}$ values as described [8], employing a 1.5 kHz spin-lock field. Order parameters S^2 were determined with the program FAST-Modelfree (version 4.15)[9], which has the model-free approach implemented[10-12].

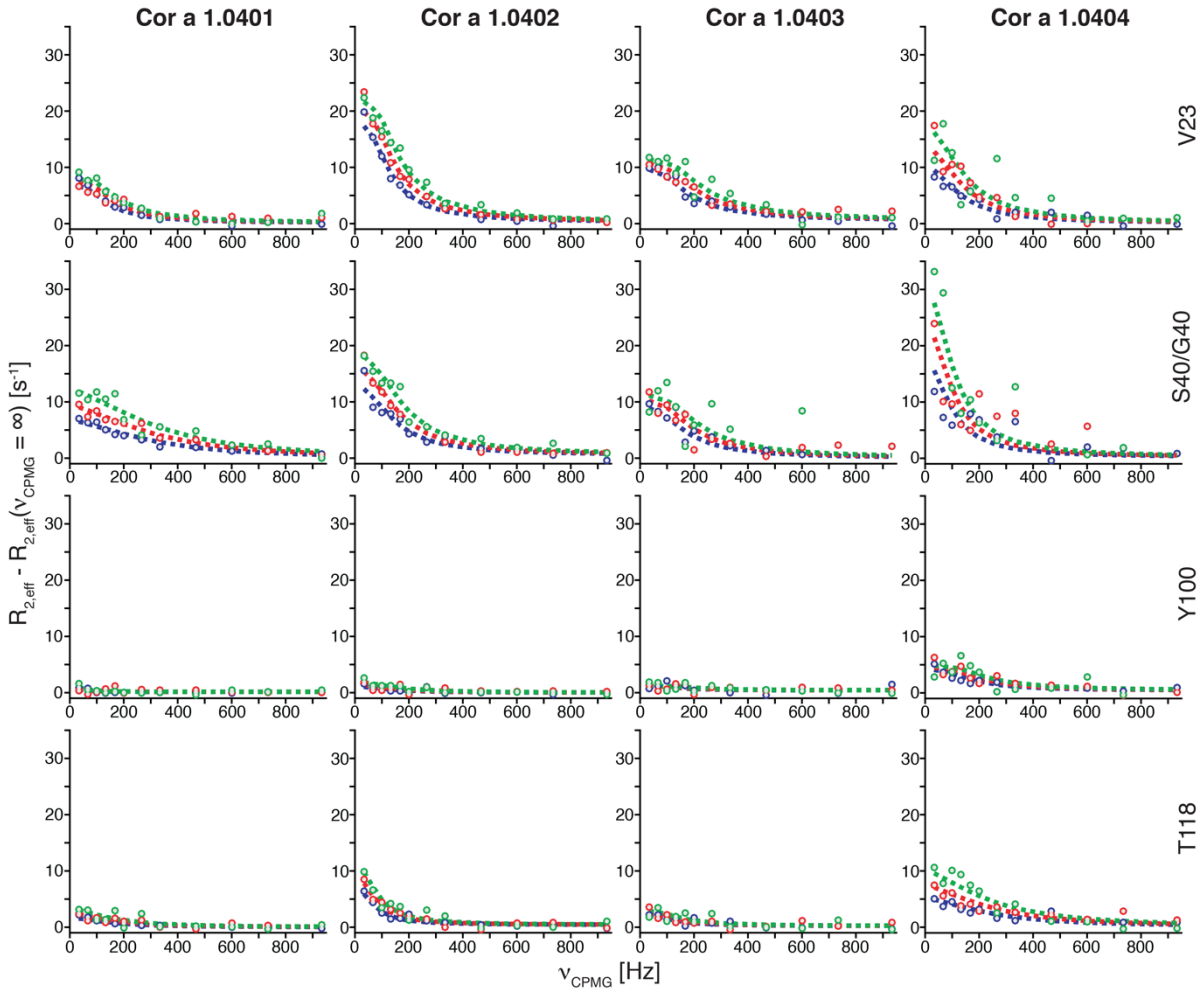
Supplementary Figures

Supplementary Figure S1



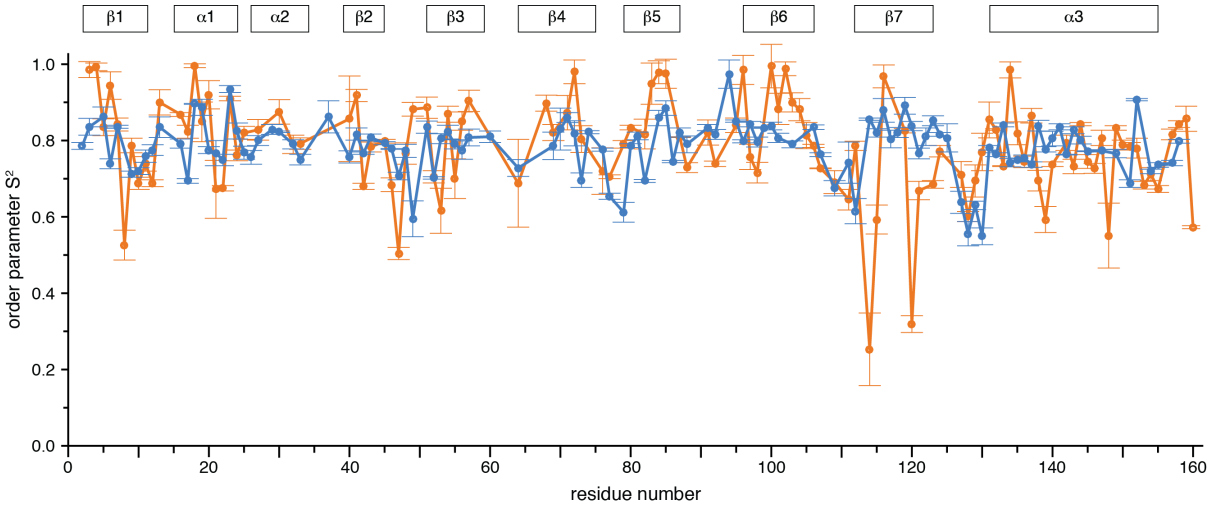
Structural comparison of the four hazelnut Cor a 1.04 isoforms with other PR-10 allergens. (a) Overlay of the solution structures of Cor a 1.0401 (blue, 6Y3H), Cor a 1.0402 (light blue, 6Y3I), Cor a 1.0403 (yellow, 6Y3K), and Cor a 1.0404 (orange, 6Y3L). (b) Overlay of the structures of Bet v 1.0101 (teal, 4A88), Mal d 1.0101 (pale green, 5MMU), Pru av 1.0101 (purple, 1E09), and Gly m 4.0101 (gray, 2K7H).

Supplementary Figure S2



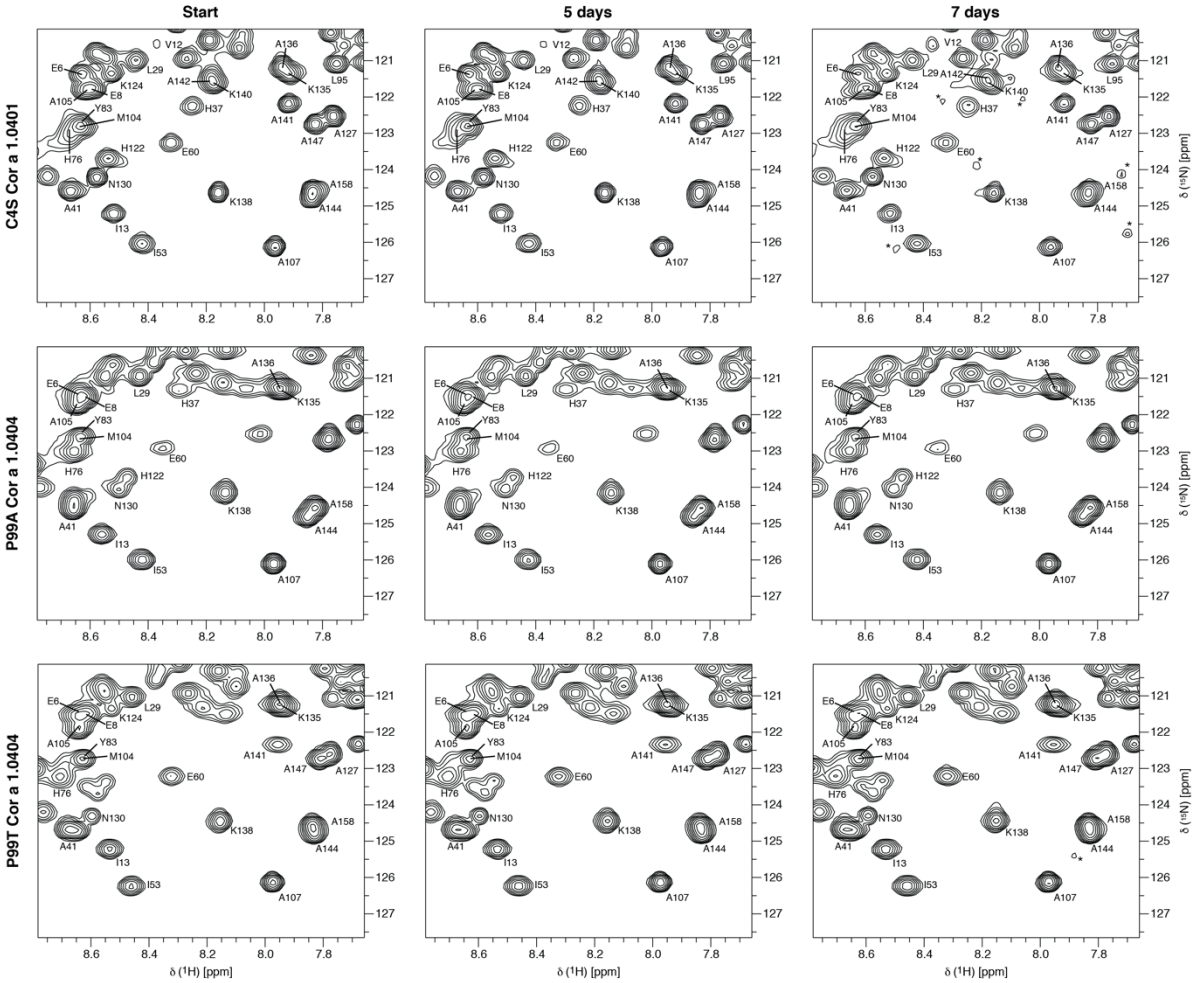
Backbone amide NH relaxation dispersion profiles of the four hazelnut Cor a 1.04 isoforms. $R_{2,eff}$ - $R_{2,eff}$ ($\nu_{CPMG} = \infty$) values at variable CPMG field strengths ν_{CPMG}, acquired at 500 MHz (blue), 600 MHz (red), and 700 MHz (green) are shown along with best-fit dashed lines. Data are shown for four representative amino acid residues in the short helix α1 (Val23), in strand β2 (Ser40 in Cor a 1.0401 and Cor a 1.0404, Gly40 in Cor a 1.0402 and Cor a 1.0403) and in the vicinity of position 99 in strands β6 (Tyr100) and β7 (Thr118). Relaxation dispersion amplitudes ($\Delta R_{2,eff}$) used in Figure 3 are given by $R_{2,eff} - R_{2,eff}(\nu_{CPMG} = \infty)$ at $\nu_{CPMG} = 0$.

Supplementary Figure S3



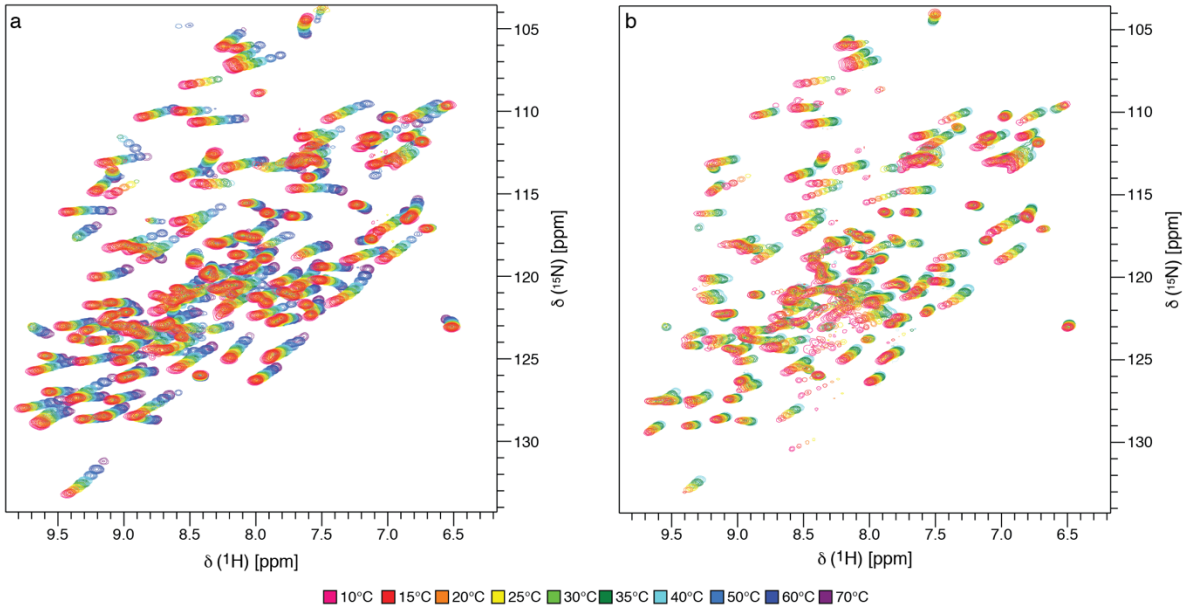
NMR-derived backbone amide NH order parameters S^2 of Cor a 1.0401 (blue) and Cor a 1.0404 (orange). Residue specific S^2 values, obtained with the program FAST-Modelfree[9], are depicted as circles and error bars are shown. Solid lines connecting residues for which experimental data are available are drawn for better visualization. Low order parameters indicate high flexibility on the picosecond to nanosecond time scale. Secondary structure elements of the hazelnut Cor a 1.04 isoforms are shown on top.

Supplementary Figure S4



Time-dependent stability of three mutant forms of Cor a 1.0401 and Cor a 1.0404. Sections of ^{1}H-^{15}N-HSQC spectra of C4S Cor a 1.0401 (top), P99A Cor a 1.0404 (middle), and P99T Cor a 1.0404 (bottom) directly after purification, as well as five and seven days later. Tentative assignments are indicated by single letter codes and signals labeled with an asterisk indicate aggregation and degradation.

Supplementary Figure S5



Temperature-dependent stability of Cor a 1.04 isoforms. Temperature dependent 700 MHz ^{1}H-^{15}N-HSQC spectra of Cor a 1.0401 (a) and Cor a 1.0404 (b) are superimposed (referenced to DSS). Cor a 1.0401 spectra could be acquired up to 70 °C, while Cor a 1.0404 resonances vanished above 40 °C.

Supplementary Figure S6

a

```
Cor a 1.0401        1  ATGGGCGTGTTCTGCTACGAAGATGAGGCGACCAGCGTTATCCCGCCGGCGCGTCTGTTC 60
                       ||||||||||||*||||||||||||||||||||||||||||||||||||||||||||||||
C4S Cor a 1.0401    1  ATGGGCGTGTTCTCCTACGAAGATGAGGCGACCAGCGTTATCCCGCCGGCGCGTCTGTTC 60

Cor a 1.0401       61  AAAAGCTTTGTGCTGGACGCGGATAACCTGATTCCGAAGGTTGCGCCGCAGCACTTTACC 120
                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
C4S Cor a 1.0401   61  AAAAGCTTTGTGCTGGACGCGGATAACCTGATTCCGAAGGTTGCGCCGCAGCACTTTACC 120
```

b

```
C4S Cor a 1.0401   241 AAGTACTGCTATAGCATCATTGAAGGTGGCCCGCTGGGTCACACCCTGGAAAAAATCAGC 300
                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||**|
P99A Cor a 1.0404  241 AAGTACTGCTATAGCATCATTGAAGGTGGCCCGCTGGGTCACACCCTGGAAAAAATCGCC 300

C4S Cor a 1.0401   301 TACGAGATTAAAATGGCGGCGGCGCCGCACGGTGGCGGTAGCATCCTGAAGATTACCAGC 360
                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
P99A Cor a 1.0404  301 TACGAGATTAAAATGGCGGCGGCGCCGCACGGTGGCGGTAGCATCCTGAAGATTACCAGC 360
```

c

```
C4S Cor a 1.0401   241 AAGTACTGCTATAGCATCATTGAAGGTGGCCCGCTGGGTCACACCCTGGAAAAAATCAGC 300
                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||*|
P99T Cor a 1.0404  241 AAGTACTGCTATAGCATCATTGAAGGTGGCCCGCTGGGTCACACCCTGGAAAAAATCACC 300

C4S Cor a 1.0401   301 TACGAGATTAAAATGGCGGCGGCGCCGCACGGTGGCGGTAGCATCCTGAAGATTACCAGC 360
                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
P99T Cor a 1.0404  301 TACGAGATTAAAATGGCGGCGGCGCCGCACGGTGGCGGTAGCATCCTGAAGATTACCAGC 360
```

Coding sequence alignment of the Cor a 1.04 mutants (a) C4S Cor a 1.0401, (b) P99A Cor a 1.0404, and (c) P99T Cor a 1.0404. The upper line indicates the sequence of the template and the lower line the sequence of the site-directed mutagenesis product. Lines represent concordant nucleotides, whereas different nucleotides are highlighted by asterisks. Successful mutations (codons) are highlighted by red boxes. Sequencing was performed by Microsynth AG (Balgach, Switzerland).

References

1 Strohalm, M., Kavan, D., Novak, P., Volny, M. & Havlicek, V. mMass 3: a cross-platform software environment for precise analysis of mass spectrometric data. *Anal Chem* **82**, 4648-4651, doi:10.1021/ac100818g (2010).

2 Bax, A., Ikura, M., Kay, L. E., Torchia, D. A. & Tschudin, R. Comparison of different modes of two-dimensional revers-correlation NMR for the study of proteins. *J Magn Reson* **86**, 304-318, doi:10.1016/022-2364(90)90262-8 (1990).

3 Shen, Y. *et al.* De novo structure generation using chemical shifts for proteins with high-sequence identity but different folds. *Protein Sci* **19**, 349-356, doi:10.1002/pro.303 (2010).

4 Shen, Y. *et al.* Consistent blind protein structure generation from NMR chemical shift data. *Proc Natl Acad Sci U S A* **105**, 4685-4690, doi:10.1073/pnas.0800256105 (2008).

5 Shen, Y., Vernon, R., Baker, D. & Bax, A. De novo protein structure generation from incomplete chemical shift assignments. *J Biomol NMR* **43**, 63-78, doi:10.1007/s10858-008-9288-5 (2009).

6 Führer, S., Zeindl, R. & Tollinger, M. NMR resonance assignments of the four isoforms of the hazelnut allergen Cor a 1.04. *Biomol NMR Assign* **14**, 45-49, doi:10.1007/s12104-019-09918-6 (2020).

7 Mittermaier, A. & Kay, L. E. New tools provide new insights in NMR studies of protein dynamics. *Science* **312**, 224-228, doi:10.1126/science.1124964 (2006).

8 Trott, O. & Palmer, A. G., 3rd. R1rho relaxation outside of the fast-exchange limit. *J Magn Reson* **154**, 157-160, doi:10.1006/jmre.2001.2466 (2002).

9 Cole, R. & Loria, J. P. FAST-Modelfree: a program for rapid automated analysis of solution NMR spin-relaxation data. *J Biomol NMR* **26**, 203-213, doi:10.1023/a:1023808801134 (2003).

10 Clore, G. M. *et al.* Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *Journal of the American Chemical Society* **112**, 4989-4991, doi:10.1021/ja00168a070 (1990).

11 Lipari, G. & Szabo, A. Model-free apporach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *Journal of the American Chemical Society* **104**, 4546-4559, doi:10.0121/ja00381a009 (1982).

12 Lipari, G. & Szabo, A. Model-free approach to the interpretation of nucelar magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *Journal of the American Chemical Society* **104**, 4559-4570, doi:10.021/ja00381a010 (1982).