



OPEN

## Allergenicity and structural properties of new Cor a 1 isoallergens from hazel identified in different plant tissues

Julian M. Hendrich<sup>1</sup>, Andreas Reuter<sup>2</sup>, Thessa P. Jacob<sup>1</sup>, Hatice Kara<sup>1</sup>, Sherine Amer<sup>1</sup>, Katharina Rödel<sup>1</sup> & Birgitta M. Wöhrl<sup>1</sup>✉

The hazel allergen Cor a 1 is a PR-10 protein, closely related to the major birch pollen allergen Bet v 1. Hazel allergies are caused by cross-reactive IgE antibodies originally directed against Bet v 1. Despite the importance of PR-10 proteins in allergy development, their function and localization in the plant remain largely elusive. Therefore, the presence of Cor a 1 mRNA and proteins was investigated in different tissues, i.e., the female flower, immature and mature nuts, catkins, and pollen. Four yet unknown Cor a 1 isoallergens, i.e., Cor a 1.0501–1.0801, and one new Cor a 1.03 variant were discovered and characterized. Depending on the isoallergen, the occurrence and level of mRNA expression varied in different tissues, suggesting different functions. Interestingly, Cor a 1.04 previously thought to be only present in nuts, was also detected in catkins and pollen. The corresponding Cor a 1 genes were expressed in *Escherichia coli*. The purified proteins were analysed by CD and NMR spectroscopy. Immunoblots and ELISAs to determine their allergenic potential showed that the new proteins reacted positively with sera from patients allergic to birch, hazel and elder pollen and were recognized as novel isoallergens/variants by the WHO/IUIS Allergen Nomenclature Sub-Committee.

**Keywords** PR-10 protein, Bet v 1, Cor a 1 isoallergens, Gene expression, LC-MS<sup>E</sup>, Hazel genome, *Corylus avellana*, Immunoblot, ELISA, RT-PCR,  $T_m$ -values

Anemophilous or wind-driven pollen from certain trees can trigger allergic diseases and therefore impose serious health problems. An estimated 10–30% of the global population are afflicted by seasonal allergies and asthma caused by pollen, and the number continues to rise<sup>1</sup>. Changes in life style, environmental pollution or dietary habits as well as climate change appear to play a role<sup>2</sup>. One of the best studied allergens is the major birch pollen allergen Bet v 1, which is the founding member of the Bet v 1-like superfamily of proteins. This superfamily has representatives in all three domains of life, archaea, bacteria and eukaryotes<sup>3</sup>, and is comprised of 11 subfamilies. Bet v 1 itself belongs to the subfamily of the pathogenesis related proteins 10 (PR-10)-like family of proteins<sup>3–5</sup>. Most of the Bet v 1 homologous allergens from pollen and plants, including the Cor a 1 allergen from hazel, belong to this subfamily. Whereas “real” PR-10 proteins, which have been detected in various plants are only induced in pathological stress situations, suggesting a role in plant defence responses, PR-10-like proteins are constitutively expressed, indicating a more general role in plant development<sup>3,6,7</sup>.

In most cases allergic reactions to Bet v 1 and homologous proteins are not life threatening and usually result in mild symptoms like rhinitis, burning and itching of the tongue and oral mucosa, which are described as the so-called oral allergy syndrome (OAS) or pollen food syndrome (PFS). However, allergy induced asthma can also occur<sup>4,8–10</sup>. Most pollen allergy sufferers are primarily sensitized to Bet v 1<sup>11</sup>. Over time, many patients develop cross-reactivity of IgE antibodies originally directed against Bet v 1 proteins to homologous PR-10 allergens, which are found in almost all angiosperms and have at least 50% sequence identity amongst them<sup>6</sup>. Thus, birch pollen allergic individuals frequently experience IgE mediated allergic reactions upon contact with various plant food products harbouring homologous PR-10 proteins, i.e. fruits like apple, cherry or peach, vegetables and roots such as celeriac or carrots, and nuts, e.g. almonds and hazelnuts<sup>4,12</sup>.

<sup>1</sup>Lehrstuhl Biochemie IV - Biophysikalische Chemie, Universität Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany. <sup>2</sup>Division of Allergology, Paul-Ehrlich-Institut, 63225 Langen, Germany. ✉email: birgitta.woehr@uni-bayreuth.de

The structural similarity of PR-10 allergens is the basis of the development of cross-reactive IgE antibodies<sup>13</sup>. The three-dimensional structure of many PR-10 and PR-10-like proteins is known and is almost identical<sup>6</sup>. They exhibit conserved, common structural elements, suggesting a general and maybe indispensable function in plants. They consist of a seven-stranded, antiparallel  $\beta$ -sheet, a long C-terminal  $\alpha$ -helix and two short helices arranged in V-shape, which together form a predominantly hydrophobic or amphiphilic pocket, being composed of a combination of hydrophilic and hydrophobic residues, that can bind small ligands<sup>14–18</sup>. A variety of small molecules were shown to bind to different PR-10 allergens, e.g. flavonoids, steroids or cytokinins, suggesting that they are involved in transport and/or storage of secondary plant compounds, UV protection and germination<sup>15,17,19–23</sup>. However, for most PR-10 allergens the precise functions in the plant remain elusive.

IgE antibodies of Bet v 1 allergic individuals can also cross-react with the corresponding Cor a 1 allergen from hazel (*Corylus avellana*)<sup>24,25</sup>. Consequently, contact with hazel pollen can trigger an allergic reaction similar to that described above, and in addition, hazelnut is a common cause of food allergy.<sup>10,25–28</sup>

Due to the high relevance for the development of cross-allergies, it is imperative to obtain a more complete picture of the occurrence of known and novel Cor a 1 isoallergens in different hazel tissues. So far, four Cor a 1 isoallergens (>67% amino acid sequence identity), termed Cor a 1.01 to Cor a 1.04, have been identified<sup>29,30</sup>. Four different Cor a 1.01 variants (>90% identity), named Cor a 1.0101 to Cor a 1.0104, have been found in hazel pollen<sup>29</sup>. Furthermore, four variants of the isoallergen Cor a 1.04 have been detected in nuts<sup>30</sup>. In contrast, Cor a 1.02 and Cor a 1.03 have been suggested to occur in leave tissue, however in that work the corresponding genes were only detected in the genome, but no expression analyses, neither on the mRNA nor on the protein level, were performed<sup>31</sup>. Thus, the presence of those isoallergens in different hazel tissues and in pollen has not yet been confirmed.

A recent study produced a fully assembled genome sequence and annotation for the hazelnut species *C. avellana* cv ‘Tombul’, an important Turkish variety. In addition to the known Cor a 1 genes, several new Cor a 1 related sequences were detected, some of which with only 40–50% identity to Cor a 1<sup>32</sup>. The study implies that there are many more than the already known Cor a 1 isoallergens and variants. However, similar to most other PR-10 allergens, the exact functions and localization of Cor a 1 proteins in the plant are not known. Some of the Cor a 1 proteins appear to be involved in transport and/or storage of flavonoids. We have previously identified a natural ligand of the Cor a 1.04 isoallergen, quercetin-3-O-(2<sup>o</sup>-O- $\beta$ -D-glucopyranosyl)- $\beta$ -D-galactopyranoside (Q3O-(Glc)-Gal), and suggested a function of the protein/ligand complex in fertilization<sup>17</sup>.

As mentioned above, some Cor a 1 proteins have been shown to be localized in pollen or nuts<sup>29</sup>. However, no comprehensive study has yet been conducted to demonstrate which Cor a 1 proteins are present in different plant tissues. Therefore, we wanted to identify new Cor a 1 isoallergens and analyse their biochemical and biophysical properties as well as the allergenic potential to contribute to a better understanding of hazel allergy.

## Results and discussion

### Identification and mRNA expression profiles of Cor a 1 isoallergens

There appear to be more Cor a 1 isoallergens and variants present in hazel than have been described so far, as indicated by similarities of genes found in the hazel genome<sup>32</sup>. Although the presence of the isoallergens Cor a 1.01 and Cor a 1.04 in pollen and kernel, respectively, has been shown, the expression and localization of the other known isoallergens Cor a 1.02 and Cor a 1.03 in different tissues has not yet been clarified.

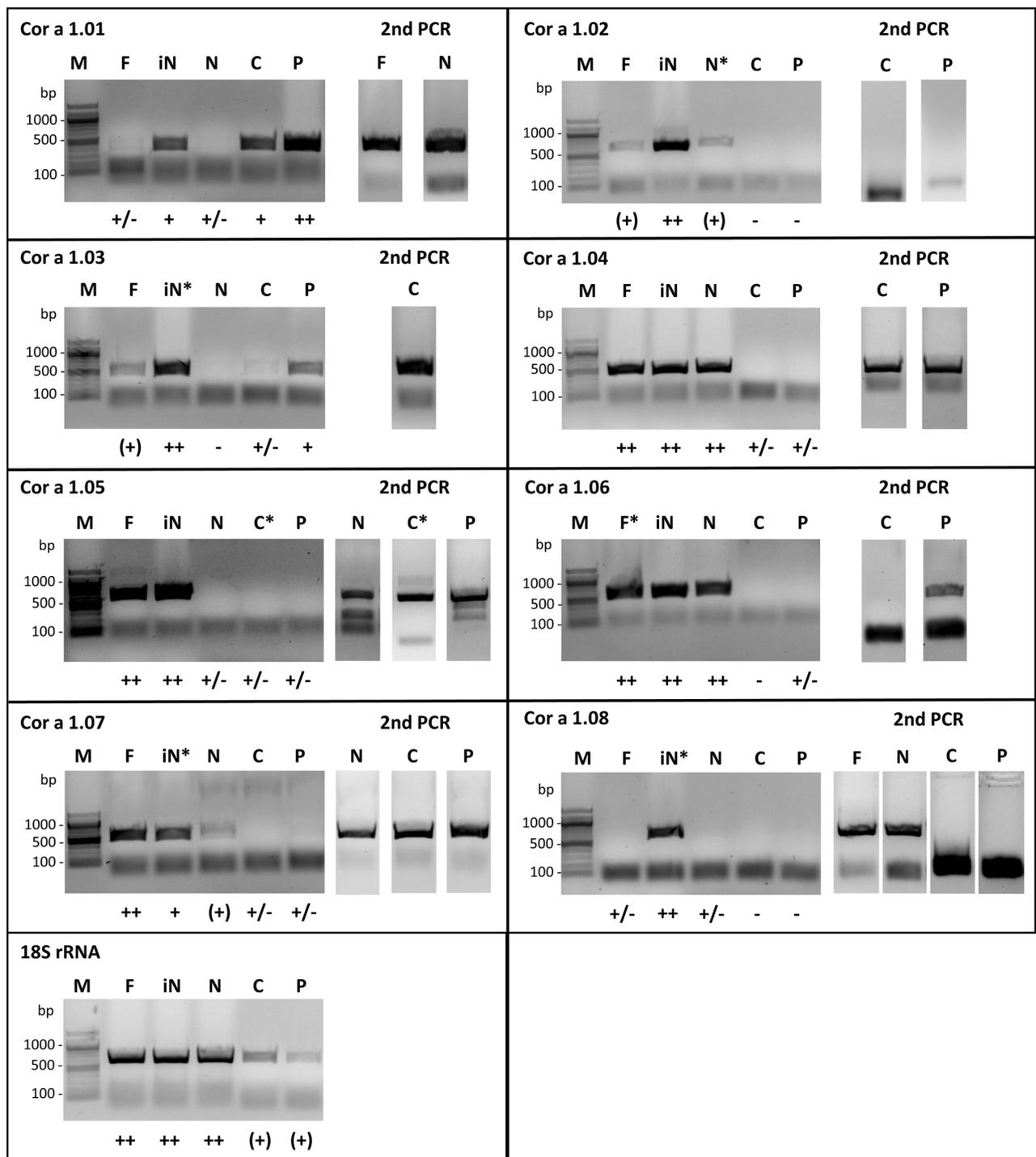
To shed more light on the occurrence and distribution of Cor a 1 isoallergens, we determined the presence of Cor a 1 mRNA in different hazel plant tissues and in pollen. Since the fully assembled and annotated genome sequence of a European hazel species, *C. avellana* cv ‘Tombul’ was recently published<sup>32</sup> it was possible to search specifically for novel Cor a 1 genes and their expression in different hazel tissues.

Total RNA was isolated from the female flower, immature and mature nuts, male catkins, and pollen. After cDNA production with an mRNA-specific poly(T) primer, PCR was performed using specific primers binding either to the 5' or 3' end of the gene (only Cor a 1.0302) or to the 5' and 3' untranslated regions (UTRs). Since the UTR regions of variants of a certain isoallergen are rather similar or even identical<sup>32</sup>, it was not possible to distinguish different variants using this method. Thus, only the existence of a specific isoallergen, possibly consisting of several variants, could be determined in a specific plant material (Figs. 1, S6). However, this procedure allowed us to use the PCR sample for cloning and for sequencing the complete amplified gene for the identification of new isoallergens (Table S1). To verify that the amplicates contained Cor a 1 DNA, all PCR probes were subjected to DNA sequencing. To detect even extremely low specific mRNA levels, a 2nd PCR amplification step was carried out in such cases with the gel purified DNA of the first PCR.

So far, the Cor a 1.01 protein has been detected in hazel pollen<sup>29</sup>. However, the mRNA expression analyses presented here indicate that Cor a 1.01 is, in addition, expressed in minimal amounts in the female flower and in the mature nut as indicated by the 2nd PCR amplification step. Significantly stronger specific mRNA levels could be detected in the immature nut as well as in catkins and pollen.

Cor a 1.02 mRNA was present in the female flower and the mature nut, and high mRNA concentrations were visible in the immature nut, but no specific mRNA was discovered in the male tissues. Moreover, a new, putative Cor a 1.02 variant, designated Cor a 1.02- Cav01g11820 was detected via sequencing of the PCR product (Figure S1). However, since it was not possible to purify the protein from *E. coli* to determine its allergenic properties, this variant was not considered further in our studies.

Cor a 1.03 mRNA was present in all tissues with the exception of the mature nut even though high mRNA levels were detected in the immature nut. Although we could detect a very faint band after the first PCR round of the mature nut, sequencing of the amplicate of the 2<sup>nd</sup> PCR revealed no Cor a 1.03 gene sequence (Fig. 1). In addition, using LC-MS<sup>E</sup> (liquid chromatography–mass spectrometry<sup>elevated energy</sup>) variants of Cor a 1.01 and Cor a 1.03 could also be discovered on the protein level by LC-MS<sup>E</sup> in pollen extract (Figure S2).



**Figure 1.** Expression of Cor a 1 genes in different hazel tissues. The RT-PCR products using gene specific primers to determine Cor a 1 mRNA expression were loaded on 1.5% agarose gels. In case no band was visible, a gel slice at the corresponding height was cut out. The purified mRNA was then used for the 2nd PCR. F: female flower; iN: immature nut; N: mature nut; C: catkin; P: pollen; ((+)) very low, (+): low, +: moderate, ++: strong expression levels. The RT-PCR samples used for cloning of the genes into expression vectors are marked with \*M: 100 bp DNA standard (New England Biolabs; Frankfurt, Germany). 18S rRNA isolated in a similar fashion was used as a control.

It has been shown previously that Cor a 1.04 is present in nuts<sup>30</sup>. Since Cor a 1.04 was found to bind to a ligand isolated from pollen extracts it was suggested that it might play a role in fertilization<sup>17</sup>. Thus, it was interesting to determine whether Cor a 1.04 is localized in other tissues as well. Surprisingly, the 2nd PCR amplification step

followed by sequencing revealed the presence of Cor a 1.04 specific mRNA also in catkins and pollen, albeit at very low amounts (Fig. 1). Thus, we show for the first time that Cor a 1.04 expression also occurs in the male tissues. This result provides a reasonable explanation why a ligand isolated from pollen extracts bound specifically to Cor a 1.0401 which was thought to be only present in the nut.

In addition, Cor a 1.04 mRNA could also be detected in the female flower and the immature nut (Fig. 1). Furthermore, LC-MS<sup>E</sup> data showed the presence of the Cor a 1.04 protein in extracts of the mature nut (Figure S2). However, only a few peptides could be confirmed, suggesting that protein isolation from nuts might be less efficient than mRNA extraction. In addition, no amplification step is performed before the MS analysis. This might also explain why other isoallergens (Fig. 1) could not be identified at the protein level by LC-MS<sup>E</sup>. Moreover, it is also possible that protein amounts in nuts are lower and do not necessarily correspond to mRNA levels.

Cor a 1.05 mRNA was found in the female flower and in the immature nut, however, in the mature nut, as well as in catkins and pollen specific mRNA was only visible after the 2nd PCR procedure (Fig. 1).

Cor a 1.06 mRNA was detected in the female flower, as well as in the immature and mature nut, and was also visible in pollen after the 2nd PCR. Cor a 1.07 mRNA levels were very low in all tissues. In contrast, Cor a 1.08 mRNA concentration was high only in the immature nut, whereas it was weak in the female flower and mature nut, and no expression could be verified in the male tissues (Fig. 1).

It is important to note, that our intention was to identify new expressed Cor a 1 genes by detecting the corresponding mRNAs in different tissues and to clone the genes. Our data represent only one specific time point, which means that Cor a 1 isoallergens that were not visible in our experiments might nevertheless be present at an earlier or later stage of development.

The observation that the mRNA expression of the individual isoallergens differs so strongly implies different properties and functions in the plant, e.g. during fertilization and kernel development, that still need to be explored. Different roles of Cor a 1 isoallergens have already been suggested as only the Cor a 1.04 isoallergen but not Cor a 1.01 binds the flavonoid derivative Q3O-(Glc)-Gal<sup>17</sup>. In summary, the data show that there are more Cor a 1 isoallergens than previously known and that their presence in mature nuts and pollen could play a role in hazelnut allergy (see below).

### Identification of additional variants

In order to identify new isoallergens and variants, the amplified DNA products from specific tissues were cloned into the pCR-Blunt vector (Thermo Fisher, Schwerte, Germany) and sequenced. It turned out that several highly similar variants of a certain isoallergen could be detected in the different tissues tested (Figures S1, and S3). For some of them, only gene fragments could be amplified (data not shown). Due to the high similarity of the variants only one variant per isoallergen was selected and included in our studies to verify the biochemical and biophysical properties as well as the allergenicity (Figure S1). Furthermore, we identified variants of several isoallergens that could not be detected in the genome of the Turkish cultivar “Tombul”<sup>32</sup> by BLAST analyses, indicating genetic differences of *C. avellana* used in our studies. These genes are labelled with a GenBank accession number only, whereas a Cav01 number represents genes also found in the “Tombul cultivar”. Our data indicate that for each isoallergen additional potential variants are present (Figure S3). However, due to their high similarity we did not include different variants of a certain isoallergen in our study.

As described above, we were able to identify four new isoallergens, namely Cor a 1.0501, Cor a 1.0601, Cor a 1.0701 and Cor a 1.0801 on the mRNA level (Figure S1). In addition to those isoallergens a new Cor a 1.03 variant, designated Cor a 1.0302 was detected. These proteins were included in biophysical and immunological studies. Due to their allergenic potential (see below), the four new isoallergens were recognized by the WHO/IUIS Allergen Nomenclature Sub-Committee and named as described.

Amino acid sequence alignments and the identity matrix of Cor a 1 isoallergens and variants revealed that the new proteins possess sequence identities to the already known isoallergens of 46.88 – 84.38%. The new Cor a 1.0302 variant is 96.23% identical to Cor a 1.0301 (Table 1).

### Biochemical and biophysical analyses

The identified genes were cloned into the expression vectors pETGB1a or pET11a (only Cor a 1.0302) and expressed in *E. coli* (Table S2). The SDS polyacrylamide gel of the purified proteins exhibits apparent molecular masses of ca. 16 kDa to 18 kDa (Figure S4).

To determine the refolding capacity and stability of the proteins, CD spectra were recorded (Fig. 2, panel A, B). At 25 °C the spectra of the recombinant (r)Cor a 1 isoallergens are characteristic for proteins harboring  $\alpha$ - and  $\beta$ -like secondary structures (Fig. 2, panel A). After heating to 95 °C, rCor a 1.0302, 1.0501, 1.0601, and 1.0701 unfolded as indicated by the minimum at 200 nm. In contrast, rCor a 1.0801 appeared to be more stable. The curve after heating to 95 °C (Fig. 2, panel A, red curve) changed only slightly and there was no discrete minimum at 200 nm, signifying that only a few sections of the protein became unfolded. After cooling to 25 °C, rCor a 1.0302, 1.0501, 1.0601 and 1.0701 almost completely refolded whereas rCor a 1.0801 was still folded and the proportion of unfolded protein moieties decreased as compared to the spectrum at 95 °C (Fig. 2, panel A, green curves).

rCor a 1.0302, 1.0601 and 1.0701 revealed relatively similar  $T_m$ -values of 68.29, 54.08 or 62.68 °C. Cor a 1.0501 exhibited the highest  $T_m$ -value of 84 °C (Fig. 2, panel B). rCor a 1.0801 displayed a biphasic melting behaviour during heating to 95 °C. However, complete denaturation could not be achieved. Due to the small differences of the denaturation and renaturation curves, the noise in the measurement of the melting curve of Cor a 1.0801 is quite high (Fig. 2, panel B).

Cor a 1.0801 shows a lower sequence identity to Bet v 1.0101 and also to most of the other Cor a 1 isoallergens (Table 1). This difference could be the reason for its higher thermostability. Previously, the structure of the

Identity matrix	Bet v 1.0101	Cor a 1									
		0101	0201	0301	0302	0401	0501	0601	0701	0801	
Bet v 1.0101	100.00	72.50	83.12	71.25	69.18	67.50	51.25	78.12	54.38	46.88	
Cor a 1	0101	72.50	100.00	75.00	65.00	64.78	63.13	48.75	78.12	50.00	46.88
	0201	83.12	75.00	100.00	71.25	69.18	70.00	50.63	84.38	53.12	50.00
	0301	71.25	65.00	71.25	100.00	96.23	71.25	48.75	69.38	58.12	53.12
	0302	69.18	64.78	69.18	96.23	100.00	71.07	50.31	68.55	59.12	54.09
	0401	67.50	63.13	70.00	71.25	71.07	100.00	52.50	68.12	60.62	51.25
	0501	51.25	48.75	50.63	48.75	50.31	52.50	100.00	48.75	51.88	55.62
	0601	78.12	78.12	84.38	69.38	68.55	68.12	48.75	100.00	51.25	47.50
	0701	54.38	50.00	53.12	58.12	59.12	60.62	51.88	51.25	100.00	46.88
0801	46.88	46.88	50.00	53.12	54.09	51.25	55.62	47.50	46.88	100.00	

**Table 1.** Identity matrix of Cor a 1 isoallergens and variants. Numbers indicate the amino acid sequence identity in %

hypothetical protein TTHA0849 from the thermophilic bacterium *Thermus thermophilus* had been determined, which also shows the typical PR-10 fold of Bet v 1<sup>33</sup>. However, sequence comparisons revealed no evidence of conserved regions in Cor a 1.0801 and TTHA0849 that might explain its thermostability. Despite the similar three-dimensional structures, the sequence identity between Cor a 1.08 and TTHA0849 is only 20.98%, which is even lower than that of all the other Cor a 1 isoallergens to TTHA0849, except Cor a 1.05 (20.42%) (Figure S7).

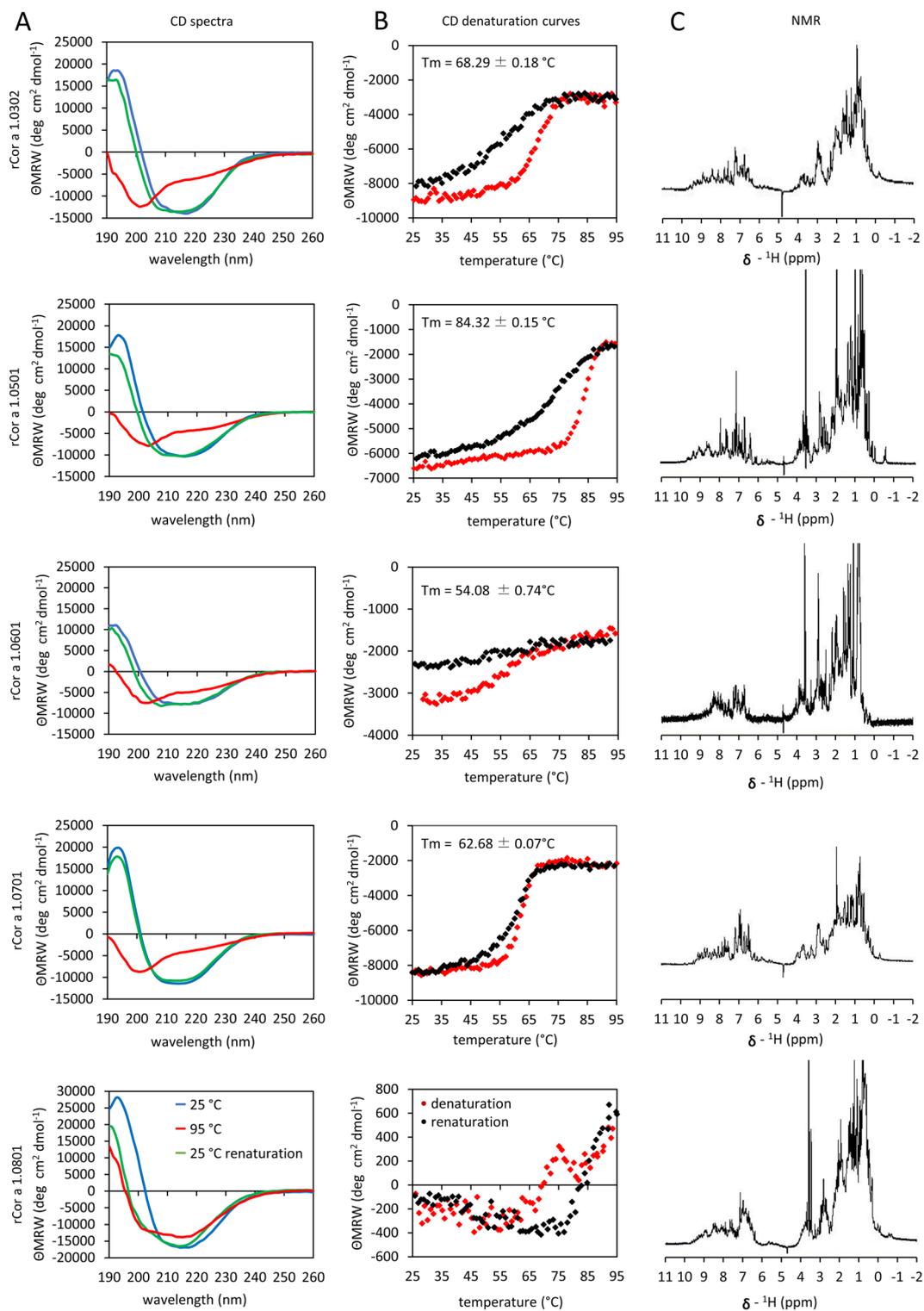
While CD spectra can be used to determine secondary structures, 1D NMR spectroscopy can reveal three-dimensional folding of a protein. The 1D NMR spectra showed dispersion of the amide proton signals and high-field shifted signals of the methyl groups, which confirmed a three-dimensional folding for all isoallergens (Fig. 2, panel C).

These data suggest that the high stability and refolding capacity of the proteins could have an impact on their allergenicity. Heating might not completely abolish the allergenicity in hazel nut containing food stuff as Cor a 1 proteins could refold if the food is cooled down.

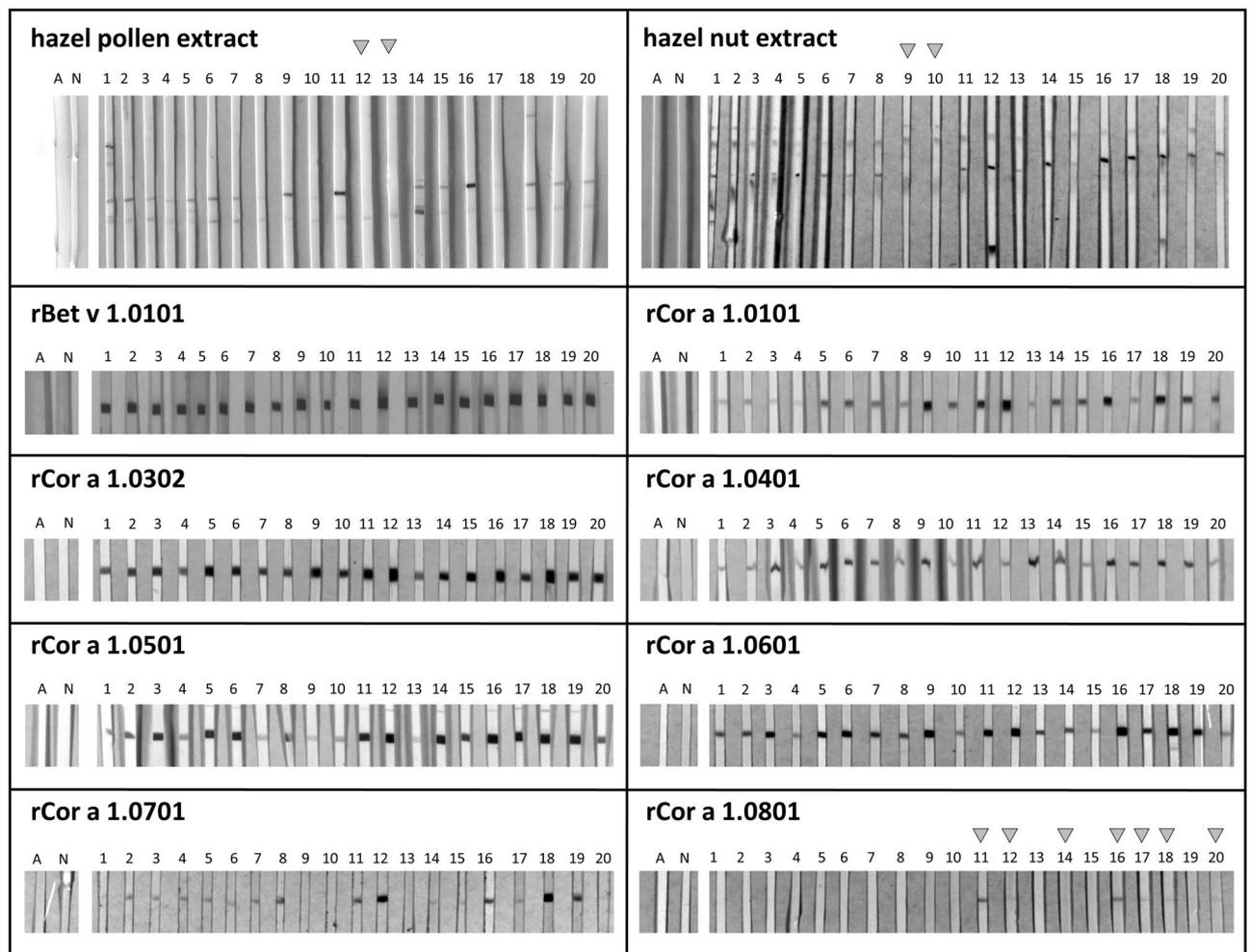
### IgE binding

The rCor a 1 proteins purified from *E. coli* were further used for immunological studies in order to determine their allergenicity. Sera from 20 tree-pollen allergic individuals that had been tested positive for specific IgE (Spez. IgE Kit from Gold Standard Diagnostics, Kassel, Germany) against a mixture of birch, hazel and alder pollen were used for immunoblot analyses with the recombinant isoallergens. Only sera reaching values  $\geq 3.5$  U/ml corresponding to the enzyme-allergo-sorbent-test (EAST) class 3 or higher were selected. As expected, all sera reacted strongly with the Bet v 1 control from birch since this is usually the sensitizing agent<sup>4</sup> (Fig. 3, Figure S5). Furthermore, all sera reacted with rCor a 1.0101, albeit with varying intensities. Similarly, IgE binding could be detected with all sera employing rCor a 1.0302, 1.0401, 1.0501, and 1.0601, indicating that all of them are potent allergens. Using rCor a 1.0701, serum #1, #9, #10, and #13 displayed no IgE binding. rCor a 1.0801 proved to be the weakest allergen and showed IgE binding only with seven sera, namely #11, #12, #14, #16, #17, #18, and #20. Obviously, the sera responded better to those Cor a 1 isoallergens which possess a high identity with the sensitizing Bet v 1.0101 protein (Table 1). The weak allergens Cor a 1.0701 and Cor a 1.0801 have only 54.38 and 46.88% identity with Bet v 1.0101, respectively.

Since we have previously described that denaturation of an allergen during the blotting procedure can impair IgE binding<sup>34</sup>, ELISAs were chosen as an additional method to confirm the allergenicity of the isoallergens (Table 2). Here, all sera reacted with rCor a 1.0101 to rCor a 1.0601. With the exception of serum #12, all sera appeared to be less reactive using Cor a 1.0701. However, in contrast to the Western Blots, rCor a 1.0701 exhibited



**Figure 2.** Physicochemical characterization of Cor a 1 isoallergens. (A) CD spectroscopy. CD spectra at 25 °C (blue curve), after heating to 95 °C (red curve), followed by stepwise cooling to 25 °C (green curve). (B) CD melting curves. CD denaturation (red) and renaturation (black) curves were recorded at 217 or 220 nm in 10 mM Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub> pH 7.0. The  $T_m$  value is displayed on top of each spectrum. No  $T_m$  could be calculated for Cor a 1.0801. (C) <sup>1</sup>H-NMR-spectroscopy. Spectra of Cor a 1.0302 (160 μM), Cor a 1.0501 (200 μM), Cor a 1.0601 (60 μM), Cor a 1.0701 (190 μM) and Cor a 1.0801 (70 μM) were recorded in a Bruker Avance 600 MHz spectrometer at 295 K in 10% (v/v) D<sub>2</sub>O, 10 mM Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub>, pH7.0.



**Figure 3.** Immunoblots. SDS-PAGE (16% PA gels) was performed with the Cor a 1 proteins, followed by blotting onto a 0.2 µm nitrocellulose membrane as indicated on top of the immunoblots. 1.5 µg/cm protein was blotted except for Cor a 1.0801 with 0.8 µg/cm. The membranes were cut into stripes and incubated with patients' sera. Bound specific IgE antibodies were detected using a mouse-anti-human IgE antibody coupled to alkaline phosphatase followed by nitroblue tetrazoliumchloride (NBT)/5-Bromo-4-chloro-3-indolylphosphate (BCIP) staining. The grey arrows indicate weak IgE binding for nCor a 1 from hazel pollen and hazel nut extract and for rCor a 1.0801. Labels on top of the immunoblots: A: secondary antibody control; N: horse serum; 1–20: sera of patients allergic against tree pollen.

IgE binding with all sera, indicating that denaturation of the antigen in the Blots could play a role. Furthermore, the ELISA might be more sensitive than the Blot. Interestingly serum #12 was highly reactive with all isoallergens other than Cor a 1.0801. In addition to the sera reacting with Cor a 1.0801 in the above Western Blot, Cor a 1.0801 also showed weak IgE binding using serum #19. These results confirm that Cor a 1.0801 is a weaker allergen than the others but nevertheless several individuals appear to possess reactive IgE antibodies.

### Concluding remarks

Our data show that there exist more isoallergens and variants of Cor a 1 than previously known, e.g., four potential variants of the newly identified isoallergen Cor a 1.05 as well as three additional variants of Cor a 1.03 could be detected (Figs. 4, S1, S3). We have already shown that Cor a 1.0401 binds to the specific natural ligand Q3O-(Glc)-Gal<sup>17</sup>. To obtain further information on ligand binding, we are currently conducting experiments to show whether the newly identified isoallergens have different ligand binding affinities.

Since little was previously known about the expression of Cor a 1 in different plant tissues the results presented here (Fig. 1, Table 1) show for the first time, that the expression of the isoallergens studied is dependent on the plant tissue, e.g. Cor a 1.01 could be found in all tissues, whereas Cor a 1.08 was only present in the female tissues, albeit at different mRNA expression levels. Obviously, the localization of a certain isoallergen has an impact on its expression level (Fig. 1). These differences suggest different functions of the isoallergens, e.g., during plant development, fertilization and germination, that remain to be elucidated in the future.

The mRNA levels of Cor a 1.01 and Cor a 1.03 are high in pollen (Fig. 1) and thus the proteins appear to play an important role as triggers of pollen allergy. In contrast, the other isoallergens, present in mature nuts, might be relevant for allergic reactions caused after intake of hazel nut containing foodstuff (Fig. 1).



isoallergen from carrot exhibits a pronounced heat stability which can be reduced but not completely abolished after heating of the food matrix<sup>35,37</sup>.

The immunoblots and ELISAs suggest that different Cor a 1 isoallergens are recognized by different pollen allergic patients. It depends on the individual immune system against which isoallergen the IgE antibodies are directed. Consequently, pollen allergic patients need to test whether it is safe to consume hazel nuts that have been baked, roasted or even untreated.

## Methods

### Identification of Cor a 1 genes in the hazel genome

Identification of Cor a 1 homologous genes was performed via BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) using the known isoallergen sequences of Cor a 1 on the nucleotide and protein level. At least four unknown potential isoallergens were identified in the genome of *Corylus avellana* cv Tombul (GenBank assembly accession number: GCA\_901000735.2)<sup>32</sup>, with the following GenBank accession numbers: OQ450370 (Cor a 1.0302), OQ230635 (Cor a 1.0501), OQ450371 (Cor a 1.0601), OQ230636 (Cor a 1.0701) and OQ230637 (Cor a 1.0801). The program Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo>) was used to create multiple sequence alignments and to calculate sequence identities.

### Primer design

Based on the *C. avellana* genome, primers were designed derived from the 5' and 3' untranslated regions or the 5' and 3' ends of the genes. Primers used are listed in Table S1.

### Expression analyses by mRNA isolation from different plant tissues and cDNA production

Immature and mature hazelnuts, pollen, hazel catkins, and female flowers were collected from hazel trees (*C. avellana*) on the Campus of the University of Bayreuth, in accordance with the Ethics Committee of the University of Bayreuth and with relevant institutional, national, and international guidelines and legislation. Female flowers were harvested early in January 2021 before the onset of the pollen flight. Pollen was shaken off from mature blooming catkins and cleaned through a fine metal sieve. The different plant materials were stored at  $-80^{\circ}\text{C}$ . As previously described<sup>34</sup>, total RNA was isolated from  $\sim 200$  mg plant material using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany), followed by reverse transcription with a poly(T) primer to obtain cDNA from mRNA using the SuperScript IV First-Strand Synthesis System with ezDNase enzyme (Thermo Fisher Scientific, Germany) and the designed primer pairs (Table S1). 18S rRNA and the specific PCR products thereof (Table S1) were obtained similarly and used as a control to determine the efficiency of RNA isolation.

### Cloning procedures, sequencing and plasmid preparation

PCR amplicates using Cor a 1 specific primers (Table S1) were loaded onto 1.5% agarose gels to visualize gene expression and purified via gel extraction (Qiaquick Gel Extraction Kit, Qiagen, Germany). In case no band could be detected on the gels, slices were cut out from the gel at the corresponding heights and also purified via gel extraction. 5  $\mu\text{l}$  of the extracted sample were used as a template for a second PCR using the same conditions as described above for the first PCR. To determine whether the expected Cor a 1 gene was present in the PCR amplicate, the PCR products were either sequenced directly using the amplification primers and/or they were cloned into the pCR-Blunt vector using the Zero Blunt PCR Cloning Kit (Thermo Fisher Scientific Germany) as previously described<sup>34</sup>. Sequences were analyzed by Sanger sequencing of the isolated plasmids or by colony sequencing (Microsynth AG, Lindau, Germany). Four to ten clones were subjected to sequencing for each Cor a 1 gene detection.

The genes of Cor a 1.0501, 1.0601, 1.0701 and 1.0801 identified by sequencing of the pCRBlunt plasmids were cloned into the bacterial expression vector pET-GB1a (G. Stier, EMBL) via Gibson cloning<sup>38</sup> using the primers shown in Table S2. Thus, the proteins contain an N-terminal 6His-GB1a fusion which can be cleaved off by tobacco etch virus (TEV) protease. All cleaved off Cor a 1 proteins start N-terminally with the amino acid sequence GMGV. The Cor a 1.0302 gene was ordered in the pET11a expression vector (Eurofins Genomic, Ebersberg, Germany), resulting in a tagless protein starting with a methionine.

### LC-MS<sup>E</sup> analysis

LC-MS<sup>E</sup> analysis of pollen extracts was performed applying published protocols<sup>17,34,35</sup>. Differing from this a Uniprot database was used for MS database search. The database consisted of reviewed entries of all species as of 2022.09 and amino acid sequences translated from nucleic acid sequences mentioned above. To obtain hazelnut extract for LC-MS<sup>E</sup> analysis, procedures were similar to those previously described<sup>24</sup>. The LC-MS<sup>E</sup> analysis was performed as described above for pollen extracts. In brief, mature hazelnuts (Seeberger GmbH, Ulm, Germany) from the supermarket were crushed in a blender, ground in liquid nitrogen, and stirred overnight at  $4^{\circ}\text{C}$  in 10 mM  $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$  pH 7.0, 2% (w/v) polyvinylpyrrolidone, 2 mM EDTA, and 2 mM sodium azide. After centrifugation (20 min, 6000 g,  $4^{\circ}\text{C}$ ), the supernatant was passed through a paper filter (Rotilabo-folded filters, type 113P, Carl Roth GmbH, Karlsruhe, Germany). After dialysis against 5 L 10 mM  $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$  pH 7.0 overnight, the protein sample was separated on a 19% SDS-PA gel. A protein band of the size of ca. 18 kDa corresponding to nCor a 1 was cut out and prepared for LC-MS<sup>E</sup> as described<sup>34</sup>.

### Protein purification of Cor a 1 proteins recombinantly expressed in *E. coli*

Purification of rCor a 1.0501, 0601, 0701 and 0801 was performed as previously described<sup>37</sup> omitting the size exclusion chromatography. Elution of the fusion proteins was performed on a 5 ml HisTrap HP column (Cytiva

Europe GmbH, Freiburg, Germany) by a stepwise increase (15%, 25%, 50%, 70%, 100%) of the imidazole concentration.

pET11a\_Cor a 1.0302 was expressed in BL21(DE3). Bacteria were grown in lysogeny broth (LB) supplemented with 100 µg/ml ampicillin at 37 °C. At an optical density at 600 nm of 0.6–0.7, the temperature was decreased to 18 °C and overexpression was induced overnight for ca. 18 h with 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). After harvesting (6000 × g, 10 min, 4 °C), the cell pellet was resuspended in 20 mM TrisHCl pH 8.0, supplemented with DNase I and EDTA-free protease inhibitor and lysed as previously described<sup>37</sup>. The supernatant was purified by ion exchange chromatography using a 5 ml HiTrap Q XL column (Cytiva Europe GmbH) in 20 mM TrisHCl pH 8.0. rCor a 1.0302 was eluted using a stepwise increase (15%, 25%, 50%, 75%, 100%) of the elution buffer (20 mM TrisHCl, 500 mM NaCl pH 7.0). Fractions containing the Cor a 1.0302 protein were pooled. After the addition of 1 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> the protein was further purified by hydrophobic interaction chromatography with a 5 ml HiTrap Octyl FF column (Cytiva Europe GmbH). After washing with 10 mM Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub> buffer containing 1 M ammonium sulfate, Cor a 1.0302 was eluted with 10 mM Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub>.

### CD und NMR spectroscopy

Circular dichroism (CD) spectra and standard 1D <sup>1</sup>H NMR spectra were recorded as previously described<sup>37</sup>.

### Patients' Sera

Sera were obtained from patients after informed consent and approval by the Ethics Committee of the University of Bayreuth. All experiments were performed in accordance with relevant guidelines and regulations. Sera from patients diagnosed with seasonal (early spring) rhinitis, cough and breathing difficulties were tested for specific IgE against a tree pollen mixture including birch, hazel and alder pollen (Spez. IgE Kit from Gold Standard Diagnostics, Kassel, Germany). Only sera reaching values ≥ 3.5 U/ml corresponding to the enzyme-allergo-sorbent-test (EAST) class 3 or higher were selected and tested further against Bet v 1.0101 (used as a control), as well as hazel pollen extract and an extract from mature hazel nuts to confirm the presence of specific IgE antibodies against Cor a 1 proteins. Those sera were used in this study to detect new Cor a 1 isoallergens.

### IgE immunoblots and IgE ELISAs

IgE immunoblots and ELISAs were performed similarly to previously described procedures<sup>37,39</sup>. Mouse anti-human IgE coupled to alkaline phosphatase (BD Bioscience, Heidelberg, Germany) and AP Conjugate Substrate Kit (Bio-Rad, Germany) was used for immunoblots. Horse serum was used as a negative control and a commercially available mixture of sera from birch pollen and hazel pollen allergic patients (Gold Standard Diagnostics Kassel GmbH) as a positive control.

For ELISAs 2 µg/ml antigen (Cor a 1 proteins or Bet v 1.0101) was coated on high binding ELISA plates F (Sarstedt, Nümbrecht, Germany) in coating buffer (50 mM Na<sub>2</sub>CO<sub>3</sub>/NaHCO<sub>3</sub> pH 9.6) overnight at 4 °C. After washing three times with PBS-T (phosphate buffered saline, 0.05% Tween 20) and blocking with blocking buffer (PBS-T containing 2% BSA) for 1 h at room temperature (RT), sera (diluted 1:10 with blocking buffer) were added and incubated overnight at RT. After washing four times with PBS-T, bound specific IgE was detected with horse radish peroxidase (HRP) labeled mouse-anti-human IgE-Fc-HRP (SouthernBiotech, Birmingham, USA) (diluted 1:5000 with blocking buffer). After 2 h at RT and washing with PBS-T, 100 µl substrate solution of the KPL TMB Microwell Peroxidase Substrate System (2-C) (Sera care Life Science, MA, USA) was added. The color reaction was stopped with 25% sulfuric acid (50 µl/well) and absorbance was read at 450 nm and 620 nm as a reference.

### Data availability

All data generated or analysed during this study are included in this published article (and its Supplementary Information files). *Accession numbers of proteins*: Cor a 1.0302: *Acc. no. OQ450370*; Cor a 1.0501: *Acc. no. OQ230635*; Cor a 1.0601: *Acc. no. OQ450371*; Cor a 1.0701: *Acc. no. OQ230636*; Cor a 1.0801: *Acc. no. OQ230637*.

Received: 16 November 2023; Accepted: 28 February 2024

Published online: 07 March 2024

### References

- White Book on Allergy: Update 2013 (2013).
- East, R. Allergies. *Nature* **479**, S1–S1 (2011).
- Breiteneder, H. & Kraft, D. The history and science of the major birch pollen allergen Bet v 1. *Biomolecules* **13**, 1151 (2023).
- Biedermann, T. *et al.* Birch pollen allergy in Europe. *Allergy* **74**, 1237–1248 (2019).
- Breiteneder, H. *et al.* The gene coding for the major birch pollen allergen Betv1, is highly homologous to a pea disease resistance response gene. *EMBO J.* **8**, 1935–1938 (1989).
- Fernandes, H., Michalska, K., Sikorski, M. & Jaskolski, M. Structural and functional aspects of PR-10 proteins. *FEBS J.* **280**, 1169–1199 (2013).
- Antoniw, J. F., Ritter, C. E., Pierpoint, W. S. & Van Loon, L. C. Comparison of three pathogenesis-related proteins from plants of two cultivars of tobacco infected with TMV. *J. Gen. Virol.* **47**, 79–87 (1980).
- Breiteneder, H. & Ebner, C. Molecular and biochemical classification of plant-derived food allergens. *J. Allergy Clin. Immunol.* **106**, 27–36 (2000).
- Breiteneder, H. & Clare Mills, E. N. Plant food allergens—structural and functional aspects of allergenicity. *Biotechnol. Adv.* **23**, 395–399 (2005).
- Calamelli, E., Trozzo, A., Di Blasi, E., Serra, L. & Bottau, P. Hazelnut allergy. *Medicina (Kaunas)* **57**, 67 (2021).
- Lin, R. Y., Clauss, A. E. & Bennett, E. S. Hypersensitivity to common tree pollens in New York City patients. *Allergy Asthma Proc.* **23**, 253–258 (2002).
- Popescu, F.-D. Cross-reactivity between aeroallergens and food allergens. *World J. Methodol.* **5**, 31–50 (2015).

13. Brazhnikov, G. *et al.* Natural human Bet v 1-specific IgG antibodies recognize non-conformational epitopes whereas IgE reacts with conformational epitopes. *Allergy* <https://doi.org/10.1111/all.15865> (2023).
14. Gajhede, M. *et al.* X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy. *Nat. Struct. Biol.* **3**, 1040–1045 (1996).
15. Marković-Housley, Z. *et al.* Crystal structure of a hypoallergenic isoform of the major birch pollen allergen Bet v 1 and its likely biological function as a plant steroid carrier. *J. Mol. Biol.* **325**, 123–133 (2003).
16. Seutter von Loetzen, C., Schweimer, K., Schwab, W., Rösch, P. & Hartl-Spiegelhauer, O. Solution structure of the strawberry allergen Fra a 1. *Biosci. Rep.* **32**, 567–575 (2012).
17. Jacob, T. *et al.* Identification of a natural ligand of the hazel allergen Cor a 1. *Sci. Rep.* **9**, 8714 (2019).
18. Führer, S. *et al.* Inverse relation between structural flexibility and IgE reactivity of Cor a 1 hazelnut allergens. *Sci. Rep.* **11**, 4173 (2021).
19. 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 (2002).
20. Hurlburt, B. K. *et al.* Structure and function of the peanut panallergen Ara h 8. *J. Biol. Chem.* **288**, 36890–36901 (2013).
21. 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 (2012).
22. Seutter von Loetzen, C. *et al.* Ligand recognition of the major birch pollen allergen Bet v 1 is isoform dependent. *PLoS ONE* **10**, e0128677 (2015).
23. Chruszcz, M. *et al.* Allergens and their associated small molecule ligands-their dual role in sensitization. *Allergy* **76**, 2367–2382 (2021).
24. 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 (1992).
25. Eriksson, N. E., Formgren, H. & Svenonius, E. Food hypersensitivity in patients with pollen allergy. *Allergy* **37**, 437–443 (1982).
26. Ortolani, C., Spano, M., Pastorello, E., Bigi, A. & Ansaloni, R. The oral allergy syndrome. *Ann. Allergy* **61**, 47–52 (1988).
27. de Groot, H., de Jong, N. W., Vuijk, M. H. & Gerth van Wijk, R. Birch pollinosis and atopy caused by apple, peach, and hazelnut; comparison of three extraction procedures with two apple strains. *Allergy* **51**, 712–718 (1996).
28. Ortolani, C. *et al.* Hazelnut allergy: A double-blind, placebo-controlled food challenge multicenter study. *J. Allergy Clin. Immunol.* **105**, 577–581 (2000).
29. Breiteneder, H. *et al.* Four recombinant isoforms of Cor a 1, the major allergen of hazel pollen, show different IgE-binding properties. *Eur. J. Biochem.* **212**, 355–362 (1993).
30. 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 (2002).
31. 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 (1997).
32. Lucas, S. J., Kahraman, K., Avşar, B., Buggs, R. J. A. & Bilge, I. A chromosome-scale genome assembly of European hazel (*Corylus avellana* L.) reveals targets for crop improvement. *Plant J.* **105**, 1413–1430 (2021).
33. Nakabayashi, M. *et al.* Structure of a conserved hypothetical protein, TTHA0849 from *Thermus thermophilus* HB8, at 2.4 Å resolution: A putative member of the StAR-related lipid-transfer (START) domain superfamily. *Acta Crystallogr. F Struct. Biol. Cryst. Commun.* **61**, 1027–1031 (2005).
34. Jacob, T. *et al.* A novel isoallergen Dau c 1.0401 in carrot: Stability, allergenicity, and comparison with other isoallergens. *Mol. Nutr. Food Res.* **65**, e2001110 (2021).
35. Jacob, T. *et al.* Food processing does not abolish the allergenicity of the carrot allergen Dau c 1: Influence of pH, temperature, and the food matrix. *Mol. Nutr. Food Res.* <https://doi.org/10.1002/mnfr.202000334> (2020).
36. Machado, Y. *et al.* Fold stability during endolysosomal acidification is a key factor for allergenicity and immunogenicity of the major birch pollen allergen. *J. Allergy Clin. Immunol.* **137**, 1525–1534 (2016).
37. Hendrich, J. M. *et al.* Allergenicity and IgE recognition of new Dau c 1 allergens from carrot. *Mol. Nutr. Food Res.* **67**, e2200421 (2023).
38. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
39. Wangorsch, A. *et al.* Identification of a Dau c PRPlike protein (Dau c 1.03) as a new allergenic isoform in carrots (cultivar Rodelika). *Clin. Exp. Allergy* **42**, 156–166 (2012).

## Acknowledgements

We thank Victoria Ogungbemi-Alt, Ulrike Persau, Andrea Hager, Verena Millitzer, and Anna-Maria Engin for excellent technical assistance, and Dr. Kristian Schweimer for technical support recording NMR spectra. We thank Dr. med. Angelika Reich (lung specialist and allergologist, Bayreuth) and Christine Seubert for collecting and initial characterization of sera from tree pollen allergic patients.

## Author contributions

B.M.W. supervised and coordinated the project. A.R. designed, analysed, and interpreted the mass spectrometry data and wrote the corresponding part of the manuscript. B.M.W. and J.M.H. designed all other experiments and evaluated the data. T.P.J. designed initial experiments, purified Cor a 1.0101 and 1.0401 and provided conceptual input. S.A. purified Cor a 1.0302 and 1.0801. H.K. cloned and purified Cor a 1.0701. K.R. purified Cor a 1.0601 and performed CD experiments with the protein. J.M.H. performed RNA extractions, PCR reactions, CD experiments, immunoblots and ELISAs. B.M.W. and J.M.H. wrote the paper. All authors contributed to, read and approved the final version of the manuscript.

## Funding

Open Access funding enabled and organized by Projekt DEAL. This work was supported by the Deutsche Forschungsgemeinschaft (DFG), project number 449125263, the Open Access Publishing Fund of the University of Bayreuth, and the Servicestelle Chancengleichheit, University of Bayreuth.

## 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/s41598-024-55856-2>.

**Correspondence** and requests for materials should be addressed to B.M.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://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) 2024