

Gastrointestinal digestion of hazelnut allergens on molecular level: Elucidation of degradation kinetics and resistant immunoactive peptides using mass spectrometry

Robin Korte^{1*}, Julia Bräcker^{2*}, Jens Brockmeyer²

¹Institute of Food Chemistry, University of Münster, Münster, Germany

²Analytical Food Chemistry, University of Stuttgart, Stuttgart, Germany

* These authors contributed equally to this work

Correspondence: Jens Brockmeyer (jens.brockmeyer@lc.uni-stuttgart.de).

Keywords: Allergen / Digestion / Food allergy / Hazelnut / Mass spectrometry

Abbreviations: aa, amino acid; FA, formic acid; FDR, false discovery rate; GI, gastrointestinal; IgE, immunoglobulin E; LTP, lipid transfer protein; MW, molecular weight; PTM, post-translational-modification

Received: 07/02/2017; Revised: 06/05/2017; Accepted: 01/06/2017

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/mnfr.201700130](#).

This article is protected by copyright. All rights reserved.

Abstract

Scope: Allergy to hazelnut seeds ranks among the most prevalent food allergies in Europe. The aim of this study was to elucidate the gastrointestinal digestion of hazelnut allergens on molecular level.

Methods and results: Hazelnut flour was digested *in vitro* following the Infogest consensus model. For six allergenic proteins, the time-dependent course of digestion was monitored by SDS-PAGE and HPLC-MS/MS, and degradation products were characterized by a bottom-up proteomics approach. Depending on the molecular structure, a specific biochemical fate was observed for each allergen, and degradation kinetics were traced back to the peptide level. 1183 peptides were characterized, including 130 peptides that carry known IgE-binding epitopes and may represent sensitizers for hazelnut allergy. The kinetics of peptide formation and degradation were determined by label free quantification and follow a complex multi-stage mechanism.

Conclusion: We present a comprehensive survey on the gastrointestinal digestion of a relevant allergenic food on level of the peptidome, including the first systematic characterization and quantification of degradation products. This provides information on the differential resistance of plant food allergens and their structural elements undergoing digestion and forms the basis for a deeper understanding of the molecular principles responsible for sensitization to food allergy.

1 Introduction

Food allergies are of growing concern in westernized countries, and allergic reactions to food regularly result in severe consequences for allergic consumers [1, 2]. The development of a type I food allergy is preceded by a sensitization process that induces the production of specific IgE against particular structural elements of food proteins. In general, linear and conformational structural motifs can serve as IgE-binding epitopes. For tree nut and peanut allergens, several epitopes of both types have been identified so far [3–8], but the prerequisites that render a food protein as an allergen are still not well understood. .

Resistance of a protein against gastrointestinal (GI) digestion has, among other structural and functional characteristics, been discussed as an important factor for allergenicity. Following the report of a comprehensive study on food allergen digestion *in vitro* by Astwood et al. in 1996 [9], stability under gastric conditions has in particular been regarded as a valid parameter for the identification of allergens [10–12], and *in vitro* assays for pepsin digestion were included into a 2001 FAO/WHO protocol for the allergenicity assessment of novel food proteins [13]. Since only proteins persisting the gastric transit in intact conformation were accepted to sensitize via the oral route [14], gastric digestion assays have in consequence frequently been utilized as a tool to differentiate between complete and incomplete allergens, the former having the potential to both sensitize and elicit allergic reactions [15]. Nevertheless, reports of potent allergens that were not stable in such assays, and the localization of many IgE-binding epitopes in structural elements of the native protein that are not surface exposed suggest that a partial digestion of food allergens in the stomach does not necessarily eliminates their sensitizing capacities [6, 16, 17]. Furthermore, more recent studies have demonstrated that products from the *in vitro* digestion of peanut allergens are able to induce IgE production in rats [18, 19].

While most previous reports on the GI digestion of allergens focused on the fate of the intact proteins and their IgE binding potential [20–23], some more recent studies also reported the characterization of several peptides as degradation products via mass spectrometry (MS) [24–26]. However, up to now no report exists on the GI fate of relevant food allergens which systematically characterizes the digestion process on the level of the degradation products. We hereby aim to close this gap by presenting a comprehensive investigation on the digestion of hazelnut allergens, including a thorough identification and relative quantification of the derived peptides by HPLC–MS/MS. Different to most previous studies, we used consumption-ready food and report utilization of the newly developed Infogest consensus model for multi-phase *in vitro* digestion, which efficiently mimics the luminal phase (Fig. 1) [27].

Hazelnut was chosen as the allergenic target food, because it covers all important families of plant food allergens: Cor a 1 and Cor a 2 [28, 29], belonging to the Bet v 1 homologue and profilin family, represent incomplete allergens, which elicit allergic responses only after sensitization to tree and grass pollen [11]. More severe symptoms and a sensitization via the GI tract have been associated with the 11S and 7S globulins Cor a 9 and Cor a 11 (cupin

superfamily) [30, 31], the lipid transfer protein (LTP) Cor a 8 and the 2S albumin Cor a 14 (prolamin superfamily) [32, 33], and the oleosins Cor a 12 and Cor a 13 [34]. Furthermore, hazelnut allergy is one of the most prevalent food allergies in Europe [34–36]. With this study, we therefore aim to elucidate the molecular dynamics of this relevant allergenic food during the GI passage. By comparison of the degradation products of relevant allergens with *bona fide* known linear IgE epitopes, we seek to provide a molecular basis for the evaluation of their immunogenic potential and for further investigations on the molecular prerequisites responsible for food allergies.

2 Materials and Methods

2.1 *In vitro* digestion

Hazelnut kernels were purchased unroasted from a local food store and ground with a cutting mill (Grindomix GM 200, Retsch) to simulate chewing. *In vitro* simulated human digestion was then carried out on the basis of the standardized static model proposed by the COST Infogest action [27] with simulated salivary fluid (SSF), simulated gastric fluid (SGF), and simulated intestinal fluid (SIF) prepared according to the authors' instructions. To allow for quick and contamination-free sampling, the sample material was split into aliquots of 500 mg in preparation for the experiment, and all digestion steps were performed under agitation at 37 °C.

To simulate all consecutive stages of human digestion (oral, gastric, and intestinal), each aliquot was first mixed with 1 mL of SSF containing 75 U/mL α -Amylase (α -Amylase from *Bacillus licheniformis*, A3403, Sigma-Aldrich, St Louis, MO, USA), and incubated for 2 min. Then, 1 mL of SGF was added to the oral bolus, together with pepsin (pepsin from porcine gastric mucosa, P6887, Sigma-Aldrich), yielding an activity of 1950 U/mL in the final mixture. The pH was adjusted to 3 by the addition of HCl, and digestion was stopped after 0, 5, 10, 15, 20, 25, 30, 60, 90, and 120 min by raising pH of the sample to 10 with NaOH, followed by sample preparation for SDS-PAGE or HPLC-MS/MS.

For intestinal digestion, the gastric digesta ($t = 120$ min) were mixed with 2 mL of SIF containing bile salts (sodium glycodeoxycholate and sodium taurocholate, each concentrated 5 mmol/L in the final mixture) and pancreatin (pancreatin from porcine pancreas, P7545, Sigma-Aldrich). The concentration of pancreatin was adjusted on the activity of trypsin, determined by TAME assay [38], to yield 100 U/mL in the final mixture, and the pH was adjusted to 7 with NaOH. Digestion was inhibited after 5, 10, 15, 20, 25, 30, 60, 90, and 120 min by raising the pH to 10, which did not stop pancreatin activity entirely, and was thus immediately followed by sample preparation for SDS-PAGE or HPLC-MS/MS analysis. Given the resulting time delay, the intestinal 0 min sample was obtained using previously heat-denatured pancreatin.

2.2 SDS-PAGE analysis

Supernatants of the gastric and intestinal digesta (after centrifugation) were incubated with Laemmli sample buffer, containing dithiothreitol, at 95°C immediately after sampling. SDS-PAGE was run on a Mini-Protean cell, using 4–20% TGX Precast Protein Gels (Bio-Rad, Munich, Germany), followed by staining with Coomassie Brilliant Blue.

Gels were stored in a fridge at 4°C until in-gel tryptic digestion was carried out according to the protocol by Shevchenko et al. [39] with slight modifications. The obtained samples were directly subjected to HPLC-MS/MS analysis.

This article is protected by copyright. All rights reserved.

2.3 Sample preparation for mass spectrometry

Peptides from GI digestion samples were purified by solid-phase extraction (SPE), using Strata-X 33u RP 30 mg/mL columns (Phenomenex, Aschaffenburg, Germany). The cartridges were activated and equilibrated according to the manual. Supernatants of the gastric (200 µL) and intestinal (411 µL, compensating for previous dilution steps) digesta were loaded onto the cartridges, washed with 1 mL of 1% formic acid (FA), and eluted with 1 mL of 90 % MeOH + 1% FA. Eluates were dried under a gentle flow of nitrogen at 40 °C, redissolved in acetonitrile/water (3:97), and stored at 4 °C before being subjected to HPLC–MS/MS analysis within 24 hours.

For reductive analysis of the intestinal 10 min sample, 1 mL of 60 mg/mL dithiothreitol solution was added to the digest and mixed thoroughly, before an increased volume of 503 µL was applied to SPE purification to compensate for this additional dilution step.

2.4 Liquid chromatography–mass spectrometry (HPLC–MS/MS)

Data were acquired on a LTQ Orbitrap XL™ mass spectrometer coupled to an Accela HPLC system (Thermo Fisher Scientific, Dreieich, Germany). LC separation was carried out on an Accucore C18 column (2.1 x 150 mm, 2.6 µm, guard cartridge of the same material, Thermo Fisher Scientific) with an injection volume of 10 µL for the purified GI digesta. For the analysis of in-gel digesta, the complete supernatant of each sample was injected. A gradient of acetonitrile/water + 1% FA was employed at a flow rate of 250 µL/min, ranging from 3–42 % acetonitrile over 100 min for GI digesta, and from 3–45 % over 70 min for the tryptic in-gel digesta. In each case, the mass spectrometer was run in data-dependent positive ion mode. Full scan mass spectra (m/z 300–1500) were acquired in the Orbitrap ($r=30000$), and the two most intense ions were sequentially isolated, fragmented, and analyzed in the linear ion trap by collision-activated dissociation with 30% normalized collision energy and 30 ms activation time. No charge states were rejected from fragmentation, and target ions were dynamically excluded from repeated fragmentation for 45 s. Electrospray ionization conditions were: Vaporizer temperature 350 °C; capillary temperature 225 °C; sheath gas flow 40 (arbitrary units); auxiliary gas flow 20 (arbitrary units); sweep gas flow 10 (arbitrary units); source voltage 3.5 kV and tube lens 135 V.

To ensure reproducibility, HPLC–MS/MS analysis was carried out on three biological replicates, which were prepared by *in vitro* simulated digestion and analyzed on three different days.

2.5 Data analysis for peptide/protein identification and quantification

The raw spectra files were loaded into PEAKS 7.0 (Bioinform, Waterloo, Canada), and searched against a hazelnut protein database (Uniprot, *Corylus avellana*, created 2016-02-26, 394 entries), using the enhanced target-decoy method ("decoy-fusion") for false discovery rate (FDR) estimation and result validation [40, 41]. For the analysis of in-gel digest data, a pig protein database (Uniprot, *Sus scrofa*, created 2016-06-02, 63599 entries) was also employed. Search parameters for peptide identification were: parent mass tolerance 5.0 ppm, fragment mass tolerance 1.0 Da, monoisotopic precursor, unspecific cleavage, methionine oxidation allowed as variable modification. Results were limited to 1% FDR.

Label free quantification of the peptides identified in the GI digesta was performed with PEAKS Q via integration of peptide precursor ions, using automatic retention time alignment with peak identification allowing for a 5.0 ppm mass error tolerance, 6.0 min retention time tolerance, and a 1% FDR threshold. All quantification results considered for this publication were validated by manual inspection.

Of the two Cor a 9 entries available in UniProt (A0A0A0P7E3, Q8W1C2), only the more recent entry (A0A0A0P7E3), which gave more results in every sample following our evaluation approach, was considered to avoid redundancy. The isoforms share a sequence identity of 96.5%.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [42] partner repository with the dataset identifier PXD006412.

3 Results

3.1 Degradation of hazelnut allergens under GI conditions

Degradation of hazelnut allergens over the course of the GI digestion was monitored by SDS-PAGE (Fig. 2). Major protein bands from the gastric 0 min sample (after oral phase prior to the gastric digestion) were excised and analyzed by in-gel tryptic digestion followed by HPLC-MS/MS (Table 1). All six hazelnut allergens commonly accepted to sensitize via the GI tract and associated with severe symptoms (Cor a 8, Cor a 9, Cor a 11, Cor a 12, Cor a 13 and Cor a 14) were identified and assigned to protein bands. Multiple identifications were observed for Cor a 9 and Cor a 11, possibly as a result of their overall high abundance and the occurrence of fragments in planta [43].

Following the progress of the gastric phase, a continuous degradation of all protein bands was observed, except for band 8, which was assigned to the 2S albumin Cor a 14 and the nsLTP Cor a 8. Both allergens thus appear to resist gastric digestion and were confirmed in the intestinal 0 min sample (bands 9 and 10). All other allergens were continuously, but not entirely, degraded by pepsin during 120 min of gastric digestion. These results demonstrate that four out of the six hazelnut allergens are affected by gastric digestion, but significant amounts of all allergens survives the gastric phase in intact form and are transferred to the small intestine. Following the reduction of allergen band intensity, the formation of a new low molecular band (< 3.5 kDa) was observed, most probably attributable to proteolytic peptides.

The intestinal phase yielded a rapid reduction in band intensity for all remaining proteins, with all initial bands being largely removed after 5 min. Only in the MW range of Cor a 8 and Cor a 14, a faint and increasingly blurry band remained. Two new bands occurred during the intestinal digestion: One at 105 kDa, which was degraded again after 20 min and probably results from the aggregation of Cor a 9 derivatives, and one at 26 kDa, which corresponded to the identification of Cor a 11, Cor a 9, as well as pancreatic trypsin and remained observable until the end of the intestinal phase. It is therefore assumed that, at least to some degree, larger fragments of Cor a 11 and Cor a 9 persist intestinal digestion, although most of the protein is rapidly degraded.

3.2 Formation of proteolytic peptides as degradation products

While SDS-PAGE allows to estimate digestion kinetics through analysis of the remaining intact allergens, we also monitored the progress of allergen degradation on the product level, following a bottom-up proteomics approach using HPLC-MS/MS combined with software-based data evaluation. For interpretation of these data, it needs to be considered that peptides smaller than 5 amino acids (aa) are usually not identified by this approach with sufficient specificity.

Peptides from all six allergens previously identified in the SDS-PAGE gels were detected in the GI samples. A total of 1017 different peptides were detected and quantified in the gastric (0–120 min) and 686 peptides in the intestinal samples (5–120 min), respectively. The allergens showed considerable differences in their respective time course of peptide formation and degradation (Fig. 3). Notably, all allergens of one protein superfamily demonstrated a similar behavior: The number of peptides from Cor a 9 and Cor a 11, members of the cupin superfamily, rapidly increased in the early gastric phase, indicating their commencing digestion. Although the composition of this peptide mixture underwent continuous change, the amount of different degradation products soon stabilized around an early maximum, which was reached after 5 min for Cor a 9 and 20 min for Cor a 11. Upon transfer to the intestinal phase, the number of peptides decreased rapidly and stabilized again after 10–30 min, demonstrating the degradation of the gastric peptides to a narrow final product spectrum. The prolamin allergens Cor a 14 and Cor a 8 exhibited a completely different behavior: Both proteins yielded near zero peptides until the end of the gastric phase, confirming their resistance against the gastric conditions. After transfer to the intestinal phase, a rapid increase was observed in the number of detected peptides, which reached a maximum at 10–15 min, followed by a continuous degradation, demonstrating intestinal digestion of these allergens. The lower total number of identified peptides is a result of the significantly shorter protein sequence of the prolamins, which leads to an exponentially larger number of possible cleavage products for the cupin allergens due to the applied search options (notably, the unspecified cleavage sites). Another reason for the difference lies in the restriction of prolamin digestion to certain sequence elements, which is demonstrated in section 3.3. The digestion of the oleosins Cor a 12 and Cor a 13 finally showed another individual characteristic, as we observed two maxima in the number of peptides, one for each phase, which in any case was reached after 15–30 min. These two maxima indicate a differential (two-step) degradation of the oleosins' sequence over the two phases. Besides the six allergens presented in Fig. 3, we also detected single peptides from non-allergenic proteins and from the major-pollen allergen Cor a 1 in most samples. Since these were only single hits that did not show any trend, and because the focus of this article is on allergens that sensitize via the GI tract, these results will not be further discussed here.

Given that the length of a peptide carrying an IgE-binding epitope appears to play a major role concerning its potential for sensitization and IgE binding, we also investigated the proceeding degradation of the allergenic proteins through the length of the proteolytic peptides (Fig. 4). Throughout the gastric phase, only a slight change in peptide length was observed: At every time point 98% of the identified peptides ranged between 5 and 37–39 aa in length, and the average as well as the median length dropped only slightly from 16.1 to 12.7 aa and from 14 to 11 aa, respectively. This indicates, together with the results from Fig. 3, that although digestion of the proteins continued, mainly a set of larger peptides was produced that resisted further degradation. However, after transfer to the intestinal phase this set was rapidly degraded to smaller peptides: The upper 99% limit dropped to 19 aa after 15 min of intestinal digestion, while the average and median peptide length was reduced to 8.3 aa and 7 aa in this time. No

significant further change to peptide length was found at later stages of the intestinal phase. Yet, even after 120 min of intestinal digestion, 25% of the remaining peptides still showed a length of ≥ 8 aa, which is a lower limit discussed for linear IgE-binding epitopes [44]. It must, however, be remarked that only the length of peptides that were identified by the applied search algorithms is captured in Fig. 4, not that of all peptides present in the digesta. From our personal experience, a reliable identification via LC-MS is typically not possible for peptides smaller than 5 aa using the provided instrumentation.

3.3 Sequence assignment, epitopes and kinetics of gastrointestinally released peptides

In addition to the global analysis of GI degradation products described in the previous section, we also assigned each peptide to the protein sequence of the originating allergen, to find out which sequence parts were susceptible and which were resistant to digestion. The results are visualized in reduced complexity in Fig. 5, summing up all peptides detected in the gastric and intestinal (without 0 min) digest samples to total sequence coverage (for a more detailed consideration of the results we refer to Fig. SI-1, SI-2, SI-3, SI-4, and SI-5 in the online version of this article, which display all identified peptides in particular).

Again, distinct analogies were found between the two allergen representatives of each superfamily, corresponding to the trends discussed in the previous sections. Digestion products of the cupin allergens Cor a 9 and Cor a 11 showed a high coverage over almost the complete protein sequence in the gastric as well as the intestinal samples. Also, no obvious differences were observed in the distribution of coverage over protein sequence within the time course of the gastric and the intestinal digest samples respectively. Our standard data processing approach, not considering post-translational-modifications (PTM) except M-oxidation, yielded no coverage for N301 of Cor a 11, which has previously been described to be glycosylated [31], as well as for all cysteine residues (five in Cor a 9, two in Cor a 11), apparently a result of disulfide bond formation or different products of S-oxidation as six of these cysteines were covered in a reductive analysis of the intestinal 10 min sample (represented by green bars). A search including PTM and mutations identified modifications for three of the Cor a 9 cysteins (C114: S-cysteinylolation, C164: S-(2-succinyl)cysteinylolation, C325: glutathione disulfide), and also covered the Cor a 11 N-terminus, which carried a single amino acid mutation (K55E). Both allergens are thus demonstrated to be homogeneously degraded by GI digestion, as the whole protein sequence appears to be susceptible to enzymatic gastric and intestinal cleavage.

Largely different results were again found for the prolamins Cor a 14 and Cor a 8. As was shown above, no digestion products were found until commencement of the intestinal phase, and these intestinal peptides were furthermore limited to a short stretch of the protein sequence, resulting in a multitude of overlapping products (Fig. SI-4). Regarding Cor a 14, the covered sequence stretch was attributed to the "hypervariable region", a solvent-exposed loop element of 2S albumins that features a high variability in amino acid composition and structure between

Accepted Article

plant species and is described to be the most important immunogenic region of 2S albumins [45], and the two adjacent α -helices III and IV¹. In case of Cor a 8, the intestinal peptides resulted from a homologous sequence part, covering a solvent-exposed loop and the neighboring helix IV². Remarkably, no degradation products including one of the eight conserved cysteine residues, which are responsible for the compact barrel-like structure of prolamins [46, 47], were identified in the digest samples, and a second search considering all available PTM for peptide identification yielded no additional hits as well. To elucidate whether this prolamin core structure entirely resisted digestion or just proved undetectable because of remaining disulfide bonds, we analyzed the intestinal 10 min sample after reduction with dithiothreitol and found a far greater product diversity, including the majority of the protein sequence and all cysteine residues (green bars in Fig. 5, all products are displayed in Fig. SI-5). We therefore conclude that Cor a 14 and Cor a 8 are in fact digested to significant degree in the intestinal phase, which correlates to the SDS-PAGE results. The resulting peptides, however, are assumed to remain associated via the skeleton of disulfide bonds, largely retaining the allergens' core structure, and only the solvent-exposed loop region between helices III and IV is released as single peptides. For the oleosins Cor a 12 and Cor a 13, degradation was focused on different parts of the protein sequence dependent on the phase of digestion. In the gastric phase, peptides from the amphipathic C-terminus were predominantly observed (compare Fig. SI-4 for a clear visualization), and data evaluation including PTM also covered an A-acetylated N-terminal peptide. The intestinal phase showed an increased degradation of the hydrophilic N-terminus and particularly of the hydrophobic central domain³. This core domain of oleosins is normally embedded into oil bodies, which have been discussed to play a role in protecting the allergens against proteolysis [34], while the N- and C-termini are exposed to the cytosol [48]. The increased digestion of this domain in the intestinal phase may therefore result from a disruption of the oil bodies by bile salts and lipase, and explain the two-step degradation of oleosins that was already observed in Fig. 3.

For an assessment of allergenicity, not only stability of the allergenic proteins, but also the immunogenic and sensitizing capacity of the digestive peptides reaching the intestinal mucosa need to be taken into account. Therefore, we manually surveyed the sequences of all 1183 peptides that were identified in the intestinal digesta, and compared them to previously described linear epitopes. Cor a 9 epitope sequences were taken from Robotham et al. (determined via linear epitope mapping) [6], Cor a 11 epitopes from Barré et al. (determination through homology modelling to Ara h 1 epitopes) [5], and the Cor a 14 epitope core sequence

¹ Cor a 14 secondary structure was assigned on the basis of an alignment to the peanut 2S albumin Ara h 6, Ara h 6 3D structure and structural elements determined by Lehmann et al [46, 63].

² Cor a 8 secondary structure based on the 3D model from Offermann et al [47].

³ Oleosin domains were assigned on basis of an alignment to sesame oleosine [64].

RGEExxE was adopted from the walnut 2S albumin Jug r 1, for which it was determined by mutational epitope analysis [49] (no comparable linear epitope data were available for Cor a 8, Cor a 12 and Cor a 13). A condensed version of the results, focusing on the most relevant epitopes, is given in Table 2, for all matches the reader is referred to Table SI-1 in the supporting information. 130 matches with an overlap of ≥ 8 aa between intestinal peptides and suggested linear epitopes were found, among these 121 matches to Cor a 9 (the most in-depth studied hazelnut allergen in terms of linear epitopes), five to Cor a 11 and four to Cor a 14. This corresponds to a 72% (18 out of 25) match of the previously described epitopes with peptides found in the intestinal samples, including all three Cor a 9 epitopes shown to exhibit the strongest and broadest binding of patient IgE [6]. In case of Cor a 11, only one out of five homology deduced epitopes matched with intestinal peptides, while in case of Cor a 14 the only linear epitope that can be expected based on current literature was identified in multiple peptides. 61% (79 out of 130) of the epitope matching peptides were found in the 0 min sample of the intestinal phase and, therefore, might get in contact with the intestinal mucosa directly after gastric emptying. Among these were all Cor a 9 deduced peptides that matched the most strongly interacting epitope 15 [6]. While a predominant portion of these peptides (60 out of 79, 76%) was immediately degraded by intestinal proteases, 19 peptides were still detectable after 5 min, and another 51 epitope matching peptides were generated at later stages of the intestinal phase, some of them showing stabilities up to 60 min and above.

As expected and demonstrated in Fig. 4, longer precursor peptides were shortened as a result of proteolytic cleavage. This is exemplified for several peptides identified from the regions of the Cor a 9 epitopes 10 and 11 (Fig. 6A). It is obvious that the formation and degradation of digestive peptides follows a complex multi-stage mechanism, with more than one possible reaction pathway for almost every product (Fig. 6B). Furthermore, it can be seen that while many gastric peptides are quickly degraded in the intestinal phase (C_1 , D_1 , D_2 , E_2), a considerable stability is observed for a number of peptides that harbor IgE epitopes, although carrying internal tryptic (E_1) or chymotryptic (F, G, H) cleavage sites. This differential stability under intestinal condition is, in addition, exemplified in Fig. 7 for a selection of peptides from the most relevant epitopes (compare Table 2).

4 Discussion

4.1 Differential stability of allergen super(families) under gastrointestinal conditions

Taking into account the results from the SDS-PAGE (protein level) as well as the HPLC–MS/MS analysis (peptide level), a connection between the molecular (super)family of hazelnut allergens and their fate in the GI digestion processes can be drawn.

The cupin allergens Cor a 9 and Cor a 11 were comprehensively digested in the gastric as well as the intestinal phase, and the released peptides were distributed over the complete protein sequence. A relevant share of intact protein, however, survived proteolysis by gastric pepsin, and reached the intestinal phase. This is in contrast to the results of Vieths et al., who reported the only comprehensive study on the digestion of hazelnut allergens available so far, and found that all proteins > 14 kDa were entirely degraded after 2 h of gastric digestion, leading to destroyed IgE reactivity with most patients under study [20]. Similar results were reported by Mattison et al. for the digestion of the cupin allergens Ana o 1 and Ana o 2 from cashew [24] and by van Bostel et al. for the legumin allergens Ara h 3 and glycinin from peanut [22]. These studies are, however, not entirely comparable to our data, because they used different model systems to simulate GI digestion, including less complex simulative fluids. Most importantly, all previous studies applied *in vitro* digestion to various protein extracts obtained under differential conditions, whereas crude hazelnut flour was utilized in this study. With the Infogest model [27] now available, we consider its application the international standard. In addition, the use of food instead of protein extracts is far closer to the situation *in vivo*, taking into account interactions of food protein and enzymes with other food compounds, such as polyphenols or lipids.

Different to the cupin allergens, the prolamins Cor a 14 and Cor a 8 demonstrated a high stability in the gastric phase, with almost no degradation of the intact proteins. The respective protein families, 2S albumins and LTPs, are commonly classified as particularly resistant against heat and protease treatment, given their closely packed molecular structure and stabilizing disulfide bonds [46, 47]. A high stability against *in vitro* digestion has been described for 2S albumins and LTPs from several allergenic foods [23, 25, 50, 51]. Concerning the digestion of hazelnut allergens, Vieths et al. also observed a high resistance of the prolamins fraction, with retained IgE binding for proteins in the prolamins molecular size range [20], and Pfeifer et al. recently reported that isolated Cor a 14 withstood simulated gastric and intestinal digestion [52]. Differently to these former studies, Cor a 14 and Cor a 8 were shown to be degraded by intestinal enzymes in our model system, a deviation that might again be explained by our model system, which more closely reflects the physiological situation, and by subjecting the whole food matrix towards digestion (instead of isolated proteins), which has been discussed as an influence on single protein digestibility in another recent study [26]. Interestingly, all peptides identified via HPLC–MS/MS without prior reduction of disulfide bonds were attributable only to a narrow range of the protein sequence close to the C-terminus (Fig. 5), which is considerably

exposed and lacks cysteine residues. It is thus assumed that the prolamin core structure of 4–5 disulfide bridged α -helices remains associated, despite enzymatic proteolysis in the gut, through the conserved skeleton of disulfide bonds. Such intestinal breakdown of the protein core structure into peptides, held together through disulfide bonds, was also reported by Moreno et al. for the intestinal digestion of the 2S albumin Ber e 1 from brazil nut [50].

Only little information on the fate of oleosins during GI digestion is available so far. Makkhun et al. recently reported that oleosins of purified sunflower oil bodies were only partially degraded under gastric conditions [53]. The authors suggested that the hydrophobic central domain was protected from pepsin activity, as a result of its anchoring inside the oil droplets, and showed that incubation with bile salts led to a displacement of oleosin from oil bodies. The effect of bile salts could thus explain the two-step degradation observed for Cor a 12 and Cor a 13 in this study and the increased release of proteolytic peptides from the hydrophobic central domain, which we observed in the intestinal phase.

4.2 Sensitizing capacity and immunological relevance of digested food allergens

Our results suggest that a relevant portion of those hazelnut allergens that sensitize and elicit reactions through contact with the GI mucosa reach the intestinal phase in intact form (Fig. 2). Beyond that, it is demonstrated that, as a consequence of protease digestion, a complex mixture of proteolytic peptides is either transferred to or produced in the intestinal compartment, including a relevant percentage of peptides with a length of 8–20 aa (Fig. 4). Knowledge of an allergen's structure, as it is presented to the effector immune system, is essential for understanding which epitopes are most likely to elicit an allergic reaction [6], with the allergenic potential considerably depending on an IgE epitope's resistance against GI digestion [54]. Although the effective absorption of intact peanut protein through the oral mucosa has also been reported [55], the GI tract represents the biggest immune organ of the human body and the major site of interaction with (and thus sensitization to) food allergens [14]. Multiple routes for the passage of intact food allergens through the epithelial barrier have been proposed, including paracellular diffusion or translocation through enterocytes [11, 56], and several studies detected intact food proteins in the blood circulation up to 3 hours after food consumption [12]. Nevertheless, it is to be expected that the bulk of food allergens undergoes enzymatic cleavage in the gut and passes the intestinal mucosa in the form of peptides [12]. Sensitization to food allergens could, thus, be either induced by a small amount of intact protein (or large protein fragments) or by digestive peptides that carry intact IgE epitopes. It has long been commonly accepted that "complete" food allergens, which have the potential to both sensitize and elicit allergic reactions, are resistant against GI digestion, while allergens susceptible to digestion are nonsensitizing elicitors [9–11]. This argumentation does, however, not explain the direct sensitization capacity of some allergens, e.g. from fruits, vegetables, or seafood, that have been shown to be digestion-labile in assays [14]. Recently, Bøgh et al. demonstrated that digested Ara h 1, which was entirely broken down into peptides < 2 kDa and did not contain any intact protein, induced the production of IgE antibodies in Brown Norway

rats, and thus retained the potential to cause both sensitization and elicitation of an allergic response [18, 19]. Substantial knowledge about the proteolytic peptides reaching the gut, especially those carrying epitopes, will therefore be of interest, concerning a deeper understanding of sensitization to food allergy.

Among the hazelnut allergens discussed in this study, linear IgE-binding epitopes have so far only been systematically determined for Cor a 9. Concerning their localization in a 3D model, the authors showed that a substantial number of epitopes are not exposed in the native hexameric protein as well as its intact monomeric form, and therefore already suggested a significant degree of GI processing prior to immune interaction [6]. We found that 72% of the previously identified Cor a 9 epitopes overlapped to at least 8 aa with proteolytic peptides identified in samples of the intestinal phase. Although peptides as short as 5 aa have also been reported to exhibit IgE-binding capacities [57, 58], the majority of characterized linear epitopes are ≥ 8 aa in length [44]. Notably, this includes all three epitopes that were categorized to exhibit particularly strong and widely distributed IgE binding (epitopes 10, 11, and 15 in Table 2). If such peptides are potential allergy-sensitizers, the resistance of these epitopes against gastric digestion, in several cases also combined with a considerable stability in the intestinal phase, might therefore be connected to their high sensitization rate. Furthermore, it would explain the broad topological distribution of the Cor a 9 derived digestive peptides over the protein sequence, which is in agreement to the distribution of the linear epitopes of Cor a 9 [6] as well as other 11S legumins [4, 54, 59, 60]. The distribution is different for 2S albumins, like Cor a 14, where it is basically accepted that the hypervariable region represents the most important site for IgE recognition [46], although IgE binding by other regions has also been described [61]. While the high immunogenicity of the hypervariable region can on the one hand be explained through its high exposure, protruding on the outer surface of the 2S albumin tertiary structure, its localization also corresponds to the intestinal Cor a 14 peptides identified in this study without previous reduction, which were all limited to the hypervariable and closely adjacent regions. Given that the immunodominant epitope RGEExxE (identified for the walnut allergen Jur r 1 [49] but identical to the Cor a 14 sequence) was part of several of these peptides, this could also support the concept that digestive peptides may promote allergenic sensitization. Significantly fewer matches of epitopes with intestinal peptides were found for Cor a 11. In this case, the available epitope data are, however, based solely on homology modeling to Ara h 1 and were not additionally verified with patient sera [5].

There has been considerable debate about the significance of linear over conformational epitopes in food allergy, with studies supporting either side [59]. It has been suggested that conformational epitopes are more of relevance to aeroallergen-mediated reactions, whereas linear epitopes are the elicitors of most GI-mediated food allergies, mainly because interaction with the immune system takes place only after at least partial digestion of food proteins [44, 62]. Given the fundamentally different GI fate of the six hazelnut allergens investigated in this study, including the assumed stability of prolamin conformation, such considerations might also depend on an allergen's susceptibility towards GI digestion.

This article is protected by copyright. All rights reserved.

Irrespective of this discussion, the MS-based identification of allergen digestion products, as presented in this study, constitutes a powerful tool to promote further understanding of the molecular structures responsible for food allergies. Elucidation of the formation and degradation kinetics of gastric and intestinal peptides (Fig. 6 & 7) enhances the understanding of these complex multi-stage reaction pathways and can yield valuable information as to the fate of food allergens and the relationship between certain epitopes and immunologic processes in the future. While the presented LC-MS/MS workflow provided an in-depth analysis of the peptide fraction of the digest, other methods like, e.g., gel filtration could be utilized to obtain additional information about larger polypeptides or protein fragments, which were observed in the SDS-PAGE analysis (Fig. 2) and could also represent relevant sensitizers or even elicitors of food allergy. Furthermore, a thorough investigation on the capacity of the identified allergen fragments to induce IgE formation *in vivo* would especially be of great interest for a more profound understanding of the molecular origin of food allergies.

The research strategy of this project was devised by Robin Korte and Jens Brockmeyer. Experiments were performed by Robin Korte and Julia Bräcker. Data analysis and interpretation were performed by Robin Korte, Julia Bräcker and Jens Brockmeyer. Robin Korte and Jens Brockmeyer have written the manuscript. All authors gave their approval to the final version.

This work was funded by a Ph. D. scholarship for Robin Korte provided by the German National Academic Foundation (Studienstiftung des Deutschen Volkes).

The authors have declared no conflicts of interest.

5 References

- [1] Sicherer, S.H., Sampson, H.A., Food allergy: epidemiology, pathogenesis, diagnosis, and treatment. *J. Allergy Clin. Immunol.* 2014, 133, 291–307.
- [2] Husain, Z., Schwartz, R.A., Food allergy update: more than a peanut of a problem. *Int. J. Dermatol.* 2013, 52, 286–294.
- [3] Burks, A.W., Shin, D., Cockrell, G., Stanley, J.S., et al., Mapping and mutational analysis of the IgE-binding epitopes on Ara h 1, a legume vicilin protein and a major allergen in peanut hypersensitivity. *Eur. J. Biochem.* 1997, 245, 334–339.
- [4] Wang, F., Robotham, J.M., Teuber, S.S., Sathe, S.K., et al., Ana o 2, a major cashew (*Anacardium occidentale* L.) nut allergen of the legumin family. *Int. Arch. Allergy Immunol.* 2003, 132, 27–39.
- [5] Barre, A., Sordet, C., Culerrier, R., Rancé, F., et al., Vicilin allergens of peanut and tree nuts (walnut, hazelnut and cashew nut) share structurally related IgE-binding epitopes. *Mol. Immunol.* 2008, 45, 1231–1240.
- [6] Robotham, J.M., Hoffman, G.G., Teuber, S.S., Beyer, K., et al., Linear IgE-epitope mapping and comparative structural homology modeling of hazelnut and English walnut 11S globulins. *Mol. Immunol.* 2009, 46, 2975–2984.
- [7] Willison, L.N., Zhang, Q., Su, M., Teuber, S.S., et al., Conformational epitope mapping of Pru du 6, a major allergen from almond nut. *Mol. Immunol.* 2013, 55, 253–263.
- [8] Zhang, Q., Willison, L.N., Tripathi, P., Sathe, S.K., et al., Epitope mapping of a 95 kDa antigen in complex with antibody by solution-phase amide backbone hydrogen/deuterium exchange monitored by Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* 2011, 83, 7129–7136.
- [9] Astwood, J.D., Leach, J.N., Fuchs, R.L., Stability of food allergens to digestion in vitro. *Nat. Biotechnol.* 1996, 14, 1269–1273.
- [10] Bredehorst, R., David, K., What establishes a protein as an allergen? *J. Chromatogr. B* 2001, 756, 33–40.
- [11] Moreno, F.J., Gastrointestinal digestion of food allergens: effect on their allergenicity. *Biomed. Pharmacother.* 2007, 61, 50–60.
- [12] Untersmayr, E., Jensen-Jarolim, E., The role of protein digestibility and antacids on food allergy outcomes. *J. Allergy Clin. Immunol.* 2008, 121, 1301–1310.
- [13] Food and Agriculture Organization of the United Nations (FAO), Evaluation of

- allergenicity of genetically modified foods, report of a joint FAO/WHO expert consultation on allergenicity of foods derived from biotechnology, Rome 2001.
- [14] Untersmayr, E., Jensen-Jarolim, E., Mechanisms of type I food allergy. *Pharmacol. Ther.* 2006, *112*, 787–798.
- [15] Aalberse, R.C., Structural biology of allergens. *J. Allergy Clin. Immunol.* 2000, *106*, 228–238.
- [16] Yagami, T., Haishima, Y., Nakamura, A., Osuna, H., et al., Digestibility of allergens extracted from natural rubber latex and vegetable foods. *J. Allergy Clin. Immunol.* 2000, *106*, 752–762.
- [17] Fu, T.-J., Abbott, U.R., Hatzos, C., Digestibility of food allergens and nonallergenic proteins in simulated gastric fluid and simulated intestinal fluid – a comparative study. *J. Agric. Food Chem.* 2002, *50*, 7154–7160.
- [18] Bøgh, K.L., Kroghsbo, S., Dahl, L., Rigby, N.M., et al., Digested Ara h 1 has sensitizing capacity in Brown Norway rats. *Clin. Exp. Allergy* 2009, *39*, 1611–1621.
- [19] Bøgh, K., Barkholt, V., Rigby, N., Mills, C., et al., Digested Ara h 1 loses sensitizing capacity when separated into fractions. *J. Agric. Food Chem.* 2012, *60*, 2934–2942.
- [20] Vieths, S., Reindl, J., Müller, U., Hoffmann, A., et al., Digestibility of peanut and hazelnut allergens investigated by a simple in vitro procedure. *Eur. Food Res. Technol.* 1999, *209*, 379–388.
- [21] Lang, G., Kagiya, Y., Kitta, K., Multiplex comparison of the digestibility of allergenic and non-allergenic proteins in rice grains by in vitro digestion. *Food Chem.* 2015, *168*, 606–614.
- [22] van Boxtel, E.L., van Den Broek, L.A.M., Koppelman, S.J., Gruppen, H., Legumin allergens from peanuts and soybeans: effects of denaturation and aggregation on allergenicity. *Mol. Nutr. Food Res.* 2008, *52*, 674–682.
- [23] Berecz, B., Mills, E.N.C., Parádi, I., Láng, F., et al., Stability of sunflower 2S albumins and LTP to physiologically relevant in vitro gastrointestinal digestion. *Food Chem.* 2013, *138*, 2374–2381.
- [24] Mattison, C.P., Grimm, C.C., Wasserman, R.L., In vitro digestion of soluble cashew proteins and characterization of surviving IgE-reactive peptides. *Mol. Nutr. Food Res.* 2014, *58*, 884–893.
- [25] Mamone, G., Nitride, C., Picariello, G., Addeo, F., et al., Tracking the fate of pasta (T. durum semolina) immunogenic proteins by in vitro simulated digestion. *J. Agric. Food Chem.*

- 2015, 63, 2660–2667.
- [26] De Angelis, E., Pilolli, R., Bavaro, S.L., Monaci, L., Insight into the gastro-duodenal digestion resistance of soybean proteins and potential implications for residual immunogenicity. *Food Funct.* 2017, 8, 1599–1610.
- [27] Minekus, M., Alming, M., Alvito, P., Ballance, S., et al., A standardised static in vitro digestion method suitable for food – an international consensus. *Food Funct.* 2014, 5, 1113–1124.
- [28] Breiteneder, H., Ferrera, F., Hoffmann-Sommergruber, K., Ebner, C., et al., Four recombinant isoforms of Cor a I, the major allergen of hazel pollen, show different IgE-binding properties. *Eur. J. Biochem.* 1993, 212, 355–362.
- [29] Jimenez-Lopez, J.C., Morales, S., Castro, A.J., Volkmann, D., et al., Characterization of profilin polymorphism in pollen with a focus on multifunctionality. *PLoS One* 2012, 7, e30878.
- [30] Beyer, K., Grishina, G., Bardina, L., Grishin, A., et al., Identification of an 11S globulin as a major hazelnut food allergen in hazelnut-induced systemic reactions. *J. Allergy Clin. Immunol.* 2002, 110, 517–523.
- [31] Lauer, I., Foetisch, K., Kolarich, D., Hazelnut (*Corylus avellana*) vicilin Cor a 11: molecular characterization of a glycoprotein and its allergenic activity. *Biochem. J.* 2004, 383, 327–334.
- [32] Schocker, F., Lüttkopf, D., Scheurer, S., Petersen, A., et al., Recombinant lipid transfer protein Cor a 8 from hazelnut: A new tool for in vitro diagnosis of potentially severe hazelnut allergy. *J. Allergy Clin. Immunol.* 2004, 113, 141–147.
- [33] Garino, C., Zuidmeer, L., Marsh, J., Lovegrove, A., et al., Isolation, cloning, and characterization of the 2S albumin: a new allergen from hazelnut. *Mol. Nutr. Food Res.* 2010, 54, 1257–1265.
- [34] Akkerdaas, J.H., Schocker, F., Vieths, S., Versteeg, S., et al., Cloning of oleosin, a putative new hazelnut allergen, using a hazelnut cDNA library. *Mol. Nutr. Food Res.* 2006, 50, 18–23.
- [35] Datema, M.R., Zuidmeer-Jongejan, L., Asero, R., Barreales, L., et al., Hazelnut allergy across Europe dissected molecularly: a EuroPrevall outpatient clinic survey. *J. Allergy Clin. Immunol.* 2015, 136, 382–391.
- [36] McWilliam, V., Koplin, J., Lodge, C., Tang, M., et al., The prevalence of tree nut allergy: a systematic review. *Curr. Allergy Asthma Rep.* 2015, 15, 54.

- [37] Willison, L.N., Sathe, S.K., Roux, K.H., Production and analysis of recombinant tree nut allergens. *Methods* 2014, 66, 34–43.
- [38] Hummel, B.C.W., A modified spectrophotometric determination of chymotrypsin, trypsin, and thrombin. *Can. J. Biochem. Physiol.* 1959, 37, 1393–1399.
- [39] Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V, et al., In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* 2006, 1, 2856–2860.
- [40] Ma, B., Zhang, K., Hendrie, C., Liang, C., et al., PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* 2003, 17, 2337–2342.
- [41] Zhang, J., Xin, L., Shan, B., Chen, W., et al., PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol. Cell. Proteomics* 2012, 11, 1–8.
- [42] Vizcaíno, J.A., Csordas, A., Del-Toro, N., Dienes, J.A., et al., 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* 2016, 44, D447–D456.
- [43] Downs, M.L., Semic-Jusufagic, A., Simpson, A., Bartra, J., et al., Characterization of low molecular weight allergens from English walnut (*Juglans regia*). *J. Agric. Food Chem.* 2014, 62, 11767–11775.
- [44] Bannon, G.A., Ogawa, T., Evaluation of available IgE-binding epitope data and its utility in bioinformatics. *Mol. Nutr. Food Res.* 2006, 50, 638–644.
- [45] Pantoja-Uceda, D., Bruix, M., Giménez-Gallego, G., Rico, M., et al., Solution structure of RicC3, a 2S albumin storage protein from *Ricinus communis*. *Biochemistry* 2003, 42, 13839–13847.
- [46] Moreno, F.J., Clemente, A., 2S albumin storage proteins: what makes them food allergens? *Open Biochem. J.* 2008, 2, 16–28.
- [47] Offermann, L.R., Bublin, M., Perdue, M.L., Pfeifer, S., et al., Structural and functional characterization of the hazelnut allergen Cor a 8. *J. Agric. Food Chem.* 2015, 63, 9150–9158.
- [48] Tzen, J.T.C., Lie, G.C., Huang, A.H.C., Characterization of the charged components and their topology on the surface of plant seed oil bodies. *J. Biol. Chem.* 1992, 267, 15626–15634.
- [49] Robotham, J.M., Teuber, S.S., Sathe, S.K., Roux, K.H., Linear IgE epitope mapping of the English walnut (*Juglans regia*) major food allergen, Jug r 1. *J. Allergy Clin. Immunol.* 2002, 109, 143–149.

This article is protected by copyright. All rights reserved.

- [50] Moreno, F.J., Mellon, F.A., Wickham, M.S.J., Bottrill, A.R., et al., Stability of the major allergen Brazil nut 2S albumin (Ber e 1) to physiologically relevant in vitro gastrointestinal digestion. *FEBS J.* 2005, 272, 341–352.
- [51] Moreno, F.J., Maldonado, B.M., Wellner, N., Mills, E.N.C., Thermostability and in vitro digestibility of a purified major allergen 2S albumin (Ses i 1) from white sesame seeds (*Sesamum indicum* L.). *Biochim. Biophys. Acta* 2005, 1752, 142–153.
- [52] Pfeifer, S., Bublin, M., Dubiela, P., Hummel, K., et al., Cor a 14, the allergenic 2S albumin from hazelnut, is highly thermostable and resistant to gastrointestinal digestion. *Mol. Nutr. Food Res.* 2015, 59, 2077–2086.
- [53] Makkhun, S., Khosla, A., Foster, T., McClements, D.J., et al., Impact of extraneous proteins on the gastrointestinal fate of sunflower seed (*Helianthus annuus*) oil bodies: a simulated gastrointestinal tract study. *Food Funct.* 2015, 6, 124–133.
- [54] Rougé, P., Culierrier, R., Sabatier, V., Granier, C., et al., Mapping and conformational analysis of IgE-binding epitopic regions on the molecular surface of the major Ara h 3 legumin allergen of peanut (*Arachis hypogaea*). *Mol. Immunol.* 2009, 46, 1067–1075.
- [55] Dirks, C.G., Pedersen, M.H., Platzer, M.H., Bindslev-Jensen, C., et al., Does absorption across the buccal mucosa explain early onset of food-induced allergic systemic reactions? *J. Allergy Clin. Immunol.* 2005, 115, 1321–1323.
- [56] Brandtzaeg, P., Mechanisms of gastrointestinal reactions to food. *Environ. Toxicol. Pharmacol.* 1997, 4, 9–24.
- [57] Banerjee, B., Greenberger, P.A., Fink, J.N., Conformational and linear B-cell epitopes of Asp f 2, a major allergen of *Aspergillus fumigatus*, bind differently to immunoglobulin E antibody in the sera of allergic bronchopulmonary aspergillosis patients. *Infect. Immun.* 1999, 67, 2284–2291.
- [58] Beezhold, D.H., Hickey, V.L., Slater, J.E., Gordon, L., Human IgE-binding epitopes of the latex allergen Hev b 5. *J. Allergy Clin. Immunol.* 1999, 103, 1166–1172.
- [59] Willison, L.N., Tripathi, P., Sharma, G., Teuber, S.S., et al., Cloning, expression and patient IgE reactivity of recombinant Pru du 6, an 11S Globulin from almond. *Int. Arch. Allergy Immunol.* 2011, 156, 267–281.
- [60] Helm, R.M., Cockrell, G., Connaughton, C., Sampson, H.A., et al., A soybean G2 glycinin allergen. *Int. Arch. Allergy Immunol.* 2000, 123, 213–219.
- [61] Otsu, K., Guo, R., Dreskin, S.C., Epitope analysis of Ara h 2 and Ara h 6: characteristic patterns of IgE-binding fingerprints among individuals with similar clinical histories. *Clin. Exp. Allergy* 2014, 45, 471–484.

- [62] Lin, J., Sampson, H.A., The role of immunoglobulin E-binding epitopes in the characterization of food allergy. *Curr. Opin. Allergy Clin. Immunol.* 2009, 9, 357–363.
- [63] Lehmann, K., Schweimer, K., Neudecker, P., Rösch, P., Sequence-specific ¹H, ¹³C and ¹⁵N resonance assignments of Ara h 6, an allergenic 2S albumin from peanut. *J. Biomol. NMR* 2004, 29, 93–94.
- [64] Chen, J.C., Lin, R.H., Huang, H.C., Tzen, J.T., Cloning, expression and isoform classification of a minor oleosin in sesame oil bodies. *J. Biochem.* 1997, 122, 819–824.

Figure legends

Figure 1: Experimental procedure for the *in vitro* digestion and analysis of hazelnut allergens.

Figure 2: SDS-PAGE (4–20%) analysis of *in vitro* digestion of hazelnuts, also showing untreated hazelnut extract (#), pepsin (Pep), and pancreatin (Pan) as a reference. Labeled bands (1–13) were cut and analyzed by in-gel digest for protein identification (Table 1).

Figure 3: Formation and degradation of proteolytic peptides from hazelnut allergens over the course of *in vitro* digestion, expressed as the number of different peptides identified in the digest samples. At each time point, three biological replicates were taken.

Figure 4: Average length of proteolytic peptides identified by the search algorithms over the course of *in vitro* digestion (boxes include the central 50% and whiskers 98% of the data, asterisks indicate the average and bold bars the median value).

Figure 5: Sequence coverage of proteolytic peptides identified in the digest samples by HPLC–MS/MS. Overlapping gastric (blue bars) and intestinal peptides (yellow bars) at 5–120 min are summarized to total coverage to reduce complexity. Only peptides identified in ≥ 2 replicates and without PTM are considered except in particular cases (purple bars) that are discussed in the manuscript. Green bars represent coverage of peptides including cysteine residues observed in the intestinal 10 min sample after reduction. Vertical arrows indicate posttranslational cleavage sites, asterisk indicates N-glycosylation.

Figure 6: Formation and degradation of proteolytic peptides, following digestion of Cor a 9 in the sequence range from N181 to Q214 (non-covered adjacent sequence in grey, epitopes 10 and 11 identified by Robotham et al [6] are underscored). The relative concentration of each peptide was determined as peak area in triplicate at each stage of gastric and intestinal digestion, following a label free quantification approach (A). The data indicate that formation and degradation reactions of the digestive peptides follow multi-stage mechanisms, which can only partially be explained and probably underlie multiple concurring reaction pathways (B).

Figure 7: Formation and degradation kinetics of proteolytic peptides with stability in the intestinal phase. All peptides listed in Table 2 that were detectable at ≥ 5 min of intestinal digestion are shown (except peptides Cor a 9 AA 187–196, 188–195, and 195–211, which are presented in Fig. 6). Non-covered adjacent sequence in grey, epitopes are underscored (for epitope references compare Table 2).

Tables

Table 1: Identification of protein bands from Fig. 1 by in-gel digest.

Band	Samples	MW	Proteins (protein family, accession) ^{a)}	Score (-10lgP)	Coverage	Peptides
1	gas 0 min	48 kDa	Cor a 11 (7S vicilin, Q8S4P9)	326.48	63 %	52
			Cor a 9 (11S legumin, A0A0A0P7E3) ^{b)}	265.40	47 %	18
			Aktin (<i>Sus scrofa</i> , B6VNT8)	129.59	14 %	2
2	gas 0 min	35 kDa	Cor a 9 (7S legumin, A0A0A0P7E3)	293.67	33 %	14
			Cor a 11 (11S vicilin, Q8S4P9)	218.02	20 %	4
3	gas 0 min	28 kDa	Cor a 9 (11S legumin, A0A0A0P7E3)	424.92	47 %	58
			Cor a 11 (7S vicilin, Q8S4P9)	145.99	9 %	2
4	gas 0 min	25 kDa	Cor a 9 (11S legumin, A0A0A0P7E3)	442.61	42 %	30
			Pepsin (<i>Sus scrofa</i> , P00791)	100.46	12 %	3
5	gas 0 min	22 kDa	Cor a 9 (11S legumin, A0A0A0P7E3)	313.76	27 %	14
6	gas 0 min	18 kDa	Cor a 9 (11S legumin, A0A0A0P7E3)	325.41	24 %	11
			Aktin (<i>Sus scrofa</i> , B22FN7)	92.10	10 %	2
7	gas 0 min	15 kDa	Cor a 9 (11S legumin, A0A0A0P7E3)	268.92	30 %	11
			Cor a 13 (oleosin, Q84T91)	125.91	21 %	3
			Cor a 12 (oleosin, Q84T21)	90.40	13 %	2
8	gas 0 min	6–12 kDa	Cor a 14 (2S albumin, D0PWG2)	165.87	17 %	3
			Cor a 8 (nsLTP, Q9ATH2)	161.56	23 %	2
9	int 0 min	10–14 kDa	Cor a 8 (nsLTP, Q9ATH2)	117.86	21 %	2
			Trypsin (<i>Sus scrofa</i> , P00761)	83.63	12 %	2

This article is protected by copyright. All rights reserved.

Band	Samples	MW	Proteins (protein family, accession) ^{a)}	Score (-10lgP)	Coverage	Peptides
10	int 0 min	6–10 kDa	Cor a 14 (2S albumin, D0PWG2)	185.34	17 %	3
11	int 10–30 min	105 kDa	Cor a 9 (11S legumin, A0A0A0P7E3)	286.18	17%	9
12	int 10–120 min	26 kDa	Cor a 11 (7S vicilin, Q8S4P9)	405.46	39 %	25
			Cor a 9 (11S legumin, A0A0A0P7E3)	373.84	44 %	27
			Trypsin (<i>Sus scrofa</i> , P00761)	249.97	52 %	9
13	int 30–90 min	6–14 kDa	Trypsin (<i>Sus scrofa</i> , P00761)	222.56	39 %	8
			α -Amylase (<i>Sus scrofa</i> , I3LLP2)	201.42	21 %	7
			Colipase (<i>Sus scrofa</i> , D5KJH8)	173.88	47 %	3

Notes: ^{a)} only proteins with -10lgP score >50 and >1 matched peptide are listed and sorted by score

^{b)} isoform [Q8W1C2] not listed for Cor a 9 to avoid redundancy

Table 2: Hazelnut allergen derived peptides identified from intestinal digest samples that overlap with known linear epitopes to ≥ 8 aa. Only positive identifications from all three replicates and only Cor a 9 epitopes with the strongest IgE binding (+++) and/or IgE binding over all three patient pools (+) in Robotham et al [6] are presented (for all overlapping peptides, the reader is referred to Table SI-1 in the supporting information). Non-covered adjacent sequence is given in blue, epitopes are underscored, and deviations from the Cor a 9 reference sequence in [6] are indicated in red.

Allergen	AA position	AA sequence	Samples (intestinal)	Epitope reference / IgE binding pattern
Cor a 9	183–194	<u>YANQLDENPRHFYLAGNP</u> DDEHQRQGQQQ	0 min	
	184–192	<u>YANQLDENPRHFYLAGNP</u> DDEHQRQGQQQ	5–60 min	
	184–193	<u>YANQLDENPRHFYLAGNP</u> DDEHQRQGQQQ	5 min	
	185–192	<u>YANQLDENPRHFYLAGNP</u> DDEHQRQGQQQ	5–60 min	Robotham et al (2009) [6]
	187–194	<u>YANQLDENPRHFYLAGNP</u> DDEHQRQGQQQ	0 min	Epitope 10
	187–196	<u>YANQLDENPRHFYLAGNP</u> DDEHQRQGQQQ	0–20 min	(+ / + / +)
	188–211	<u>YANQLDENPRHFYLAGNP</u> DDEHQRQGQQQ	0 min	
	188–195	<u>YANQLDENPRHFYLAGNP</u> DDEHQRQGQQQ	5–30 min	
	188–196	<u>YANQLDENPRHFYLAGNP</u> DDEHQRQGQQQ	0 min	
	188–211	DENPRHFYLAGNP <u>DDEHQRQGQQQ</u> FGQRRRQQQH	0 min	
	195–206	YLAGNP <u>DDEHQRQGQQQ</u> FGQRRRQQQH	5–30 min	Robotham et al (2009) [6]
	195–209	YLAGNP <u>DDEHQRQGQQQ</u> FGQRRRQQQH	0 min	Epitopes 11,12
	195–210	YLAGNP <u>DDEHQRQGQQQ</u> FGQRRRQQQH	0 min	(+++ / + / -)
	195–210	YLAGNP <u>DDEHQRQGQQQ</u> FGQRRRQQQH	0 min	(++ / - / -)
	195–211	YLAGNP <u>DDEHQRQGQQQ</u> FGQRRRQQQH	0–5 min	

This article is protected by copyright. All rights reserved.

Allergen	AA position	AA sequence	Samples (intestinal)	Epitope reference / IgE binding pattern
	195–212	<u>YLAGNPDDDEHQROGQQQFGQRRRQQQH</u>	0–5 min	
	195–213	<u>YLAGNPDDDEHQROGQQQFGQRRRQQQH</u>	0 min	
	195–215	<u>YLAGNPDDDEHQROGQQQFGQRRRQQQH</u>	0 min	
	196–206	<u>LAGNPDDDEHQROGQQQFGQRRRQQQH</u>	5 min	
	196–212	<u>LAGNPDDDEHQROGQQQFGQRRRQQQH</u>	0 min	
	197–206	<u>AGNPDDDEHQROGQQQFGQRRRQQQH</u>	5–15 min	
	197–211	<u>AGNPDDDEHQROGQQQFGQRRRQQQH</u>	0 min	
	197–212	<u>AGNPDDDEHQROGQQQFGQRRRQQQH</u>	0 min	
	197–216	<u>AGNPDDDEHQROGQQQFGQRRRQQQH</u>	0 min	
	200–208	<u>NPDDEHQROGQQQFGQRRRQQQH</u>	5–120 min	
	207–214	<u>NPDDEHQROGQQQFGQRRRQQQH</u>	5 min	
	240–247	<u>FSGFDAEFLADAFNV</u>	5–120 min	Epitope 13 (+++ / +++ / -)
	279–289	<u>QMVRPERSQEWERQE</u>	0 min	
	279–290	<u>QMVRPERSQEWERQE</u>	0 min	Robotham et al (2009) [6]
	279–293	<u>QMVRPERSQEWERQE</u>	0 min	Epitope 15
	279–294	<u>QMVRPERSQEWERQE</u>	0 min	(+++ / +++ / +++)
	280–290	<u>MVRPERSQEWERQE</u>	0 min	

Allergen	AA position	AA sequence	Samples (intestinal)	Epitope reference / IgE binding pattern
Cor a 11	379–398	FEVNAHGNSRFPLAGKGNIVNE	0 min	Barre et al (2008) [5] Epitope 19 (homology match)
	379–400	FEVNAHGNSRFPLAGKGNIVNE	0 min	
	380–400	EVNAHGNSRFPLAGKGNIVNE	0 min	
	382–400	NAHGNSRFPLAGKGNIVNE	0 min	
	387–400	SRFPLAGKGNIVNE	0 min	
Cor a 14	105–125	QAVMQQGEMRGEEMREVMETAR	5–10 min	Robotham et al (2002) [49] (for Jug r 1)
	109–125	QQQGEMRGEEMREVMETAR	5–10 min	
	110–125	QQGEMRGEEMREVMETAR	5 min	
	115–127	RGEEMREVMETAR	5 min	

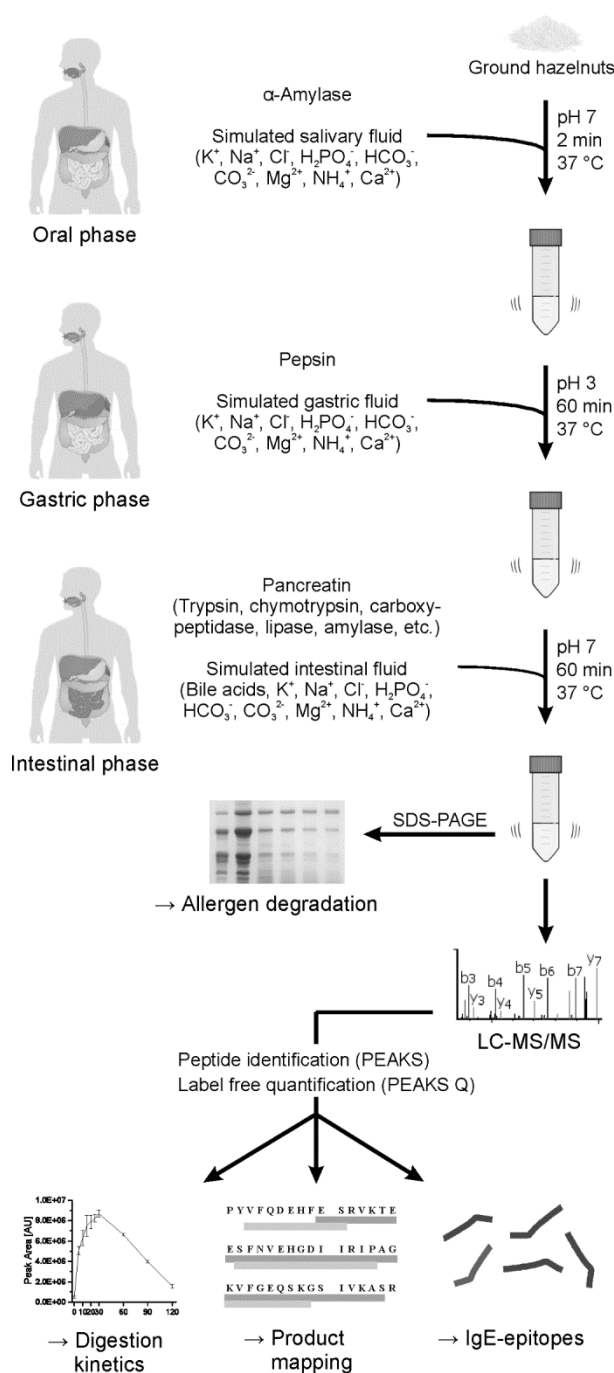


Figure 1: Experimental procedure for the *in vitro* digestion and analysis of hazelnut allergens.

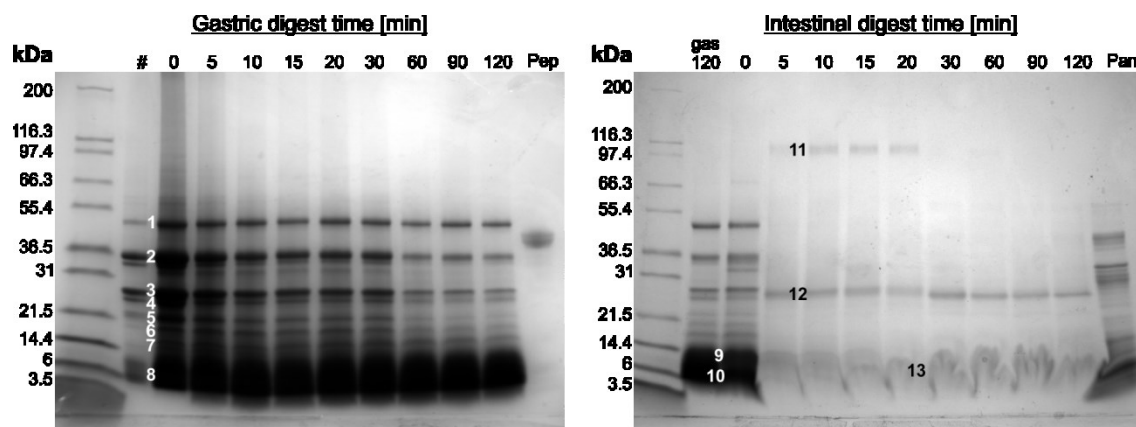


Figure 2: SDS-PAGE (4–20%) analysis of *in vitro* digestion of hazelnuts also showing untreated hazelnut extract (#), pepsin (Pep) and pancreatin (Pan) as a reference. Labeled bands (1-13) were cut and analyzed by in-gel digest for protein identification (Table 1).

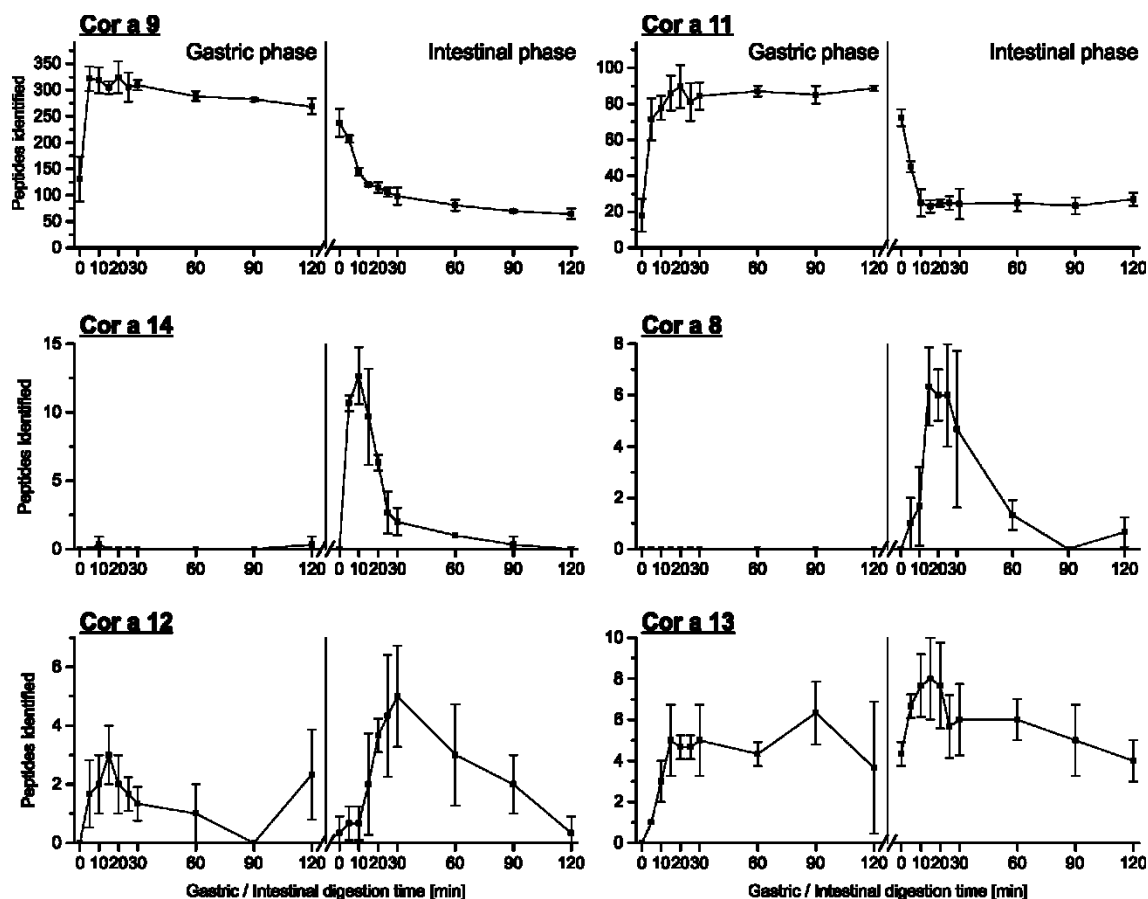


Figure 3: Formation and degradation of proteolytic peptides from hazelnut allergens over the course of *in vitro* digestion, expressed as the number of different peptides identified in the digest samples. At each time point three biological replicates were taken.

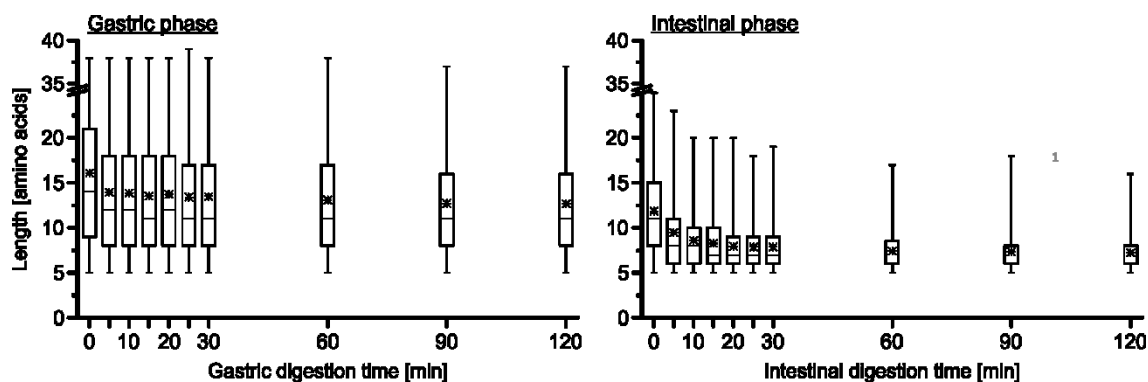


Figure 4: Average length of proteolytic peptides identified by the search algorithms over the course of *in vitro* digestion (boxes include the central 50% and whiskers 98% of the data, asterisks indicate the average and bold bars the median value).

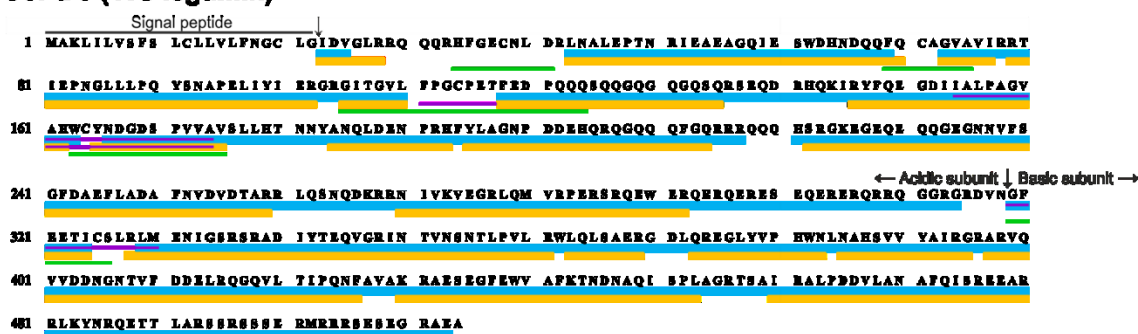
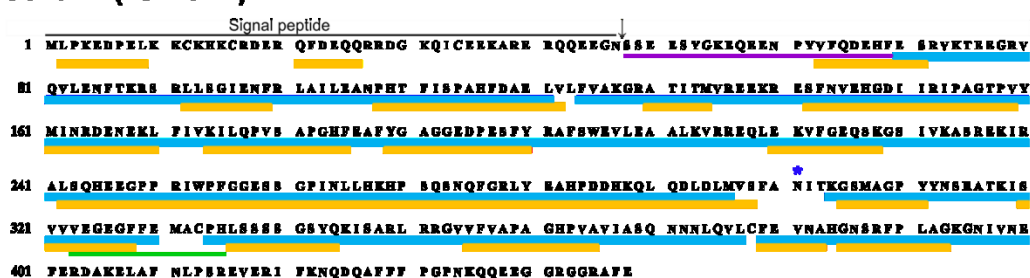
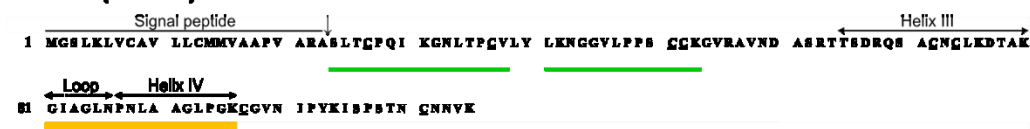
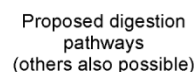
Cor a 9 (11S legumin)**Cor a 11 (7S vicilin)****Cor a 14 (2S albumin)****Cor a 8 (nsLTP)****Cor a 12 (oleosin)****Cor a 13 (oleosin)**

Figure 5: Sequence coverage of proteolytic peptides identified in the digest samples by HPLC–MS/MS. Overlapping gastric (blue bars) and intestinal peptides (yellow bars) at 5–120 min are summarized to total coverage to reduce complexity. Only peptides identified in ≥ 2 replicates and without PTM are considered except in particular cases (purple bars) that are discussed in the manuscript. Green bars represent coverage of peptides including cysteine residues observed in the intestinal 10 min sample after reduction. Vertical arrows indicate posttranslational cleavage sites, asterisk indicates N-glycosylation.



This article is protected by copyright. All rights reserved.

et al [6] are underscored). The relative concentration of each peptide was determined as peak area in triplicate at each stage of gastric and intestinal digestion following a label free quantification approach (A). The data indicate that formation and degradation reactions of the digestive peptides follow multi-stage mechanisms, which can only partially be explained and probably underlie multiple concurring reaction pathways (B).

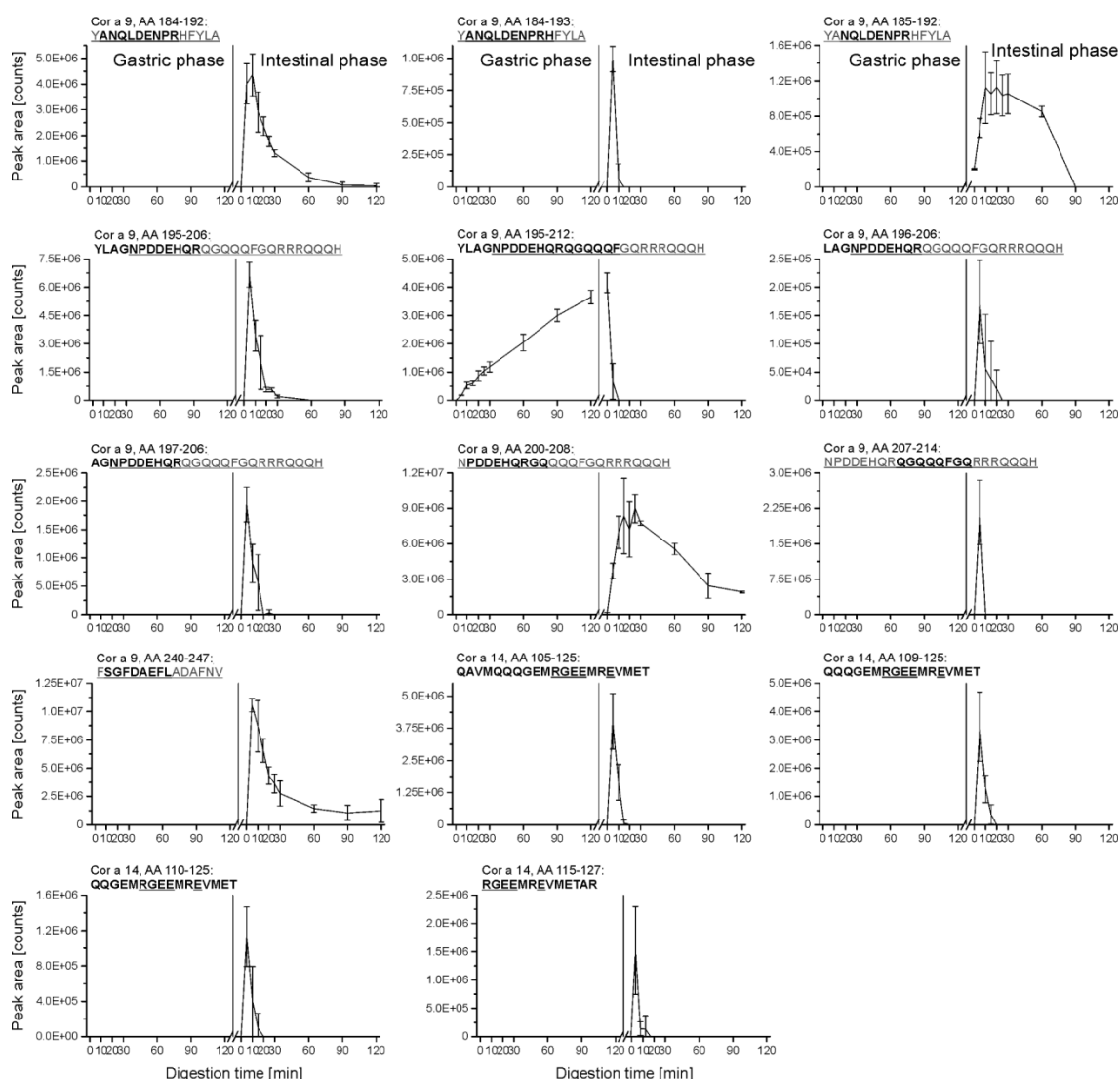
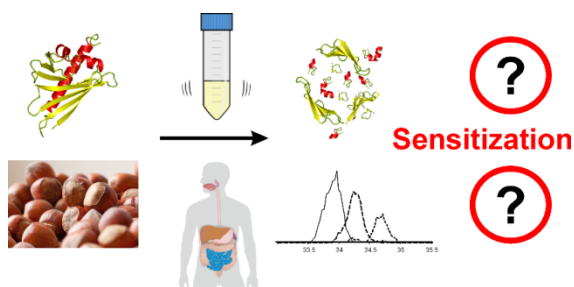


Figure 7: Formation and degradation kinetics of proteolytic peptides with stability in the intestinal phase. All peptides listed in Table 2 that were detectable at ≥ 5 min of intestinal digestion are shown (except peptides Cor a 9 AA 187–196, 188–195, and 195–211, which are presented in Fig. 6). Non-covered adjacent sequence in grey, epitopes are underscored (for epitope references compare Table 2).

Sensitization to hazelnut allergens via the gastrointestinal tract has been associated with severe clinical symptoms. We therefore studied the biochemical fate of six relevant hazelnut allergens undergoing digestion *in vitro*. Product structures and digestion kinetics were elucidated by HPLC–MS/MS. 1183 digestion resistant peptides were characterized that may represent sensitizers or elicitors for hazelnut allergy, and a comprehensive insight into the molecular dynamics of the proteolytic degradation of plant food allergens is provided.



Gastrointestinal digestion of hazelnut allergens on molecular level: Elucidation of degradation kinetics and resistant immunoactive peptides using mass spectrometry

Robin Korte¹, Julia Bräcker^{2*}, Jens Brockmeyer²

¹Institute of Food Chemistry, University of Münster, Münster, Germany

²Analytical Food Chemistry, University of Stuttgart, Stuttgart, Germany

* These authors contributed equally to this work

Correspondence: Jens Brockmeyer (jens.brockmeyer@lc.uni-stuttgart.de).

Keywords: Allergen / Digestion / Food allergy / Hazelnut / Mass spectrometry

Abbreviations: **aa**, amino acid; **FA**, formic acid; **FDR**, false discovery rate; **GI**, gastrointestinal; **IgE**, immunoglobulin E; **LTP**, lipid transfer protein; **MW**, molecular weight; **PTM**, post-translational-modification

Table SI-1: Hazelnut allergen derived peptides identified from intestinal digest samples that overlap with known linear epitopes to ≥ 8 amino acids. Only positive identifications from all three replicates are presented. Non-covered adjacent sequence is given in blue, epitopes are underscored, and deviations from the Cor a 9 reference sequence in [6] are indicated in red. Cor a 9 epitopes were determined by linear epitope mapping [6], Cor a 11 epitopes by homology modelling to Ara h 1 linear epitopes [5], and the Cor a 14 epitope core sequence (RGEExxE) by mutational epitope analysis (to Jug r 1 reactive sera) [49].

Allergen	AA position	AA Sequence	Samples (intestinal)	Reference
Cor a 9	44–68	NALEPTNRIEAEAGQIESWDHNDQQFQC	0–5 min	Robotham et al (2009) [6] Epitope 2 (+ / - / -)
	44–70	NALEPTNRIEAEAGQIESWDHNDQQFQC	5 min	
	50–68	NRIEAEAGQIESWDHNDQQFQC	5 min	
	52–67	IEAEAGQIESWDHNDQQFQC	5 min	
	52–68	IEAEAGQIESWDHNDQQFQC	5–10 min	
	52–70	IEAEAGQIESWDHNDQQFQC	5–30 min	
	53–68	EAEAGQIESWDHNDQQFQC	5–10 min	
	53–70	EAEAGQIESWDHNDQQFQC	5 min	
	54–68	AEAGQIESWDHNDQQFQC	0–10 min	
	54–70	AEAGQIESWDHNDQQFQC	5–10 min	
	55–68	EAGQIESWDHNDQQFQC	0–20 min	
	55–70	EAGQIESWDHNDQQFQC	5–10 min	
	56–68	AGQIESWDHNDQQFQC	0–10 min	
	56–70	AGQIESWDHNDQQFQC	5 min	
	58–68	QIESWDHNDQQFQC	0–5 min	
	59–67	QIESWDHNDQQFQC	5–60 min	
	59–68	QIESWDHNDQQFQC	0–60 min	
	59–70	QIESWDHNDQQFQC	5–60 min	
	61–68	QIESWDHNDQQFQC	0–60 min	
	61–69	QIESWDHNDQQFQC	0–5 min	
	61–70	QIESWDHNDQQFQC	5–30 min	
	62–70	QIESWDHNDQQFQC	5–20 min	
	76–96	VIRRTIEPNGLLLPQYSNAPELIYIERG	0 min	Robotham et al (2009) [6] Epitope 3 (+ / - / -)
	76–97	VIRRTIEPNGLLLPQYSNAPELIYIERG	0 min	
	77–97	IRRTIEPNGLLLPQYSNAPELIYIERG	0 min	
	79–98	RTIEPNGLLLPQYSNAPELIYIERG	5 min	
	80–97	TIEPNGLLLPQYSNAPELIYIERG	5–10 min	
	90–98	TIEPNGLLLPQYSNAPELIYIERG	5–25 min	
	90–102	TIEPNGLLLPQYSNAPELIYIERG	5–10 min	
	85–97	GLLLPQYSNAPELIYIERG	0–5 min	
	85–98	GLLLPQYSNAPELIYIERG	0–10 min	
	87–97	LLPQYSNAPELIYIERG	0–5 min	
	88–97	LPQYSNAPELIYIERG	0 min	
	89–97	PQYSNAPELIYIERG	0–5 min	
	89–98	PQYSNAPELIYIERG	0 min	
	91–98	PQYSNAPELIYIERG	0–30 min	
	99–107	LIYIERGRGITGVLF	0 min	Robotham et al (2009) [6] Epitope 4 (- / + / +)
	99–108	LIYIERGRGITGVLF	0 min	
	99–110	LIYIERGRGITGVLF	0 min	
	100–107	LIYIERGRGITGVLF	0 min	
	100–108	LIYIERGRGITGVLF	0 min	
	100–109	LIYIERGRGITGVLF	0 min	
	100–110	LIYIERGRGITGVLF	0 min	

Allergen	AA position	AA Sequence	Samples (intestinal)	Reference
	100–111	LIYIERGRGITGVLF	0 min	
	101–108	LIYIERGRGITGVLF	0 min	
	118–129	FEDPQQQSQQGQ GQGGQSQRSEQDRHQKIRYFQ	0 min	
	118–131	FEDPQQQSQQGQ GQGGQSQRSEQDRHQKIRYFQ	5 min	Robotham et al (2009) [6]
	118–133	FEDPQQQSQQGQ GQGGQSQRSEQDRHQKIRYFQ	5 min	
	118–135	FEDPQQQSQQGQ GQGGQSQRSEQDRHQKIRYFQ	0–5 min	
	118–139	FEDPQQQSQQGQ GQGGQSQRSEQDRHQKIRYFQ	0 min	
	118–147	FEDPQQQSQQGQ GQGGQSQRSEQDRHQKIRYFQ	0 min	
	121–147	PQQQSQQGQGQGGQSQRSEQDRHQKIRYFQ	0 min	
	128–147	PQQQSQQGQGQGGQSQRSEQDRHQKIRYFQ	0 min	
	146–158	RYFQEGDIIALPAGVAHWCY	5 min	
	147–158	YFQEGDIIALPAGVAHWCY	5–60 min	
	148–158	FQEGDIIALPAGVAHWCY	5–15 min	
	149–158	QEGDIIALPAGVAHWCY	0–60 min	Robotham et al (2009) [6]
	150–158	EGDIIALPAGVAHWCY	5–120 min	
	151–158	GDIIALPAGVAHWCY	5–120 min	Epitope 8 (- / + / -)
	154–163	GDIIALPAGVAHWCY	0 min	
	155–162	GDIIALPAGVAHWCY	5 min	
	155–163	GDIIALPAGVAHWCY	0 min	
	156–163	GDIIALPAGVAHWCY	0 min	
	165–174	YNDGDSPVVAVSLLHTN	5–25 min	Robotham et al (2009) [6]
	166–174	NDGDSPVVAVSLLHTN	5–120 min	
	166–175	NDGDSPVVAVSLLHTN	5–20 min	Epitope 9 (+ / - / -)
	167–175	DGDSPVVAVSLLHTN	10–20 min	
	183–194	YANQLDENPRHFYLA	0 min	
	184–192	YANQLDENPRHFYLA	5–60 min	
	184–193	YANQLDENPRHFYLA	5 min	
	185–192	YANQLDENPRHFYLA	5–60 min	Robotham et al (2009) [6]
	187–194	YANQLDENPRHFYLA	0 min	
	187–196	YANQLDENPRHFYLA	0–20 min	Epitope 10 (+ / + / +)
	188–211	YANQLDENPRHFYLAGNPddehQRQGQQQ	0 min	
	188–195	YANQLDENPRHFYLA	5–30 min	
	188–196	YANQLDENPRHFYLA	0 min	
	188–211	DENPRHFYLAGNPddehQRQGQQQ FGQRRRQQQH	0 min	
	195–206	YLAGNPddehQRQGQQQ FGQRRRQQQH	5–30 min	
	195–209	YLAGNPddehQRQGQQQ FGQRRRQQQH	0 min	
	195–210	YLAGNPddehQRQGQQQ FGQRRRQQQH	0 min	
	195–211	YLAGNPddehQRQGQQQ FGQRRRQQQH	0–5 min	
	195–212	YLAGNPddehQRQGQQQ FGQRRRQQQH	0–5 min	
	195–213	YLAGNPddehQRQGQQQ FGQRRRQQQH	0 min	Robotham et al (2009) [6]
	195–215	YLAGNPddehQRQGQQQ FGQRRRQQQH	0 min	
	196–206	LAGNPddehQRQGQQQ FGQRRRQQQH	5 min	Epitopes 11,12 (+++ / + / -)
	196–212	LAGNPddehQRQGQQQ FGQRRRQQQH	0 min	
	197–206	AGNPddehQRQGQQQ FGQRRRQQQH	5–15 min	
	197–211	AGNPddehQRQGQQQ FGQRRRQQQH	0 min	
	197–212	AGNPddehQRQGQQQ FGQRRRQQQH	0 min	
	197–216	AGNPddehQRQGQQQ FGQRRRQQQH	0 min	
	200–207	NPDDEHQRQGQQQFGQRRRQQQH	5–120 min	
	207–214	NPDDEHQRQGQQQFGQRRRQQQH	5 min	
	240 - 247	FSGFDAEFLADAFNV	5–120 min	Epitope 13

Allergen	AA position	AA Sequence	Samples (intestinal)	Reference
				(+++ / +++ / -)
	258–278	ARRLQSNQDKRRNIVKVEGRL	0 min	Robotham et al (2009) [6]
	261–278	LQSNQDKRRNIVKVEGRL	0 min	
	270–278	NIVKVEGRL	0 min	
	279–289	QMVRPERSRQEWERQ	0 min	Robotham et al (2009) [6]
	279–290	QMVRPERSRQEWERQ	0 min	
	279–293	QMVRPERSRQEWERQ	0 min	
	279–294	QMVRPERSRQEWERQE	0 min	
	280–290	MVRPERSRQEWERQ	0 min	
	366–377	VLRWLQLSAERGDQLQREGL	0 min	Epitope 20 (+ / - / -)
	427–438	AVAKRAESEGFEWVAFKTN	0 min	Robotham et al (2009) [6]
	428–438	VAKRAESEGFEWVAFKTN	0 min	
	429–438	AKRAESEGFEWVAFKTN	0 min	
	430–438	KRAESEGFEWVAFKTN	0–10 min	
	431–439	RAESEGFEWVAFKTN	5–60 min	
	431–440	RAESEGFEWVAFKTN	5–120 min	
	431–441	RAESEGFEWVAFKTN	5–120 min	
	432–439	RAESEGFEWVAFKTN	5–15 min	
	432–440	RAESEGFEWVAFKTN	5–120 min	
	432–441	RAESEGFEWVAFKTN	5–15 min	
	472–479	AFQISREEARRLKYNRQETTLAR	0 min	Robotham et al (2009) [6]
	473–489	AFQISREEARRLKYNRQETTLAR	0 min	
	473–490	AFQISREEARRLKYNRQETTLAR	0 min	
	478–489	AFQISREEARRLKYNRQETTLAR	0 min	
	478–490	AFQISREEARRLKYNRQETTLAR	0 min	
	478–491	AFQISREEARRLKYNRQETTLAR	0 min	
	481–491	AFQISREEARRLKYNRQETTLAR	0 min	
	482–491	AFQISREEARRLKYNRQETTLAR	0 min	
	483–491	AFQISREEARRLKYNRQETTLAR	0 min	
Cor a 11	379–398	FEVNAHGNSRFPLAGKGNIV	0 min	Barre et al (2008) [5]
	379–400	FEVNAHGNSRFPLAGKGNIVNE	0 min	
	380–400	EVNAHGNSRFPLAGKGNIVNE	0 min	
	382–400	NAHGNSRFPLAGKGNIVNE	0 min	
	387–400	SRFPLAGKGNIVNE	0 min	
Cor a 14	105–125	QAVMQQQGEMRGEEMREVMET	5–10 min	Robotham et al (2002) [49] (for Jug r 1)
	109–125	QQQGEMRGEEMREVMET	5–10 min	
	110–125	QQGEMRGEEMREVMET	5 min	
	115–127	RGEEMREVMETAR	5 min	

Cor a 9, gastric phase





Figure SI-1: Alignment of peptides from the gastric phase (5–120 min) with Cor a 9.

Cor a 9, intestinal phase



Figure SI-2: Alignment of peptides from the intestinal phase (5–120 min) with Cor a 9.

Cor a 11, gastric phase



Cor a 11, intestinal phase

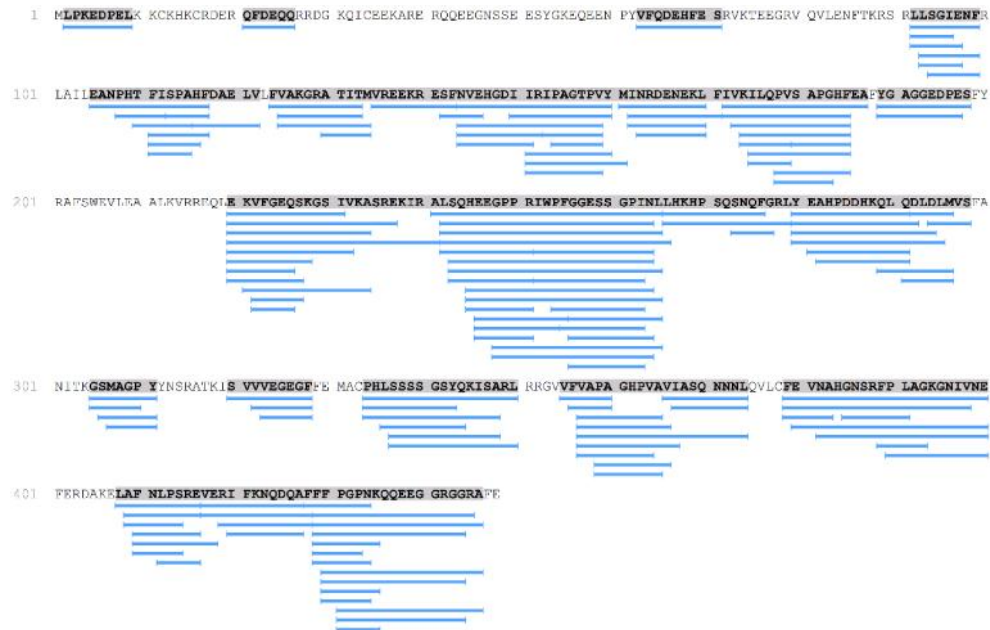
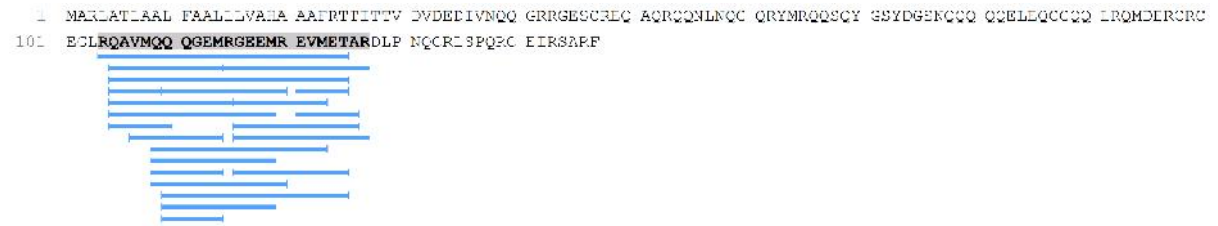


Figure SI-3: Alignment of peptides from the gastric and intestinal phases (5–120 min) with Cor a 11.

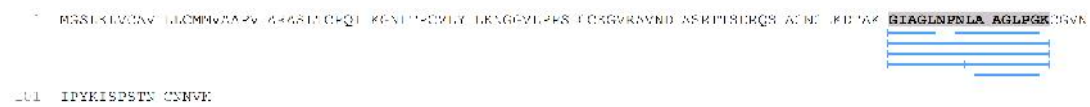
Cor a 14, gastric phase → no matches

Cor a 14, intestinal phase

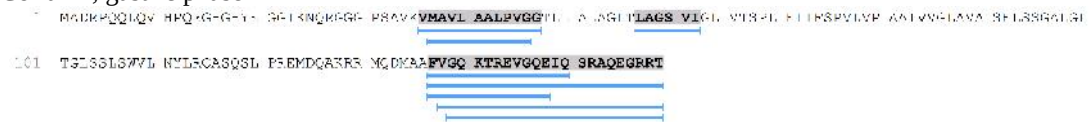


Cor a 8, gastric phase → no matches

Cor a 8, intestinal phase



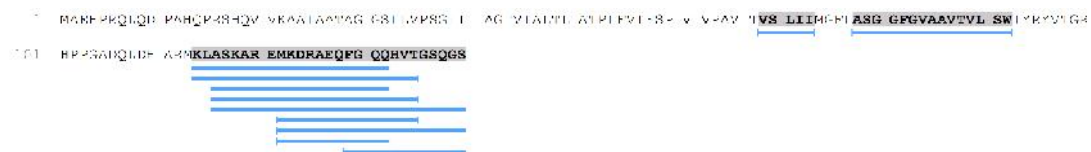
Cor a 12, gastric phase



Cor a 12, intestinal phase



Cor a 13, gastric phase

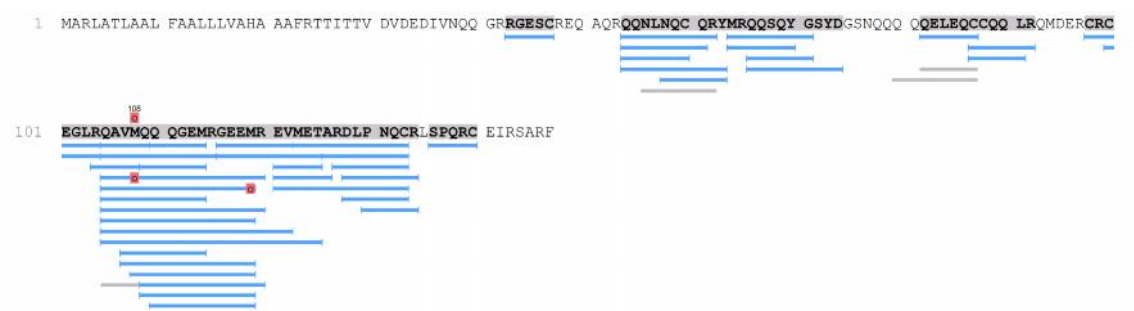


Cor a 13, intestinal phase



Figure SI-4: Alignment of peptides from the gastric and intestinal phases (5–120 min) with Cor a 14, Cor a 8, Cor a 12, and Cor a 13.

Cor a 14, intestinal phase 10 min, reduced



Cor a 8, intestinal phase 10 min, reduced

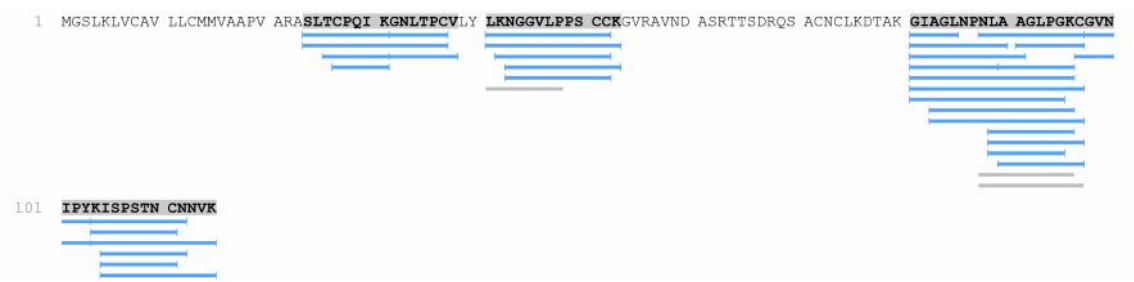


Figure SI-5: Alignment of peptides from the intestinal 10 min sample after reduction with dithiothreitol with Cor a 14 and Cor a 8.