Food and Chemical Toxicology 133 (2019) 110762

Contents lists available at ScienceDirect

Food and Chemical Toxicology

journal homepage: www.elsevier.com/locate/foodchemtox

Safety assessment of miraculin using *in silico* and *in vitro* digestibility analyses

Shahrzad Tafazoli[a], Trung D. Vo[a], Ashley Roberts[a,*], Chester Rodriguez[b], René Viñas[b], Michael E. Madonna[c], Yi-Hsuan Chiang[c], Juliana W. Noronha[c], Jonathon C. Holguin[c], Jason A. Ryder[c], Alan Perlstein[c]

[a] Intertek Scientific & Regulatory Consultancy, 2233, Argentia Road, Suite 201, Mississauga, Ontario, L5N 2X7, Canada
[b] The Coca-Cola Company, P.O. Box 1734, Atlanta, GA, 30313, United States
[c] Miraculex, 202 Cousteau Place, Suite 210, Davis, CA, 95618, United States

ARTICLE INFO

Keywords:
miraculin
Taste-modifier
SGF model
Allergenic

ABSTRACT

Miraculin is a glycoprotein with the ability to make sour substances taste sweet. The safety of miraculin has been evaluated using an approach proposed by the Food and Agriculture Organization of the United Nations and the World Health Organization for assessing the safety of novel proteins. Miraculin was shown to be fully and rapidly digested by pepsin in an *in vitro* digestibility assay. The proteomic analysis of miraculin's pepsin digests further corroborated that it is highly unlikely that any of the protein will remain intact within the gastrointestinal tract for potential absorption. The potential allergenicity and toxigenicity of miraculin, investigated using *in silico* bioinformatic analyses, demonstrated that miraculin does not represent a risk of allergy or toxicity to humans with low potential for cross-reactivity with other allergens. The results of a sensory study, characterizing the taste receptor activity of miraculin, showed that the taste-modifying effect of miraculin at the concentration intended for product development has a rapid onset and disappearance with no desensitizing impact on the receptor. Overall, the results of this study demonstrate that the use of miraculin to impact the sensory qualities of orally administered products with a bitter/sour taste profile is not associated with any safety concerns.

1. Introduction

The fruit of *Synepalum dulcificum*, also known as *Richadella dulcifica* (Uniprot, 2019a), a shrub native to tropical West Africa, has been known for over a century to make sour substances taste sweet (Kurihara and Beidler, 1968). The West African natives chewed the “miraculous berry”, also referred to as “sweet berry” (Hanelt, 2001; Roecklein and Leung, 1987), prior to food consumption to make acidic foods with an overly sour taste, such as *kankies* (sour cornbread), or intensely sour drinks, such as palm wine and *pitto* (a beer made from fermented grain), more palatable (Inglett et al., 1965). The unusual ability of these red berries to modify a sour into sweet taste earned them the name ‘Miracle Fruit’ (Morris, 1976).

The active component of the miracle fruit was isolated by 2 independent research groups in 1968 (Brouwer et al., 1968; Kurihara and Beidler, 1968). The researchers who identified the taste modifying substance demonstrated that it is a basic glycoprotein present within the thin-layered pulp of the miracle berry and named it “miraculin”. Subsequently, in 1988, Theerasilp and Kurihara (1988) successfully isolated the purified miraculin glycoprotein and determined it has a molecular weight of 28 kDa. Soon after, the same group determined the complete amino acid sequence of miraculin and demonstrated that it is a single polypeptide with 191 amino acid residues having 2 glycosylated residues, Asn-42 and Asn-186, cross-linked by a disulfide bond (Theerasilp and Kurihara, 1988; Theerasilp et al., 1989). The molecular weight of purified miraculin was calculated to be 24.6 kDa, taking into

Abbreviations: BLAST, Basic Local Alignment Search Tool; BLASTp, Basic Local Alignment Search Tool protein; CAC, Codex Alimentarius Commission; CAERS, Center for Food Safety and Applied Nutrition Adverse Event Reporting System; CFSAN, Center for Food Safety and Applied Nutrition; FAERS, United States Food and Drug Administration Adverse Event Reporting System; FAO, Food and Agriculture Organization of the United Nations; FARRP, Food Allergy Research and Resource Program; FASTA, FAST-All; FDA, United States Food and Drug Administration; GRAS, Generally Recognized as Safe; LC-MS/MS, liquid chromatography coupled to tandem mass spectroscopy; NCBI, National Center for Biotechnology Information; NDIN, new dietary ingredient notification; SGF, simulated gastric fluid; U.S., United States; WHO, World Health Organization

* Corresponding author. Intertek Scientific & Regulatory Consultancy, 2233 Argentia Road, Suite 201, Mississauga, Ontario, L5N 2X7. Canada.
E-mail address: ashley.roberts@intertek.com (A. Roberts).

https://doi.org/10.1016/j.fct.2019.110762
Received 3 June 2019; Received in revised form 26 July 2019; Accepted 14 August 2019
Available online 14 August 2019
0278-6915/ © 2019 Elsevier Ltd. All rights reserved.

account its amino acid sequence and a carbohydrate content of 13.9%, consisting mainly of glucosamine, mannose, galactose, xylose, and fucose (Theerasilp and Kurihara, 1988; Theerasilp et al., 1989). No amino acid homology was found between miraculin and sweet proteins, such as thaumatin and monellin (Theerasilp et al., 1989).

The taste-modifying effect of miraculin was shown to be species-specific with effects reported in chimpanzees, rhesus monkeys, and humans but not in rodents (Kurihara and Beidler, 1969; Diamant et al., 1972; Brouwer et al., 1983; Sanematsu et al., 2016). The exact molecular mechanism underlying the taste-modifying activity of miraculin has been the subject of much debate. Several investigators proposed that miraculin binds directly to the sweet-taste receptors, hT1R2-hT1R3, without activating them, and it is only when subjected to an acidic pH that miraculin activates these receptors (Kurihara, 1992; Ito et al., 2007; Asakura et al., 2011; Misaka, 2013; Sanematsu et al., 2016). The perceived taste of 0.1 M citrate after consuming 1 μM of miraculin was reported to correspond to the equivalent sweetness of 0.4 M sucrose, suggesting that miraculin is 400,000 times sweeter than sucrose on a molar basis (Kurihara, 1992; Kuroda, 2012). Miraculin's taste-modifying effects were reported to last from between approximately 20 min and up to 3 h in some cases with high concentrations (Kurihara and Beidler, 1969; Hellekant and van der Wel, 1989; Ito et al., 2007).

Attempts to introduce miracle berry products into the United States (U.S.) marketplace faced a number of regulatory challenges with the U.S. Food and Drug Administration (FDA) denying a petition filed in 1974 from a Massachusetts-based company, the Miralin Corporation, for affirmation that miracle fruit (*Synepalum dulcificum*) berries, concentrates, and extracts are Generally Recognized as Safe (GRAS) (U.S. FDA, 1977). The FDA issued a regulatory letter on September 19, 1974, indicating the Agency has determined that there was insufficient safety data to make a judgment of safe use of miracle fruit products, and subsequently requested that Miralin Co. "*take prompt action to cease interstate shipments of these products*" and that "*failure to do so would result in regulatory action by the agency*" (U.S. FDA, 1977). Following a careful evaluation of the available data, the Commissioner further advised that the petition submitted for GRAS affirmation was also evaluated as a food additive petition and concluded that, "*The data provided are incomplete and inadequate to establish the conditions of safe use of the additive. The substances cannot, therefore, be approved as a food additive*".

In 2009, a new dietary ingredient notification (NDIN) for 'Miracle Fruit Extract" was filed by My M Fruit, LLC for use in a dietary supplement called 'mberry' providing a serving size of 175 mg Miracle Fruit Extract per day in the form of a tablet to be consumed by allowing to dissolve on the tongue (U.S. FDA, 2009). In response to this notification, the FDA issued an objection letter on April 14, 2009, stating that the dietary supplement that would contain the NDI *"is not a dietary supplement within the meaning of 21 U.S.C. 321(ff)(2) (in particular, section (A)(i) which states: '[a dietary supplement] is intended for ingestion in a form described in section 21 U.S.C. 350(c)(1)(B)(i)."* The Agency also stated that, *"An article that is delivered orally, but that exerts its effect prior to being swallowed (for example, after dissolving on the tongue* […]*) or that is a delivery system for a substance that is absorbed buccolingually is not 'intended for ingestion*" (AHPA, FDA report NO: RPT574). Despite this, several dietary supplement-type products containing miracle berry or miracle fruit extracts are available on the U.S. market, including mberry Miracle Fruit Tablets, MiraBurst Easy Melt Miracle Berry Tablets, Miraculous Miracle Fruit Tablets, and Miracle Frooties Miracle Fruit Tablets. To date, there have been no adverse events or allergenic reactions resulting from consumption of these products that have been reported through the United States Food and Drug Administration Adverse Event Reporting System (FAERS) or the Center for Food Safety and Applied Nutrition (CFSAN) Adverse Event Reporting System (CAERS) (U.S. FDA, 2018a, b).

In an attempt to address the deficiencies highlighted by the FDA during their review of miracle fruit's GRAS affirmation petition, the safety of miraculin has been assessed by evaluating its susceptibility to *in vitro* digestion and determining its allergenic potential using bioinformatics, as well as conducting a sensory experiment to characterize the taste receptor activity. This approach follows the current strategies being applied for food safety assessment of novel proteins based on the guidelines provided by the Food and Agriculture Organization of the United Nations (FAO) and the World Health Organization (WHO) (Codex Alimentarius, 2003).

2. Materials and methods

2.1. Preparation of miraculin

Whole miracle berries were obtained from Miracle Fruit Farm (Miami, Fl) and deseeded. The pulp and skin tissue of the berries were separated by high speed mechanical agitation. The extraction of miraculin from the pulp was carried out according to method as described by Duhita et al. (2009) with some modifications. First, the subsequent mixture was homogenized *via* high speed mixing. The resulting slurry was centrifuged to separate the insoluble material from the supernatant. Any remaining insoluble material was removed from the supernatant via a 0.45 μm microfiltration step. The sediment was suspended in 10 mM citric acid with 0.5 M NaCl (pH 3.5) for extraction of miraculin. The pH of the solution was adjusted to pH 7.2 to 7.3 with 1 M Tris-base prior to column loading. All procedures were performed below 6 °C. Nickel-immobilized affinity chromatography was performed according to previous described methods, with some modifications on the binding, washing, and elution buffers (Duhita et al., 2009). The column was equilibrated and washed with binding buffer (20 mM Tris-base, 10 mM citric acid, 0.5 M NaCl at pH 7.2 to 7.3). A sample of the extracted miraculin solution was loaded onto the column and the target protein was eluted with the binding buffer at pH 4.6 to 4.7. All other experimental conditions were performed as described by Duhita et al. (2009). The resulting eluted protein solution was dialyzed with deionized water and then freeze-dried.

2.2. Digestibility of miraculin

The protein digestibility of miraculin was investigated using a simulated gastric fluid (SGF) model as described by Thomas et al. (2004). The *in vitro* digestibility study was conducted using SGF that was designed to mimic the conditions of the human stomach. A preincubation mixture consisting of 1.52 mL SGF (35 mM NaCl, 0.084 N HCl, 10 U/μg pepsin, pH 2.0) was preheated to 37 °C prior to the addition of 0.08 mL miraculin (2 mg/mL, final concentration of 0.1 mg/mL). An identical mixture was prepared with 5.45 U/μg pepsin. Samples of 0.1 mL were collected at 0, 20, 40, and 60 min of incubation. Samples were quenched with the addition of 35 μL of 200 mM sodium bicarbonate, 168.8 μL of Novex™ Tricine SDS Sample Buffer (2X) (ThermoFisher Scientific), and 33.8 μL NuPAGE™ Sample Reducing Agent (10X) (ThermoFisher Scientific). A control sample (no incubation) was quenched prior to the addition of miraculin. All quenched samples were heated at 85 °C for 10 min. Following the heating period, the digestibility of the protein (296 ng) was evaluated by gel electrophoresis (Novex® 10–20% Tricine Protein Gels, Novex® Tricine Mini Gels; Invitrogen-ThermoFisher Scientific) with Coomassie blue staining (Imperial™ Protein Stain; ThermoFisher Scientific). Serial dilutions of undigested miraculin were also prepared.

2.3. Pepsin digest mapping of miraculin

The proteolytic fate of the protein after pepsin digestion was assessed by JadeBio (La Jolla, CA) using liquid chromatography coupled to tandem mass spectroscopy (LC-MS/MS). The protein was digested by the addition of 0.08 mL miraculin (2 mg/mL) to 1.52 mL of SGF buffer with 5.45 U/μg pepsin. Samples of 0.1 mL were collected at 0 (no

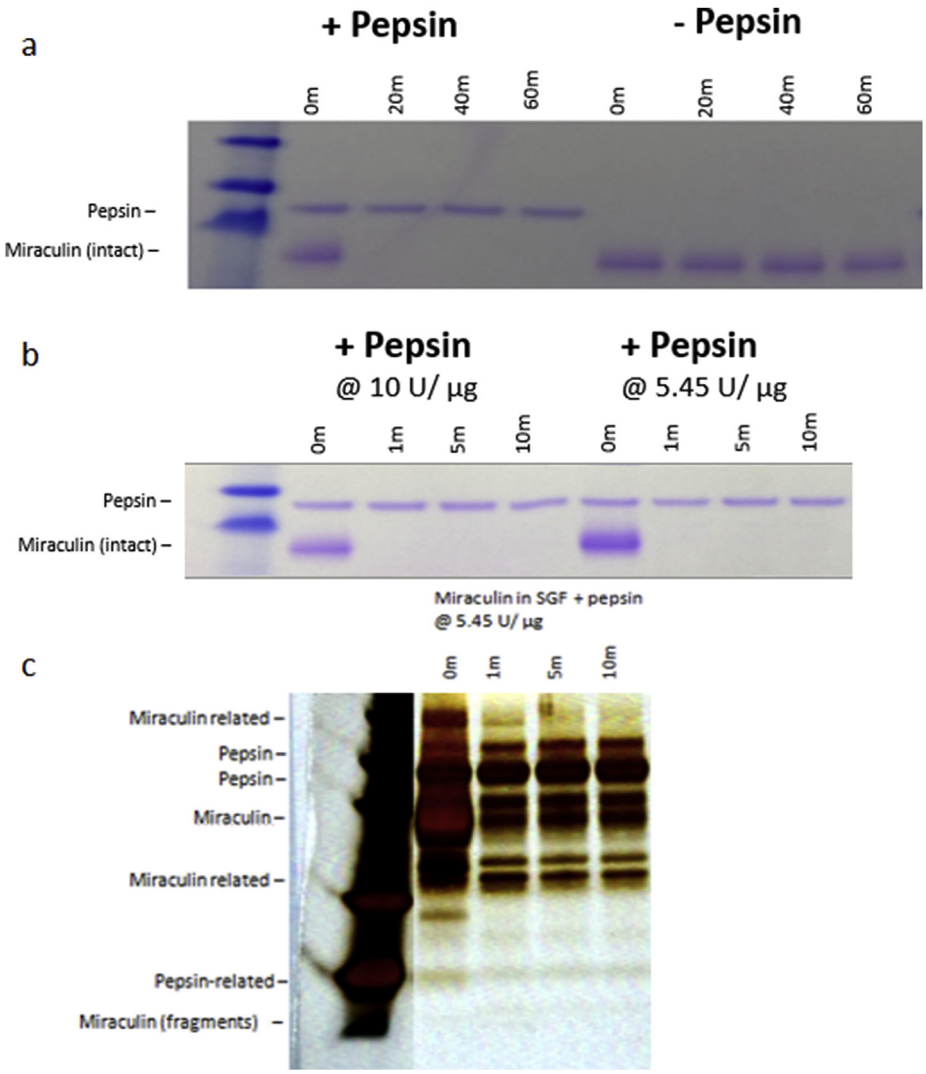

Fig. 1. Results of *in vitro* digestion of miraculin in the presence of SGF with and without pepsin for 60 min (a), with pepsin at concentrations of 5.45 U/g and 10 U/g for 10 min (b), and silver stained gels after digestion with 5.45 U/g pepsin for 10 min (c).

pepsin), 0.5, 1, and 10 min of incubation at 37 °C and quenched with 35 µL 200 mM sodium bicarbonate. Using a 10 µL (1 µg total protein) injection, samples were analyzed by C18 reverse-phase LC-MS/MS on an Agilent Q-TOF 6550. Data was assessed using SpectrumMill (B.06.00.201), comparing MS and MS/MS ions against the miraculin sequence.

2.4. In silico assessment of allergenicity potential of miraculin

The potential allergenicity of miraculin was investigated in a similar manner as that described previously by Jin et al. (2017) using the approach described by the FAO/WHO (2001) and Codex Alimentarius (2009). The allergenicity was evaluated using the well-documented allergen database (AllergenOnline, Version 19; http://www.allergenonline.org/). AllergenOnline is a peer-reviewed database maintained by the Food Allergy Research and Resource Program (FARRP) of the University of Nebraska and contains a comprehensive list of known or putative allergenic proteins. The current version of the AllergenOnline database (Version 19; February 10, 2019) contains 2,129 peer-reviewed allergenic sequences. Three searches were conducted: (1) full-length sequence alignment; (2) 80 amino acid sequence alignment; and (3) 8 amino acid exact match (FARRP, 2019). The full-length sequence alignment search queries the entire length of the protein of interest with the full sequence of a known allergen. The 80 amino acid sequence alignment queries a 'sliding window' of 80 amino acid sequences (segments 1 to 80, 2 to 81, 3 to 82, etc.) derived from each full-length sequence. The 8 amino acid exact match queries for an exact match of 8 contiguous amino acids within the full sequence of a known allergen. The expectation value (E-value) cutoff for the full-length and 80 amino acid sequence alignment searches was set as the default (E-value cut off = 1.0 to 10 and 10, for full-length and 80 amino acid sequence searches, respectively). The searches were conducted using FASTA Version 35.04 (updated in January 15, 2009). To qualify the implications of the results, significant sequence homology was defined as an identity match of greater than 35%, of which greater may suggest sequence homology and evidence for cross-reactivity (Goodman et al., 2008; Ladics et al., 2011; Jin et al., 2017). Searches were conducted on February 28, 2019.

2.5. In silico assessment of toxigenicity potential of miraculin

The potential toxigenicity of miraculin was investigated against a curated database of venom proteins and animal toxins (Jungo et al., 2012) using the Basic Local Alignment Search Tool protein (BLASTp) suite maintained by the National Center for Biotechnology Information (NCBI). The toxin database was downloaded in FASTA format and contained 6,841 sequences of toxins produced by snakes, spiders, scorpions, cone snails, jellyfish, insects, sea anemones, lizards, fish

species, and platypuses, and other venomous animals that lack venom devices (toads, ticks, and worms) (Uniprot, 2019b). A full-length sequence alignment search was conducted to ascertain whether the miraculin protein sequence share homology or structural similarity to a known toxin. The sequence alignment search was conducted on January 25, 2019 with the BLASTp suite using the default settings (maximum target sequence: 100; expect threshold: 10; matrix: BLOSUM62).

2.6. Sensory evaluation of miraculin activity

The sweetness profile of miraculin (as a fruit extract powder of the miracle berry) was evaluated by 6 panelists who had previous experience evaluating sweetness intensity. Panelists were instructed to sip lemonade juice with a sweetness intensity of 7 brix until a baseline sweetness intensity was established. Each panelist was provided 0.08 g of miracle fruit powder and were instructed to consume but hold the powder in the mouth for 1 min before swallowing. Panelists consumed 60 mL of lemonade juice every 5 min until a total of 420 mL was consumed. The sweetness of each cup was recorded at the following intervals: 0, 5, 10, 15, 20, 25, and 30 min.

3. Results

3.1. Digestibility of miraculin

The results of the *in vitro* digestibility assay using Coomassie blue staining indicate that miraculin was completely digested within 20 min in the presence of the SGF containing pepsin (Fig. 1a). However, in the presence of the SGF without pepsin no digestion of the miraculin protein was highlighted. When incubated with pepsin at concentrations of 5.45 and 10 U/μg, miraculin was fully digested within 1 min, indicating a very rapid metabolism, irrespective of the pepsin concentration suggesting that the rate of digestion of the protein is pepsin-and not low pH dependent (Fig. 1b). Furthermore, silver staining the electrophoresis gels reveals a number of different fragments, indicating that the protein would be fully digested within the human gastrointestinal tract (Fig. 1c).

3.2. Pepsin digest mapping of miraculin

Miraculin was increasingly digested by pepsin with longer digestion time, generating 61 unique peptides after 10 min of pepsin digestion, as analyzed using LC-MS/MS (Table 1). The sequence coverage was approximately 40% at 30 s, 62% at 1 min, and 75% at 10 min, suggesting that sequence coverage increased with longer digestion time. At 10 min, the only peptides that were not identified contained cysteine residues (i.e., disulfide bonds to other peptides). The peptide sequence map is shown in Fig. 2. The miraculin amino acid sequence contained several potential *N*-linked glycosylation sites. The peptides sequenced from the miraculin pepsin digests are provided in Table 2. The mapping of the miraculin protein digest study further corroborate the rapidity of the digestion and that it is highly unlikely that any of the protein will remain intact for potential absorption.

Table 1
LC-MS/MS results of miraculin pepsin digestion.

Timepoint (min)	Total Intensity	Unique Peptide	Sequence Coverage
0	7.5×10^5	5	17%
0.5	2.5×10^7	33	40%
1	7.7×10^7	54	62%
10	2.2×10^8	61	75%

3.3. In silico assessment of allergenicity potential of miraculin

A search of the full-length amino acid sequence of miraculin revealed significant similarity with known allergens from *Solanum tuberosum* (potato) and *Glycine* max (soybean), with sequence identities ranging between 26.6 and 33.7%, and E-values ranging from 0.34 to 4.1×10^{-8} (Table 3). Upon closer examination of the sequence alignment of these proteins, the highest peptide match was "YKLVFCP" that is present in both miraculin and the soybean Kunitz trypsin inhibitor KTi2. AllergenOnline identified 112 '80mer sliding windows' along the 191 amino acid length of the miraculin sequence. A search using this approach revealed 7 and 9 of '80 amino acid' matches that had greater than 35% identity with the allergenic proteins Kunitz trypsin inhibitor KTi2 from *G.* max and proteinase inhibitor from *S. tuberosum*, respectively (Table 3). Specifically, the miraculin amino acid sequence shared 38% identity with Kunitz trypsin inhibitor KTi2 and 36% identity with proteinase inhibitor. The search for identity with 8 contiguous amino acids revealed no significant similarity with any known allergens (data not shown). The peptide sequences identified following pepsin digest mapping were searched using the full-length and 80mer sliding window approaches. Searches with the full-length sequences of the pepsin-digested peptides revealed a number of hits with known allergens having identities ranging from 36% to 67%, and similarities ranging from 60 to 100% (Table S1 in Supplementary Materials). No sequence homology to known allergens from the AllergenOnline database were identified in the 80mer sliding window search. Taken together, the results of the *in silico* assessment suggest that miraculin does not pose a risk of cross-reactivity with known allergens.

3.4. In silico assessment of toxigenicity potential of miraculin

A comparison of the primary amino acid sequence of miraculin with known protein toxins, using full-length sequence alignment search, identified matches with sequence identities ranging from 28 to 54% and E-values ranging from 0.61 to 9.0 with a number of animal protein toxins (Table 4). Based on the total query cover of the miraculin protein sequence, which ranged from 6 to 26%, no or low structural homology to protein toxins is expected.

3.5. Sweetness intensity to evaluate protein activity

Miracle fruit powder significantly increased the sweetness of lemonade juice within 1–2 min (Fig. 3) in all 6 individuals tested. After 30 min of cessation of miracle fruit powder consumption, the sweetness of lemonade juice returned to baseline levels in all of the study subjects. This indicates that the onset and disappearance of the taste modifying effects of miraculin are rapid with no lasting desensitization effect.

4. Discussion

Using a standard *in vitro* pepsin SGF digestibility assay, miraculin was shown to be rapidly and completely digested under simulated physiologic gastric conditions. These findings were corroborated with the results of a follow-up proteomic analysis wherein pepsin digests of miraculin were fractionated and analyzed separately by LC-MS/MS, which demonstrated that sequence coverage increased with longer digestion time. The rapidity of the glycoprotein digestion supports the fact that the protein will undergo hydrolysis within the gastrointestinal tract following consumption and will therefore not be absorbed intact and will not bind to any other sweet taste receptors other than those on the tongue. The rapidity and the degree of digestion of the miraculin glycoprotein also indicates that the protein in its natural form will not be present systemically to elicit an allergenic or toxicogenic response. These results were further assessed in terms of the Codex Alimentarius Commission (CAC) 'weight of evidence' approach for the assessment of newly expressed proteins, which includes determination of the source

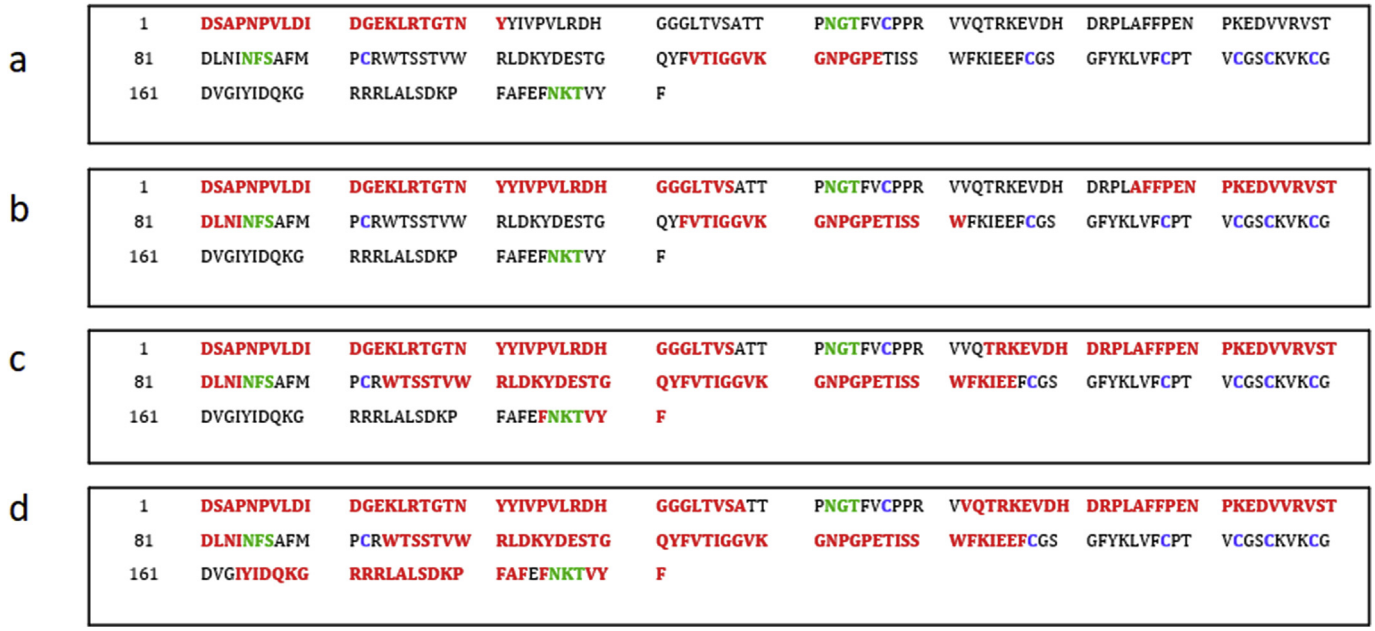

Fig. 2. Sequence Map of Miraculin Pepsin Digestion at Timepoints 0 min (a), 0.5 min (b), 1 min (c), and 10 min (d).

of the introduced protein, assessing similarities between the amino acid sequence of the novel protein and that of known allergens and toxins, as well as an evaluation of its susceptibility to enzymatic degradation (Codex Alimentarius, 2003). The CAC indicated that resistance of proteins to proteolytic digestion was an important consideration in the safety assessment paradigm of novel proteins, based on a proposed link between the stability of a protein in SGF and its status as a food allergen (Astwood et al., 1996; Taylor and Lehrer, 1996; Bannon, 2004; EFSA, 2017). It is understood that an important characteristic of food allergens is their resistance to denaturation and digestion, because food allergens must be proteolytically stable and resistant to gastric digestion by pepsin to reach the intestinal mucosa and elicit their allergenic response (Bannon, 2004; Herman et al., 2007). The standardized *in vitro* pepsin digestibility tests conducted with the miraculin protein based on the conditions described in the U.S. Pharmacopeia (USP, 1995) (i.e., pepsin concentration, pH, temperature, etc.); therefore, provide a predictive tool for protein allergenicity assessment (Astwood et al., 1996; Thomas et al., 2004). Although this pepsin digestibility assay is not meant to precisely mimic the *in vivo* conditions, they are designed to evaluate the susceptibility of a protein to digestion under fixed conditions *in vitro*, and along with other evidence, such as *in silico* analyses, provide information on the allergenicity potential of a novel protein. Therefore, the *in vitro* digestibility analysis clearly shows that the miraculin protein is not proteolytically stable and able to reach the intestinal mucosa.

While the *in vitro* digestibility analysis demonstrates rapid and complete hydrolysis of miraculin, a corroborative step in the safety assessment of novel proteins also includes comparing its primary amino acid sequence with those of known allergens (Codex Alimentarius, 2009; Ladics et al., 2011). Regardless of the source of a novel protein, shared amino acid sequence similarities to other known proteins can provide valuable information on the level of structural relatedness to proteins with a known history of safe use or to proteins that are known to be allergens (Ladics et al., 2011). These sequence comparison searches for the miraculin glycoprotein were conducted using various algorithms including the Basic Local Alignment Search Tool (BLAST) or FASTA (Ladics et al., 2007), which are designed to find evolutionary relationships between protein sequences based upon regions of shared amino acid identity and similarity to known allergenic proteins that are maintained in a database at the FARRP, University of Nebraska (http://www.allergenonline.org/). A search of the full-length amino acid sequence of the miraculin glycoprotein as reported by Theerasilp et al. (1989), with known allergens from the AllergenOnline database did reveal similarity with several allergenic proteins from *Glycine* max and *Solanum tuberosum*. While it has been determined that alignments with high identity scores (50–70%) may indicate a potential for allergenic cross-reactivity (Aalberse, 2000), an E-value larger than 1×10^{-7} is unlikely to identify proteins that may share immunologic or allergic cross-reactivity to known allergens (Hileman et al., 2002). Therefore, on the basis of these criteria, the full-length sequences of miraculin was determined to share less than 35% identity over the full length of the amino acid sequence with corresponding E-values ranging from 4.1×10^{-8} to 0.34. The protein sequence with a reported E-value of 4.1×10^{-8} was proteinase inhibitor from *Solanum tuberosum* with a corresponding identity of 27.7% and 52.5% similarity. E-values greater than 1×10^{-7} are suggestive of a "significant" similarity between proteins (Ladics, 2018). Therefore, given that the E-value was less than 1×10^{-7} and considering that the sequences shared less than 35% identity over approximately 50% of the amino acid sequence, it is unlikely that miraculin is associated with an allergenic risk potential. However, when a window of 80 amino acids was applied as a search parameter in the same allergen database, the amino acid sequence of miraculin showed a reasonable number of significant alignments (greater than 35% identity) with 2 proteins, Kunitz trypsin inhibitor KTi2 and proteinase inhibitor from *Glycine* max and *Solanum tuberosum*, respectively. Sequences sharing greater than 35% identity over an 80 amino acid window with known allergens may have a possibility of cross-reaction (Codex Alimentarius, 2009). These proteins are not recognized as major allergens in potatoes or soybean (Ogawa et al., 2000; Seppälä et al., 2001). Despite being listed as 1 of 8 major allergens in the U.S., the incidence rate of allergenic reaction to soy is reported to be as low as 0.1%–0.25% (Cordle, 2004; Loh and Tang, 2018). Potatoes are not listed as a major food allergen, and despite its high frequency of consumption across different demographic groups, reported food allergy to potatoes is uncommon (Castells et al., 1986; Seppälä et al., 2001).The pepsin digest mapping of the miraculin protein revealed that the protein would be digested into small peptides of approximately 8–31 amino acids in length after 10 min. These small peptides were searched against the AllergenOnline database using the full-length amino acid sequence approach and a number of the peptides were

Table 2
Peptides sequenced from miraculin pepsin digestion.

Starting Amino Acid	Sequence	Retention Time	Pepsin-T0		Pepsin-T30s		Pepsin-T60s		Pepsin-T10min	
			Spectra Number	Intensity	Spectra Number	Intensity	Spectra Number	Intensity	Spectra Number	Intensity
1	(-)DSAPNPVLDIDGEKLR(T)	15.4	0		2	2.86×10^5	2	9.30×10^5	1	1.10×10^5
1	(-)DSAPNPVLDIDGEKLRTGTN(Y)	15.3	1	7.76×10^4	2	1.19×10^6	2	2.62×10^6	1	3.06×10^5
1	(-)DSAPNPVLDIDGEKLRTGTNY(Y)	16	1	2.59×10^5	2	4.30×10^6	2	9.29×10^6	2	1.44×10^6
1	(-)DSAPNPVLDIDGEKLRTGTNYYIVPVLR(D)	19.1	0		2	2.63×10^5	0		0	
2	(D)SAPNPVLDIDGEKLR(T)	15	0		2	2.42×10^5	2	5.76×10^5	1	1.21×10^5
2	(D)SAPNPVLDIDGEKLRTGTN(Y)	15	1	5.60×10^4	2	1.18×10^6	2	2.58×10^6	2	4.65×10^5
2	(D)SAPNPVLDIDGEKLRTGTNY(Y)	15.7	2	3.08×10^5	2	5.04×10^6	2	1.07×10^7	2	1.78×10^6
2	(D)SAPNPVLDIDGEKLRTGTNYYIVPVLR(D)	18.7	0		1	1.07×10^5	1	1.19×10^5	0	
3	(S)APNPVLDIDGEKLRTGTN(Y)	14.8	0		1	5.80×10^4	1	1.04×10^5	0	
3	(S)APNPVLDIDGEKLRTGTNY(Y)	15.5	0		1	1.41×10^5	2	4.66×10^5	1	9.29×10^4
4	(A)PNPVLDIDGEKLR(T)	15	0		0		2	5.11×10^4	0	
4	(A)PNPVLDIDGEKLRTGTN(Y)	15	0		0		1	9.97×10^4	0	
4	(A)PNPVLDIDGEKLRTGTNY(Y)	16	0		1	1.57×10^5	2	4.31×10^5	0	
6	(N)PVLDIDGEKLRTGTNY(Y)	15.6	0		0		1	7.27×10^4	0	
9	(L)DIDGEKLR(T)	9.6	0		0		1	7.44×10^4	1	3.31×10^5
9	(L)DIDGEKLRTGTN(Y)	10.1	0		2	2.61×10^5	2	5.92×10^5	2	4.74×10^6
9	(L)DIDGEKLRTGTNY(Y)	11.7	0		1	2.49×10^5	1	7.67×10^5	1	9.03×10^6
9	(L)DIDGEKLRTGTNYYIVPVLR(D)	17	0		0		1	7.58×10^4	0	
10	(D)IDGEKLR(T)	9.6	0		0		0		1	1.85×10^4
10	(D)IDGEKLRTGTNY(Y)	10.8	0		0		1	2.42×10^4	1	4.91×10^4
11	(I)DGEKLRTGTN(Y)	10.1	0		0		1	4.43×10^4	1	3.00×10^5
11	(I)DGEKLRTGTNY(Y)	11.7	0		1	2.48×10^4	1	8.04×10^4	1	6.00×10^5
21	(N)YYIVPVLR(D)	17	0		1	2.29×10^5	1	8.62×10^5	1	1.21×10^6
21	(N)YYIVPVLRDHGGGL(T)	17	0		0		0		2	2.25×10^5
21	(N)YYIVPVLRDHGGGLT(V)	16.6	0		0		1	3.68×10^4	1	7.52×10^4
21	(N)YYIVPVLRDHGGGLTVS(A)	17.1	0		1	7.07×10^4	1	9.11×10^4	2	1.58×10^6
22	(Y)YIVPVLR(D)	15.6	0		1	9.67×10^5	1	3.59×10^6	1	6.03×10^6
22	(Y)YIVPVLRDHGGGL(T)	15.9	0		1	6.21×10^4	1	2.94×10^4	2	1.18×10^6
22	(Y)YIVPVLRDHGGGLT(V)	15.5	0		0		1	1.27×10^5	1	2.77×10^5
22	(Y)YIVPVLRDHGGGLTV(S)	16.8	0		0		0		1	1.43×10^5
22	(Y)YIVPVLRDHGGGLTVS(A)	16.2	0		1	2.08×10^5	1	3.81×10^5	1	1.98×10^6
22	(Y)YIVPVLRDHGGGLTVSA(T)	16.5	0		0		0		1	1.72×10^5
24	(I)VPVLRDHGGGL(T)	15.9	0		0		0		1	4.62×10^4
53	(V)VQTRKEVDHDRPLAF(F)	11.7	0		0		0		1	3.29×10^5
55	(Q)TRKEVDHDRPLAF(F)	11.4	0		0		0		2	5.26×10^6
55	(Q)TRKEVDHDRPLAFFPENPKE(D)	13.9	0		0		1	7.72×10^4	0	
55	(Q)TRKEVDHDRPLAFFPENPKEDVVRVSTDL(N)	16.7	0		0		1	3.70×10^5	0	
56	(T)RKEVDHDRPLAF(F)	11.3	0		0		0		2	6.54×10^5
66	(L)AFFPENPK(E)	13.2	0		1	7.78×10^4	1	9.83×10^4	1	1.00×10^5
66	(L)AFFPENPKEDVVRVSTDL(N)	17.3	0		1	2.90×10^4	0		0	
68	(F)FPENPKE(D)	9.6	0		2	1.17×10^6	2	2.30×10^6	3	9.98×10^6
68	(F)FPENPKEDV(V)	11.7	0		0		0		2	8.78×10^5
68	(F)FPENPKEDVVRV(S)	13.5	0		0		0		1	3.69×10^5
68	(F)FPENPKEDVVRVS(T)	12.9	0		0		3	7.12×10^5	2	1.04×10^6
68	(F)FPENPKEDVVRVSTD(L)	13.1	0		1	1.95×10^5	2	9.60×10^5	2	3.10×10^6
68	(F)FPENPKEDVVRVSTDL(N)	15.5	0		3	4.06×10^6	4	2.07×10^7	4	6.55×10^7
68	(F)FPENPKEDVVRVSTDLN(I)	14.7	0		0		1	3.17×10^5	0	
68	(F)FPENPKEDVVRVSTDLNINF(S)	18.4	0		1	1.31×10^6	1	2.84×10^6	1	2.99×10^5
69	(F)PENPKEDVVRVSTD(L)	11.1	0		0		0		2	2.03×10^5
6	(F)PENPKEDVVRVSTDL(N)	14.1	0		1	7.40×10^4	2	1.00×10^6	3	4.83×10^6
69	(F)PENPKEDVVRVSTDLNINF(S)	17.5	0		0		1	3.00×10^5	0	
75	(E)DVVRVSTD(L)	10	0		0		0		1	5.26×10^5
75	(E)DVVRVSTDL(N)	14	0		1	6.53×10^5	1	1.88×10^6	1	1.37×10^7
75	(E)DVVRVSTDLN(I)	12.8	0		0		1	3.09×10^4	0	
75	(E)DVVRVSTDLNINF(S)	18.2	0		1	1.27×10^5	1	3.82×10^5	1	1.75×10^5
77	(V)VRVSTDLNINF(S)	17.1	0		0		1	4.84×10^4	0	
96	(R)WTSSTVWRLDKYDESTGQY(F)	15.9	0		0		1	1.17×10^5	1	1.73×10^5
100	(S)TVWRLDKYDESTGQY(F)	14.3	0		0		0		1	3.73×10^5
101	(T)VWRLDKYDESTGQY(F)	14	0		0		0		1	1.48×10^6
102	(W)RLDKYDESTGQY(F)	10.3	0		0		0		1	2.10×10^5
105	(L)DKYDESTGQYF(V)	13.1	0		0		0		1	1.61×10^5
116	(F)VTIGGVKGNPGPE(T)	11.7	1	5.25×10^4	1	4.63×10^4	1	4.89×10^4	0	
116	(F)VTIGGVKGNPGPETISS(W)	13.2	0		0		0		1	1.15×10^6
116	(F)VTIGGVKGNPGPETISSW(F)	16.3	0		1	2.56×10^5	1	7.62×10^5	2	5.89×10^6
118	(T)IGGVKGNPGPETISSW(F)	15.3	0		0		1	4.83×10^4	1	1.64×10^5
125	(Y)FVTIGGVKGNPGPE(T)	13.8	0		1	1.51×10^4	1	1.42×10^4	0	
125	(Y)FVTIGGVKGNPGPET(I)	13.8	0		0		1	2.58×10^4	2	3.20×10^5
125	(Y)FVTIGGVKGNPGPETIS(S)	15	0		0		1	6.27×10^4	0	
125	(Y)FVTIGGVKGNPGPETISS(W)	14.8	0		1	3.56×10^5	1	1.47×10^6	2	9.77×10^6
125	(Y)FVTIGGVKGNPGPETISSW(F)	17.6	0		2	1.60×10^6	2	7.36×10^6	2	4.08×10^7
125	(Y)FVTIGGVKGNPGPETISSWF(K)	20.4	0		0		1	1.04×10^5	2	6.16×10^5

(*continued on next page*)

Table 2 (*continued*)

Starting Amino Acid	Sequence	Retention Time	Pepsin-T0		Pepsin-T30s		Pepsin-T60s		Pepsin-T10min	
			Spectra Number	Intensity	Spectra Number	Intensity	Spectra Number	Intensity	Spectra Number	Intensity
125	(Y)FVTIGGVKGNPGPETISSWFKIEE(F)	20	0		0		1	3.34×10^4	0	
133	(S)WFKIEE(F)	15.1	0		0		0		1	5.15×10^6
133	(S)WFKIEEF(C)	18.1	0		0		0		1	1.11×10^5
170	(G)IYIDQKGRRRLAL(S)	12.5	0		0		0		1	2.13×10^5
170	(G)IYIDQKGRRRLALSDKPFA(F)	13.5	0		0		0		2	2.73×10^6
170	(G)IYIDQKGRRRLALSDKPFAF(E)	15.6	0		0		0		1	1.05×10^6
172	(Y)IDQKGRRRLALSDKPFAF(E)	14.3	0		0		0		1	1.33×10^7
183	(L)SDKPFAF(E)	14.7	0		0		0		1	3.75×10^5
191	(E)FNKTVYF(-)	14.3	0		0		1	4.14×10^5	1	4.78×10^5

Table 3
Search results of AllergenOnline database version 19 with miraculin.

Sequence G.I. #	Organism	Description	80 mer[a]		Full Length			
			% Identity	# Hits (> 35%)	Length	E-value	% Identity	Amino acid overlap
994779	*Solanum tuberosum*	proteinase inhibitor	36.00	9/112	221	4.1×10^{-8}	27.7	202
124148	*Solanum tuberosum*	Aspartic protease inhibitor 11	–	–	188	1.7×10^{-7}	28.4	201
256429	*Glycine max*	Kunitz trypsin inhibitor KTi	–	–	216	1.2×10^{-6}	31.9	191
18770	*Glycine max*	trypsin inhibitor subtype A	–	–	217	1.2×10^{-6}	31.9	191
256635	*Glycine max*	Kunitz trypsin inhibitor KTi1	–	–	203	3.3×10^{-6}	33.7	187
18772	*Glycine max*	trypsin inhibitor subtype B	–	–	217	3.4×10^{-6}	32.3	189
256636	*Glycine max*	Kunitz trypsin inhibitor KTi2	37.54	7/112	204	5.2×10^{-6}	33.0	185
510515	*Glycine max*	Kunitz trypsin inhibitor	–	–	208	2.8×10^{-5}	26.6	184
20141344	*Solanum tuberosum*	cysteine protease inhibitor 1	–	–	222	0.11	30.3	198
1575306	*Solanum tuberosum*	cysteine proteinase inhibitor 10 precursor, partial	–	–	186	0.34	27.9	197

[a] The amino acid sequence had a reported length of 191 amino acids and 112 80 mers in the sliding window search.

determined to share sequence homology with known allergens. The identities ranged from 36 to 67%, while the similarity of the sequences ranged from 60 to 100% (see Table S1 in Supplementary Materials). These *in silico* results suggest that there is a potential for allergenicity following digestion of the miraculin protein. However, the pepsin digests are relatively small in size such that the highest molecular weight peptide was in the range of 4 kDa. A minimum requirement for a food allergen to elicit an allergenic response is to have at least 2 epitopes, each of which being approximately 15 amino acid residues long (Huby et al., 2000), suggesting a size limit of, at minimum, 30 amino acid residues (corresponding to a molecular weight of approximately 3 kD). The molecular weight of food allergens was suggested to typically range

Table 4
Results of the toxigenicity search of miraculin.

Organism	Description	Sequence Length	Query Cover	E-Value	Identity
Daboia siamensis	Snaclec 5	148	21%	0.61	29%
Daboia siamensis	Snaclec 3	148	21%	0.62	29%
Tropidechis carinatus	Venom nerve growth factor 4	244	19%	0.65	33%
Tropidechis carinatus	Venom nerve growth factor 2	244	19%	0.75	33%
Hoplocephalus stephensii	Venom nerve growth factor 2	245	20%	0.82	32%
Tropidechis carinatus	Venom nerve growth factor 1	245	26%	0.83	31%
Bothrops pauloensis	L-amino-acid oxidase (Fragment)	503	18%	1.1	34%
Crotalus durissus cumanensis	L-amino acid oxidase Cdc18 (Fragment)	498	18%	1.4	33%
Crotalus atrox	L-amino-acid oxidase	516	18%	1.4	33%
Crotalus adamanteus	L-amino-acid oxidase	516	18%	1.4	33%
Conus ventricosus	Conotoxin VnMKLT1-012	77	25%	1.5	33%
Notechis scutatus scutatus	L-amino-acid oxidase	517	20%	1.8	30%
Stichodactyla helianthus	Kappa-stichotoxin-She3a	35	6%	2.1	54%
Bothriechis schlegelii	L-amino acid oxidase Bs29 (Fragment)	498	18%	2.4	34%
Bothrops jararacussu	L-amino-acid oxidase (Fragment)	497	18%	2.8	34%
Bothrops moojeni	L-amino-acid oxidase (Fragment)	478	18%	3.0	34%
Theraphosa blondi	Kappa-theraphotoxin-Tb1c	35	10%	3.8	38%
Conus leopardus	Alpha-conotoxin-like Lp1.6a	61	10%	4.8	50%
Pseudechis australis	L-amino-acid oxidase	517	20%	5.1	28%
Notechis scutatus scutatus	Venom nerve growth factor 1	244	19%	5.3	31%
Oxyuranus scutellatus scutellatus	L-amino-acid oxidase	517	20%	6.1	28%
Dendroaspis angusticeps	Muscarinic toxin 1	66	18%	7.3	42%
Crotalus adamanteus	L-amino-acid oxidase	516	18%	7.7	31%
Dendroaspis angusticeps	Muscarinic toxin 4	66	18%	8.3	42%
Thalassianthus aster	Kappa-thalatoxin-Tas2a	75	6%	8.8	54%
Heterodactyla hemprichii	Kappa-thalatoxin-Hhe2a	75	6%	9.0	54%

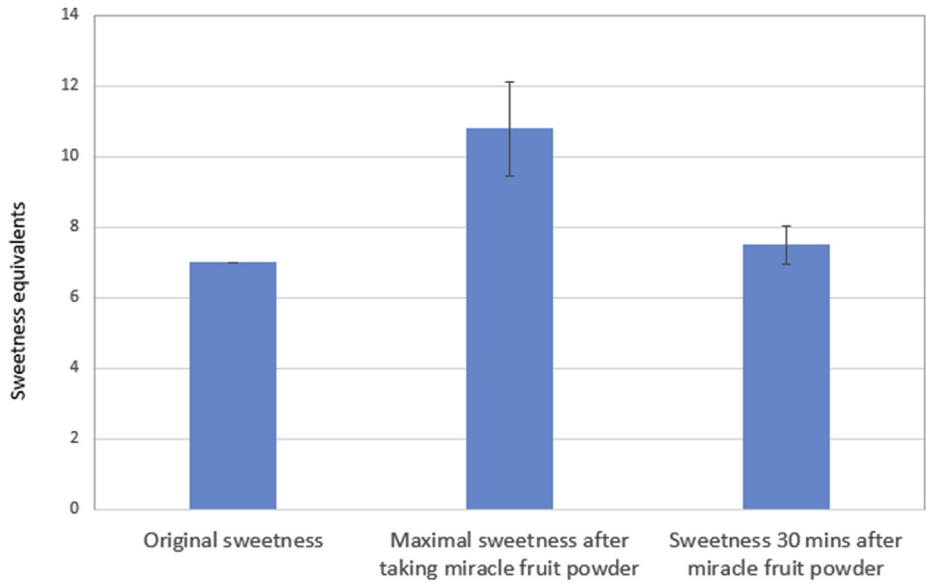

Fig. 3. Change in sweetness of lemonade with miracle fruit powder consumed by 6 panelists (mean ± SD).

from 10 to 70 kDa (Lehrer et al., 1996; Huby et al., 2000; Pekar et al., 2018) and most major allergenic proteins, such as the peanut allergens Ara h 1 and Ara h 3, have at least 23 and 4 linear epitopes, respectively, equivalent to a molecular weight of at least 60 kDa (Huby et al., 2000). Therefore, considering that the miraculin protein is digested into small peptides with molecular weights less than 4 kDa, it is not expected that the resulting small peptides would contain the epitopes required for eliciting an allergenic reaction.

The allergenic potential of miraculin and its likelihood to cause cross-reaction was further evaluated through a search of 8 segments of identical contiguous amino acids, as recommended by FAO/WHO (2001). A search of the AllergenOnline database using an exact match for 8 contiguous amino acids did not reveal any significant similarity with any known allergens. The approach using 8 contiguous amino acids was suggested to result in a greater likelihood of identifying false positives, and do not add much value to allergenicity assessments (Silvanovich et al., 2006; Herman et al., 2009). The sequence alignment examination of miraculin with that of Kunitz trypsin inhibitor KTi2 and proteinase inhibitor revealed that the longest peptide match was "YKLVFCP" present in both proteins. Importantly, this small peptide chain of 7 contiguous amino acids was not identified in the pepsin digest mapping analysis of miraculin. As highlighted in the *in vitro* digestibility analysis this information indicates that the glycoprotein is fully digested by pepsin and would therefore not pose any allergenic risk. All other sequences had less than 7 contiguous amino acid alignments with known allergens. Proteins sharing a single short identity match (8 contiguous amino acids) does not provide sufficient evidence that would indicate the likelihood of cross-reactivity of proteins, and results typically suggest a 'chance' finding, thus, the common search for 8-amino-acid identity between novel proteins and known allergens appears to be of little additional value in assessing the potential allergenicity of novel proteins (Herman et al., 2009).

Another important aspect in food safety assessment of a novel protein includes a comparison of its primary amino acid sequence with known protein toxins (Codex Alimentarius, 2009). The results of the full-length sequence alignment search of miraculin demonstrated that it does not share homology or structural similarity to any animal venom protein or toxins or virulence factors or harbor any toxic potential, considering the low query coverage and high E-values/scores for these alignments (Pearson, 2000; Bushey et al., 2014). Sequences with alignments less than 20% identity over a 100 amino acid region are not considered homologs (Hammond et al., 2013), and may be considered to occur by chance and do not indicate structural homology (Pearson, 1996).

The results of the sensory trial demonstrated that the taste modification process through reaction at the receptor level by the miraculin glycoprotein is fast with the full effect being reached within 1–2 min and then disappearing within 30 min. The fact that the onset and disappearance of the taste modifying effect is also rapid further demonstrates that there will be no lasting effects resulting from the use of the protein at levels that are consistent with those proposed for future development. These findings are similar to those reported by Kurihara and Beidler (1969), who demonstrated that the sweetening effects of miraculin reached a maximum after being held in the mouth for about 3 min and the effects of a solution of 0.02 μM disappeared in 20 min. The strength of the taste modifying effects of miraculin was also reported to be dependent on the type of acid used, with weak acids producing a more potent taste-modifying effect than strong acids (Kurihara and Beidler, 1969).

5. Conclusions

Taking into account the guidelines of the Codex Alimentarius Commission (Codex Alimentarius, 2003; 2009), the safety of miraculin has been assessed using both an *in vitro* digestibility assay and *in silico* analyses (bioinformatics). The results of the bioinformatic analyses comparing the amino acid sequence homology between miraculin and known human allergens and toxins demonstrated that miraculin does not represent a risk of allergy or toxicity to humans and has a low risk of cross-reactivity with other allergens. This conclusion was corroborated through proteomic analysis demonstrating that the miraculin glycoprotein was fully and readily digested by pepsin in an *in vitro* digestibility assay, further confirming its lack of allergenic potential. The results from sensory analyses likewise demonstrated that the taste-modifying effect of miraculin at the concentration levels to be used in product development has a rapid onset and disappearance of the effect, indicating that the overall response is limited to the function of the protein. The return of the sensory response to normality within a short period of time further indicates that the miraculin protein has no desensitizing impact on the receptor. Overall, the results of the analyses using guidance as laid down by Codex indicate a lack of a safety concern when the protein is used to alter the sensory qualities of sour substances.

Author contributions

Shahrzad Tafazoli, Trung D. Vo and Ashley Roberts were responsible for preparation of the manuscript. All authors read and approved the content of the manuscript.

Conflict of interest and funding disclosure

Intertek Scientific & Regulatory Consultancy received financial support from Coca-Cola and Miraculex to assist in the preparation of the manuscript.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.fct.2019.110762.

References

Aalberse, R.C., 2000. Structural biology of allergens. J. Allergy Clin. Immunol. 106, 228–238. https://doi.org/10.1067/mai.2000.108434.

Alimentarius, Codex, 2003. Guideline for the Conduct of Food Safety Assessment of Foods Produced Using Recombinant-DNA Microorganisms. Codex Alimentarius Commission, Rome, Italy. CAC/GL 46-2003; Adopted in 2003, Annexes II and III adopted in 2008. http://www.fao.org/fileadmin/user_upload/gmfp/docs/CAC.GL_45_2003.pdf.

Alimentarius, Codex, 2009. Foods Derived from Modern Biotechnology, second ed. Codex Alimentarius Commission, Food and Agriculture Organization of the United Nations (FAO)/World Health Organization (WHO), Geneva, Switz. http://www.fao.org/docrep/011/a1554e/a1554e00.htm.

Asakura, T., Miyano, M., Yamashita, H., Sakurai, T., Nakajima, K., Ito, K., Misaka, T., Ishimarua, Y., Abe, K., 2011. Analysis of the interaction of food components with model lingual epithelial cells: the case of sweet proteins. Flavour Fragrance J. 26, 274–278. https://doi.org/10.1002/ffj.2073.

Astwood, J.D., Leach, J.N., Fuchs, R.L., 1996. Stability of food allergens to digestion in vitro. Nat. Biotechnol. 14, 1269–1273. https://doi.org/10.1038/nbt1096-1269.

Bannon, G.A., 2004. What makes a food protein an allergen? Curr. Allergy Asthma Rep. 4, 43–46. https://doi.org/10.1007/s11882-004-0042-0.

Bartoshuk, L.M., Gentile, R.L., Molkowitz, H.R., Meiselman, H.L., 1974. Sweet taste induced by miracle fruit (*Synsepalum dulcificum*). Physiol. Behav. 12, 449–456. https://doi.org/10.1016/0031-9384(74)90122-X.

Brouwer, J.N., Van der Wel, H., Francke, A., Henning, G.J., 1968. Miraculin, the sweetness inducing protein from miracle fruit. Nature 220, 373–374. https://doi.org/10.1038/220373a0.

Brouwer, J.N., Glaser, D., Hard Af Segerstad, C., Hellekant, G., Ninomiya, Y., Van der Wel, H., 1983. The sweetness-inducing effect of miraculin; behavioural and neurophysiological experiments in the rhesus monkey *Macaca mulatta*. J. Physiol. 337, 221–240. (As cited in: Sanematsu et al., 2016 [Ref. #3]). https://doi.org/10.1113/jphysiol.1983.sp014621.

Bushey, D.F., Bannon, G.A., Delaney, B.F., Graser, G., Hefford, M., Jiang, X., Lee, T.C., Madduri, K.M., Pariza, M., Privalle, L.S., et al., 2014. Characteristics and safety assessment of intractable proteins in genetically modified crops. Regul. Toxicol. Pharmacol. 69, 154–170. https://doi.org/10.1016/j.yrtph.2014.03.003.

Castells, M.D., Pascual, C., Esteban, M.M., Ojeda, J.A., 1986. Allergy to white potato. J. Allergy Clin. Immunol. 78, 1110–1114. https://doi.org/10.1016/0091-6749(86)90258-7.

Cordle, C.T., 2004. Soy protein allergy: incidence and relative severity. J. Nutr. 134, 1213S–1219S. https://doi.org/10.1093/jn/134.5.1213S.

Diamant, H., Hellekant, G., Zotterman, Y., 1972. In: Schneider, D. (Ed.), The Effect of Miraculin on the Taste Buds of Man, Monkey and Rat. Olfaction and Taste IV Wissenschaftliche Verlagsgesellschaft MBH, Stuttgart, Germany, pp. 241–244 (As cited in: Bartoshuk et al., 1974 [Ref. #10]).

Duhita, N., Hiwasa-Tanase, K., Yoshida, S., Ezura, H., 2009. Single-step purification of native miraculin using immobilized metal-affinity chromatography. J. Agric. Food Chem. 57, 5148–5151. https://doi.org/10.1021/jf9004065.

EFSA, 2017. Guidance on allergenicity assessment of genetically modified plants (EFSA Panel on Genetically Modified Organisms/GMO) (Question no: EFSA-Q-2014-00547). EFSA J 15 (49pp), 4862. https://www.efsa.europa.eu/en/efsajournal/pub/4862.

EuropaBio, 2014. Allergenicity Assessment Applications for Authorisation of Genetically Modified Food and Feed Products. Submitted Under Regulation (EC) No 1829/2003. EFSA Technical Meeting, Oct. 28, 2014. EuropaBio the European Association for Bioindustries, Brussels, Belgium. https://www.efsa.europa.eu/sites/default/files/event/documentset/141028-p08.pdf.

FAO/WHO, 2001. Evaluation of allergenicity of genetically modified foods. reportReport of a Joint FAO/WHO Expert Consultation on Allergenicity of Foods Derived from Biotechnology 22-25 January 2001. Food and Agriculture Organization of the United Nations (FAO), Rome, Italy/World Health Organization (WHO), Geneva, Switz. http://www.who.int/foodsafety/publications/gmo-allergenicity/en.

FARRP, 2019. AllergenOnline Version 19: Home of the FARRP Allergen Protein Database. University of Nebraska-Lincoln, Food Allergy Research and Resource Program (FARRP), Lincoln, NE. http://www.allergenonline.org/, Accessed date: 10 February 2019.

Goodman, R.E., Vieths, S., Sampson, H.A., Hill, D., Ebisawa, M., Taylor, S.L., van Ree, R., 2008. Allergenicity assessment of genetically modified crops–what makes sense? [Includes letter to the editor from Rob C Aalberse and response from Richard E Goodman et al.]. Nat. Biotechnol. 26, 73–81. (erratum, 26, 241). https://doi.org/10.1038/nbt1343.

Hammond, B., Kough, J., Herouet-Guicheney, C., Jez, J.M., 2013. Toxicological evaluation of proteins introduced into food crops [ILSI international food biotechnology committee task force on use of mammalian toxicology studies in safety assessment of GM foods]. Crit. Rev. Toxicol. 43 (Suppl. 2), 25–42. https://doi.org/10.3109/10408444.2013.842956.

Synsepalum (DC). Baill., hist. Pl. 11 (1891) 286. In: 1st Engl. ed. In: Hanelt, P. (Ed.), Mansfeld's Encyclopedia of Agriculture and Horticultural Crops (Except Ornamentals) 3. Springer-Verlag, Berlin, Germany, pp. 1660.

Hellekant, G., van der Wel, H., 1989. Taste modifiers and sweet proteins. In: Cagan, R.H. (Ed.), Neural Mechanisms in Taste. CRC Press, Boca Raton, FL, pp. 85–96.

Herman, R.A., Woolhiser, M.M., Ladics, G.S., Korjagin, V.A., Schafer, B.W., Storer, N.P., Green, S.B., Kan, L., 2007. Stability of a set of allergens and non-allergens in simulated gastric fluid. Int. J. Food Sci. Nutr. 58, 125–141. (As cited in: EuropaBio, 2014 [Ref.#5]). https://doi.org/10.1080/09637480601149640.

Herman, R.A., Song, P., Thirumalaiswamysekhar, A., 2009. Value of eight-amino-acid matches in predicting the allergenicity status of proteins: an empirical bioinformatic investigation. Clin. Mol. Allergy 7 (7pp), 9. https://doi.org/10.1186/1476-7961-7-9.

Hileman, R.E., Silvanovich, A., Goodman, R.E., Rice, E.A., Holleschak, G., Astwood, J.D., Hefle, S.L., 2002. Bioinformatic methods for allergenicity assessment using a comprehensive allergen database. Int. Arch. Allergy Immunol. 128, 280–291. https://doi.org/10.1159/000063861.

Huby, R.D., Dearman, R.J., Kimber, I., 2000. Why are some proteins allergens? Toxicol. Sci. 55, 235–246. https://doi.org/10.1093/toxsci/55.2.235.

Inglett, G., Dowling, B., Albrecht, J., Hoglan, F., 1965. Taste modifiers, taste-modifying properties of miracle fruit (*Synsepalum dulcificum*). J. Agric. Food Chem. 13, 284–287. (As cited in: Bartoshuk et al., 1974 [Ref. #19]). https://doi.org/10.1021/jf60139a026.

Ito, K., Asakura, T., Morita, Y., Nakajima, K., Koizumi, A., Shimizu-Ibuka, A., Masuda, K., Ishiquro, M., Terada, T., Maruyama, J., et al., 2007. Microbial production of sensory-active miraculin. Biochem. Biophys. Res. Commun. 360, 407–411. (As cited in: Mikasa, 2013 [Ref. #9]). https://doi.org/10.1016/j.bbrc.2007.06.064.

Jin, Y., Goodman, R.E., Tetteh, A.O., Lu, M., Tripathi, L., 2017. Bioinformatics analysis to assess potential risks of allergenicity and toxicity of HRAP and PFLP proteins in genetically modified bananas resistant to Xanthomonas wilt disease. Food Chem. Toxicol. 109, 81–89. https://doi.org/10.1016/j.fct.2017.08.024.

Jungo, F., Bougueleret, L., Xenarios, I., Poux, S., 2012. The UniProtKB/Swiss-Prot Tox-Prot program: a central hub of integrated venom protein data. Toxicon 60, 551–557. https://doi.org/10.1016/j.toxicon.2012.03.010.

Kurihara, Y., 1992. Characteristics of antisweet substances, sweet proteins, and sweetness-inducing proteins. Crit. Rev. Food Sci. Nutr. 32, 231–252. (As cited in: Sun et al., 2006 [Ref. #17]). https://doi.org/10.1080/10408399209527598.

Kurihara, K., Beidler, L.M., 1968. Taste-modifying protein from miracle fruit. Science 161, 1241–1243. https://doi.org/10.1126/science.161.3847.1241.

Kurihara, K., Beidler, L.M., 1969. Mechanism of the action of taste-modifying protein. Nature 222, 1176–1179. https://doi.org/10.1038/2221176a0.

Kuroda, M., 2012. 11. Proteins and peptides with tastes. 11.2.1.6. Miraculin. In: Hettiarachchy, N.S., Sato, K., Marshall, M.R., Kanna, A. (Eds.), Food Proteins and Peptides: Chemistry, Functionality, Interactions and Commercialization. CRC Press (Taylor & Francis Group), Boca Raton, FL, pp. 277–296 (See p. 286).

Ladics, G.S., 2018. Assessment of the potential allergenicity of genetically-engineered food crops. 2018. J. Immunotoxicol. Nov. 9, 1–11. (epub ahead of print). https://doi.org/10.1080/1547691X.2018.1533904.

Ladics, G.S., Bannon, G.A., Silvanovich, A., Cressman, R.F., 2007. Comparison of conventional FASTA identity searches with the 80 amino acid sliding window FASTA search for the elucidation of potential identities to known allergens. Mol. Nutr. Food Res. 51, 985–998. https://doi.org/10.1002/mnfr.200600231.

Ladics, G.S., Cressman, R.F., Herouet-Guicheney, C., Herman, R.A., Privalle, L., Song, P., Ward, J.M., McClain, S., 2011. Bioinformatics and the allergy assessment of agricultural biotechnology products: industry practices and recommendations. Regul. Toxicol. Pharmacol. 60, 46–53. https://doi.org/10.1016/j.yrtph.2011.02.004.

Lehrer, S.B., Horner, W.E., Reese, G., 1996. Why are some proteins allergenic? Implications for biotechnology. Crit. Rev. Food Sci. Nutr. 36, 553–564. https://doi.org/10.1080/10408399609527739.

Loh, W., Tang, M.L.K., 2018. The epidemiology of food allergy in the global context. Int. J. Environ. Res. Public Health 15 (8pp), 2043. https://doi.org/10.3390/ijerph15092043.

Misaka, T., 2013. Molecular mechanisms of the action of miraculin, a taste-modifying protein. Semin. Cell Dev. Biol. 24, 222–225. https://doi.org/10.1016/j.semcdb.2013.02.008.

Morris, J.A., 1976. Sweetening agents from natural sources. Lloydia 39, 25–38 Cited in Temussi, 2006.

Ogawa, A., Samoto, M., Takahashi, K., 2000. Soybean allergens and hypoallergenic soybean products. J. Nutr. Sci. Vitaminol. 46, 271–279.
Pearson, W.R., 1996. Effective protein sequence comparison. Methods Enzymol. 266, 227–258. https://doi.org/10.1016/S0076-6879(96)66017-0.
Pearson, W.R., 2000. Flexible sequence similarity searching with the FASTA3 program package. Methods Mol. Biol. 132, 185–219. https://doi.org/10.1385/1-59259-192-2:185.
Pekar, J., Ret, D., Untersmayr, E., 2018. Stability of allergens. Mol. Immunol. 100, 14–20. https://doi.org/10.1016/j.molimm.2018.03.017.
Roecklein, J.C., Leung, P., 1987. Sweetener [Sweet berry, miraculous-berry, miraculous fruit]. A Profile of Economic Plants. Transaction Inc., New Brunswick, NJ, pp. 412.
Sanematsu, K., Kitagawa, M., Yoshida, R., Nirasawa, S., Shigemura, N., Ninomiya, Y., 2016. Intracellular acidification is required for full activation of the sweet taste receptor by miraculin. Sci. Rep. 6 (7pp), 22807. https://doi.org/10.1038/srep22807.
Seppälä, U., Majamaa, H., Turjanmaa, K., Helin, J., Reunala, T., Kalkkinen, N., Palosuo, T., 2001. Identification of four novel potato (*Solanum tuberosum*) allergens belonging to the family of soybean trypsin inhibitors. Allergy 56, 619–626. http://doi.org/10.1034/j.1398-9995.2001.00058.x.
Silvanovich, A., Nemeth, M.A., Song, P., Herman, R., Tagliani, L., Bannon, G.A., 2006. The value of short amino acid sequence matches for prediction of protein allergenicity. Toxicol. Sci. 90, 252–258. https://doi.org/10.1093/toxsci/kfj068.
Sun, H.J., Cui, M.L., Ma, B., Ezura, H., 2006. Functional expression of the taste-modifying protein, miraculin, in transgenic lettuce. FEBS Lett. 580, 620–626. https://doi.org/10.1016/j.febslet.2005.12.080.
Taylor, S.L., Lehrer, S.B., 1996. Principles and characteristics of food allergens. Crit. Rev. Food Sci. Nutr. 36 (Suppl. l), S91–S118. https://doi.org/10.1080/10408399609527761.
Temussi, P.A., 2006. Natural sweet macromolecules: how sweet proteins work. Cell. Mol. Life Sci. 63, 1876–1888.
Theerasilp, S., Kurihara, Y., 1988. Complete purification and characterization of the taste-modifying protein, miraculin, from miracle fruit. J. Biol. Chem. 263, 11536–115539.
Theerasilp, S., Hitotsuya, H., Nakajo, S., Nakaya, K., Nakamura, Y., Kurihara, Y., 1989. Complete amino acid sequence and structure characterization of the taste-modifying protein, miraculin. J. Biol. Chem. 264, 6655–6659.
Thomas, K., Aalbers, M., Bannon, G.A., Bartels, M., Dearman, R.J., Esdaile, D.J., Fu, T.J., Glatt, C.M., Hadfield, N., Hatzos, C., et al., 2004. A multi-laboratory evaluation of a common in vitro pepsin digestion assay protocol used in assessing the safety of novel proteins. Regul. Toxicol. Pharmacol. 39, 87–98. https://doi.org/10.1016/j.yrtph.2003.11.003.
Uniprot, 2019a. Taxonomy - synsepalum dulcificum (miracle fruit) (Richadella dulcifica). In: UniProtKB [Protein Knowledgebase]. UniProt Consortium. Collaboration between European Bioinformatics Institute (EMBL-EBI), Cambridge, UK, The SIB Swiss Institute of Bioinformatics, Geneva, Switz. And the Protein Information Resource (PIR), Washington, DC, . https://www.uniprot.org/taxonomy/3743, Accessed date: 28 January 2019.
Uniprot, 2019b. Animal toxin annotation project. In: UniProtKB [Protein Knowledgebase]. UniProt Consortium. Collaboration between European Bioinformatics Institute (EMBL-EBI), Cambridge, UK, the SIB Swiss Institute of Bioinformatics, Geneva, Switz. and the Protein Information Resource (PIR), Washington, DC, . https://www.uniprot.org/program/Toxins, Accessed date: 28 January 2019.
U.S. FDA, 1977. Miracle fruit berries, concentrates, and exstracts; Denial of GRAS affirmation and determination of food additive status [Docket No. 76G-0412]. Fed. Regist. 42, 26467–26468.
U.S. FDA, 2009. Memorandum Re: 75 Day premarket notification for new dietary ingredient [miracle fruit from my M fruit LLC]. U.S. Food and Drug administration (U.S. FDA), Office of Nutrition, Labeling and Dietary Supplements, Washington, DC. NDI No. 574; Docket No. FDA-1995-S-0039; Apr. 30, 2009. https://www.regulations.gov/document?D=FDA-2009-S-0608-0049.
U.S. FDA, 2018. CFSAN adverse event reporting system (CAERS). U.S. Food and Drug administration (U.S. FDA). Center for Food Safety & Applied Nutrition (CFSAN), Silver Spring, MD (Page Last Updated: 10/16/2018). https://www.fda.gov/food/complianceenforcement/ucm494015.htm.
U.S. FDA, 2018. FDA adverse event reporting system (FAERS) public dashboard. FAERS. U.S. Food and Drug administration (U.S. FDA). Center for Food Safety & Applied Nutrition (CFSAN), Silver Spring, MD. https://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/ucm070093.htm, Accessed date: 23 July 2018.
USP, 1995. U.S. Pharmacopoeia, Twenty-Third Edition & the National Formulary, eighteenth ed. US Pharmacopoeia Convention, Inc., Twinbrook Parkway, Rockville, MD, pp. 2053 (As cited in: Thomas et al., 2004 [As: Mack Printing Co., Easton, PA]).

Table S1 Search Results of AllergenOnline Database Version 19 with Peptides Following Pepsin Digest Mapping

Peptide Sequence	Description	Full Length				
		Length	E-value	% Identity	% Similarity	Amino acid overlap
DSAPNPVLDIDGEKLRT	Putative Ani s 11-like protein pr	160	0.13	50	100	12
	Allergen Aspartic protease inhibitor	188	0.16	58.3	91.7	12
	Putative Ani s 11-like protein 2	287	0.24	50	100	12
	Putative Ani s 11 allergen precur	307	0.26	50	100	12
DSAPNPVLDIDGEKLRTGTNY	Allergen Aspartic protease inhibitor	188	0.023	44.4	83.3	18
	Putative Ani s 11-like protein pr	160	0.28	50	100	12
	Putative Ani s 11-like protein 2	287	0.39	50	100	12
	Putative Ani s 11 allergen precur	307	0.41	50	100	12
	Allergen proteinase inhibitor	221	0.6	40	70	20
DSAPNPVLDIDGEKLRTGTNYY	Allergen Aspartic protease inhibitor	188	0.037	44.4	83.3	18
	Putative Ani s 11-like protein pr	160	0.4	50	100	12
	Putative Ani s 11-like protein 2	287	0.58	50	100	12
	Putative Ani s 11 allergen precur	307	0.6	50	100	12
	Putative Kunitz trypsin inhibitor KTi	203	0.62	50	75	16
	Putative Kunitz trypsin inhibitor KTi	204	0.62	50	75	16
	Putative Kunitz trypsin inhibitor	208	0.63	43.8	81.3	16
	Allergen proteinase inhibitor	221	0.86	40	70	20
DSAPNPVLDIDGEKLRTGTNYYIVPVLRD	Putative Kunitz trypsin inhibitor	208	0.0013	45.5	86.4	22
	Putative Kunitz trypsin inhibitor KTi	203	0.0016	50	81.8	22
	Putative Kunitz trypsin inhibitor KTi	204	0.0016	50	81.8	22
	Allergen Aspartic protease inhibitor	188	0.0041	40	76	25
	Allergen proteinase inhibitor	221	0.34	39.1	69.6	23
	Putative Kunitz trypsin inhibitor; KT	216	0.7	50	83.3	18
	Putative trypsin inhibitor subtype B	217	0.71	50	77.8	18
	Putative trypsin inhibitor subtype A	217	0.71	50	83.3	18
DSAPNPVLDIDGEKLRT	Putative Ani s 11-like protein pr	160	0.13	50	100	12
	Allergen Aspartic protease inhibitor	188	0.16	58.3	91.7	12
	Putative Ani s 11-like protein 2	287	0.24	50	100	12
	Putative Ani s 11 allergen precur	307	0.26	50	100	12

Table S1 Search Results of AllergenOnline Database Version 19 with Peptides Following Pepsin Digest Mapping

Peptide Sequence	Description	Full Length				
		Length	E-value	% Identity	% Similarity	Amino acid overlap
DSAPNPVLDIDGEKLRTGTNY	Allergen Aspartic protease inhibitor	188	0.023	44.4	83.3	18
	Putative Ani s 11-like protein pr	160	0.28	50	100	12
	Putative Ani s 11-like protein 2	287	0.39	50	100	12
	Putative Ani s 11 allergen precur	307	0.41	50	100	12
	Allergen proteinase inhibitor	221	0.6	40	70	20
DSAPNPVLDIDGEKLRTGTNYY	Allergen Aspartic protease inhibitor	188	0.037	44.4	83.3	18
	Putative Ani s 11-like protein pr	160	0.4	50	100	12
	Putative Ani s 11-like protein 2	287	0.58	50	100	12
	Putative Ani s 11 allergen precur	307	0.6	50	100	12
	Putative Kunitz trypsin inhibitor KTi	203	0.62	50	75	16
	Putative Kunitz trypsin inhibitor KTi	204	0.62	50	75	16
	Putative Kunitz trypsin inhibitor	208	0.63	43.8	81.3	16
	Allergen proteinase inhibitor	221	0.86	40	70	20
DSAPNPVLDIDGEKLRTGTNYYIVPVLRD	Putative Kunitz trypsin inhibitor	208	0.0013	45.5	86.4	22
	Putative Kunitz trypsin inhibitor KTi	203	0.0016	50	81.8	22
	Putative Kunitz trypsin inhibitor KTi	204	0.0016	50	81.8	22
	Allergen Aspartic protease inhibitor	188	0.0041	40	76	25
	Allergen proteinase inhibitor	221	0.34	39.1	69.6	23
	Putative Kunitz trypsin inhibitor; KT	216	0.7	50	83.3	18
	Putative trypsin inhibitor subtype B	217	0.71	50	77.8	18
	Putative trypsin inhibitor subtype A	217	0.71	50	83.3	18
SAPNPVLDIDGEKLRTGTNY	Allergen Aspartic protease inhibitor	188	0.009	44.4	83.3	18
	Putative Ani s 11-like protein pr	160	0.12	50	100	12
	Putative Ani s 11-like protein 2	287	0.22	50	100	12
	Putative Ani s 11 allergen precur	307	0.24	50	100	12
	Allergen proteinase inhibitor	221	0.32	40	70	20
	Allergen cysteine proteinase inhibitor	186	0.69	53.3	60	15
	Allergen Cysteine protease inhibitor	222	0.82	53.3	60	15
SAPNPVLDIDGEKLRTGTNYY	Allergen Aspartic protease inhibitor	188	0.021	44.4	83.3	18

Table S1 Search Results of AllergenOnline Database Version 19 with Peptides Following Pepsin Digest Mapping

Peptide Sequence	Description	Full Length				
		Length	E-value	% Identity	% Similarity	Amino acid overlap
	Putative Ani s 11-like protein pr	160	0.25	50	100	12
	Putative Ani s 11-like protein 2	287	0.36	50	100	12
	Putative Ani s 11 allergen precur	307	0.38	50	100	12
	Putative Kunitz trypsin inhibitor KTi	203	0.39	50	75	16
	Putative Kunitz trypsin inhibitor KTi	204	0.39	50	75	16
	Putative Kunitz trypsin inhibitor	208	0.4	43.8	81.3	16
	Allergen proteinase inhibitor	221	0.55	40	70	20
	Putative Kunitz trypsin inhibitor; KT	216	0.73	50	81.3	16
	Putative trypsin inhibitor subtype A	217	0.73	50	81.3	16
	Putative trypsin inhibitor subtype B	217	0.73	50	75	16
APNPVLDIDGEKLRT	Putative Ani s 11-like protein pr	160	0.072	50	100	12
	Allergen Aspartic protease inhibitor	188	0.085	58.3	91.7	12
	Putative Ani s 11-like protein 2	287	0.13	50	100	12
	Putative Ani s 11 allergen precur	307	0.14	50	100	12
APNPVLDIDGEKLRTGTNY	Allergen Aspartic protease inhibitor	188	0.0076	44.4	83.3	18
	Putative Ani s 11-like protein pr	160	0.11	50	100	12
	Putative Ani s 11-like protein 2	287	0.19	50	100	12
	Putative Ani s 11 allergen precur	307	0.21	50	100	12
	Allergen proteinase inhibitor	221	0.52	43.8	75	16
	Allergen cysteine proteinase inhibit	186	0.6	53.3	60	15
	Allergen Cysteine protease inhibito	222	0.72	53.3	60	15
APNPVLDIDGEKLRTGTNYY	Allergen Aspartic protease inhibitor	188	0.013	44.4	83.3	18
	Putative Ani s 11-like protein pr	160	0.16	50	100	12
	Putative Kunitz trypsin inhibitor KTi	203	0.28	50	75	16
	Putative Kunitz trypsin inhibitor KTi	204	0.28	50	75	16
	Putative Kunitz trypsin inhibitor	208	0.29	43.8	81.3	16
	Putative Ani s 11-like protein 2	287	0.29	50	100	12
	Putative Ani s 11 allergen precur	307	0.31	50	100	12
	Putative Kunitz trypsin inhibitor; KT	216	0.54	50	81.3	16

Table S1 Search Results of AllergenOnline Database Version 19 with Peptides Following Pepsin Digest Mapping

Peptide Sequence	Description	Full Length				
		Length	E-value	% Identity	% Similarity	Amino acid overlap
	Putative trypsin inhibitor subtype B	217	0.55	50	75	16
	Putative trypsin inhibitor subtype A	217	0.55	50	81.3	16
	Allergen proteinase inhibitor	221	0.75	43.8	75	16
	Allergen cysteine proteinase inhibit	186	0.86	53.3	60	15
	Allergen RecName: Full=Glutathione	206	0.95	41.2	70.6	17
NPVLDIDGEKLRTGTNYY	Allergen Aspartic protease inhibitor	188	0.22	43.8	81.3	16
	Putative Kunitz trypsin inhibitor KTi	203	0.23	50	75	16
	Putative Kunitz trypsin inhibitor KTi	204	0.24	50	75	16
	Putative Kunitz trypsin inhibitor	208	0.24	43.8	81.3	16
	Putative Kunitz trypsin inhibitor; KT	216	0.46	50	81.3	16
	Putative trypsin inhibitor subtype B	217	0.46	50	75	16
	Putative trypsin inhibitor subtype A	217	0.46	50	81.3	16
	Allergen proteinase inhibitor	221	0.63	43.8	75	16
	Allergen cysteine proteinase inhibit	186	0.72	53.3	60	15
	Allergen RecName: Full=Glutathione	206	0.8	41.2	70.6	17
	Allergen Cysteine protease inhibito	222	0.86	53.3	60	15
LDIDGEKLRT	No sequences with E(< 1.000000)					
LDIDGEKLRTGTNY	No sequences with E(< 1.000000)					
LDIDGEKLRTGTNYY	Putative Kunitz trypsin inhibitor KTi	203	0.79	46.7	73.3	15
	Putative Kunitz trypsin inhibitor KTi	204	0.79	46.7	73.3	15
	Putative Kunitz trypsin inhibitor	208	0.81	40	80	15
LDIDGEKLRTGTNYYIVPVLRD	Putative Kunitz trypsin inhibitor	208	0.00067	42.9	85.7	21
	Putative Kunitz trypsin inhibitor KTi	203	0.00087	47.6	81	21
	Putative Kunitz trypsin inhibitor KTi	204	0.00088	47.6	81	21
	Allergen Aspartic protease inhibitor	188	0.69	33.3	71.4	21
	Putative Kunitz trypsin inhibitor; KT	216	0.77	47.1	82.4	17
	Putative trypsin inhibitor subtype B	217	0.77	47.1	76.5	17
	Putative trypsin inhibitor subtype A	217	0.77	47.1	82.4	17
DIDGEKLRT	No sequences with E(< 1.000000)					

Table S1 Search Results of AllergenOnline Database Version 19 with Peptides Following Pepsin Digest Mapping

Peptide Sequence	Description	Full Length				
		Length	E-value	% Identity	% Similarity	Amino acid overlap
DIDGEKLRTGTNYY	Putative Kunitz trypsin inhibitor	208	0.84	42.9	78.6	14
IDGEKLRTGTNY	No sequences with E(< 1.000000)					
IDGEKLRTGTNYY	No sequences with E(< 1.000000)					
NYYIVPVLRD	Putative Kunitz trypsin inhibitor	208	0.37	62.5	100	8
	Putative Kunitz trypsin inhibitor KTi	203	0.51	62.5	100	8
	Putative Kunitz trypsin inhibitor KTi	204	0.51	62.5	100	8
NYYIVPVLRDHGGGLT	Putative Kunitz trypsin inhibitor	208	0.0032	57.1	92.9	14
	Putative Kunitz trypsin inhibitor KTi	203	0.0041	57.1	92.9	14
	Putative Kunitz trypsin inhibitor KTi	204	0.038	50	92.9	14
NYYIVPVLRDHGGGLTV	Putative Kunitz trypsin inhibitor KTi	203	0.0027	56.3	87.5	16
	Putative Kunitz trypsin inhibitor	208	0.0061	57.1	92.9	14
	Putative Kunitz trypsin inhibitor KTi	204	0.066	50	92.9	14
NYYIVPVLRDHGGGLTVSA	Putative Kunitz trypsin inhibitor KTi	203	0.0025	50	88.9	18
	Putative Kunitz trypsin inhibitor	208	0.0075	57.1	92.9	14
	Putative Kunitz trypsin inhibitor KTi	204	0.08	50	92.9	14
YYIVPVLRD	Putative Kunitz trypsin inhibitor	208	0.18	62.5	100	8
	Putative Kunitz trypsin inhibitor KTi	203	0.25	62.5	100	8
	Putative Kunitz trypsin inhibitor KTi	204	0.25	62.5	100	8
YYIVPVLRDHGGGLT	Putative Kunitz trypsin inhibitor	208	0.0021	57.1	92.9	14
	Putative Kunitz trypsin inhibitor KTi	203	0.0027	57.1	92.9	14
	Putative Kunitz trypsin inhibitor KTi	204	0.027	50	92.9	14
YYIVPVLRDHGGGLTV	Putative Kunitz trypsin inhibitor KTi	203	0.0019	56.3	87.5	16
	Putative Kunitz trypsin inhibitor	208	0.0045	57.1	92.9	14
	Putative Kunitz trypsin inhibitor KTi	204	0.051	50	92.9	14
YYIVPVLRDHGGGLTVS	Putative Kunitz trypsin inhibitor KTi	203	0.0025	56.3	87.5	16
	Putative Kunitz trypsin inhibitor	208	0.0056	57.1	92.9	14
	Putative Kunitz trypsin inhibitor KTi	204	0.062	50	92.9	14
YYIVPVLRDHGGGLTVSA	Putative Kunitz trypsin inhibitor KTi	203	0.0018	50	88.9	18
	Putative Kunitz trypsin inhibitor	208	0.0055	57.1	92.9	14

Table S1 Search Results of AllergenOnline Database Version 19 with Peptides Following Pepsin Digest Mapping

Peptide Sequence	Description	Full Length				
		Length	E-value	% Identity	% Similarity	Amino acid overlap
	Putative Kunitz trypsin inhibitor KTi	204	0.061	50	92.9	14
YYIVPVLRDHGGGLTVSAT	Putative Kunitz trypsin inhibitor KTi	203	0.0011	52.6	89.5	19
	Putative Kunitz trypsin inhibitor	208	0.0051	47.4	78.9	19
	Putative Kunitz trypsin inhibitor KTi	204	0.066	42.1	84.2	19
IVPVLRDHGGGLT	Putative Kunitz trypsin inhibitor	208	0.11	50	91.7	12
	Putative Kunitz trypsin inhibitor KTi	203	0.15	50	91.7	12
VVQTRKEVDHDRPLAFF	No sequences with E(< 1.000000)					
QTRKEVDHDRPLAFF	No sequences with E(< 1.000000)					
QTRKEVDHDRPLAFFPENPKED	No sequences with E(< 1.000000)					
QTRKEVDHDRPLAFFPENPKEDVVRVSTDLN	No sequences with E(< 1.000000)					
TRKEVDHDRPLAFF	No sequences with E(< 1.000000)					
LAFFPENPKE	No sequences with E(< 1.000000)					
LAFFPENPKEDVVRVSTDLN	No sequences with E(< 1.000000)					
FFPENPKED	No sequences with E(< 1.000000)					
FFPENPKEDVV	No sequences with E(< 1.000000)					
FFPENPKEDVVRVS	No sequences with E(< 1.000000)					
FFPENPKEDVVRVST	No sequences with E(< 1.000000)					
FFPENPKEDVVRVSTDL	No sequences with E(< 1.000000)					
FFPENPKEDVVRVSTDLN	No sequences with E(< 1.000000)					
FFPENPKEDVVRVSTDLNI	No sequences with E(< 1.000000)					
FFPENPKEDVVRVSTDLNINFS	No sequences with E(< 1.000000)					
FPENPKEDVVRVSTDL	No sequences with E(< 1.000000)					
FPENPKEDVVRVSTDLN	No sequences with E(< 1.000000)					
FPENPKEDVVRVSTDLNINFS	No sequences with E(< 1.000000)					
EDVVRVSTDL	No sequences with E(< 1.000000)					
EDVVRVSTDLN	No sequences with E(< 1.000000)					
EDVVRVSTDLNI	No sequences with E(< 1.000000)					
EDVVRVSTDLNINFS	No sequences with E(< 1.000000)					
VVRVSTDLNINFS	No sequences with E(< 1.000000)					

Table S1 Search Results of AllergenOnline Database Version 19 with Peptides Following Pepsin Digest Mapping

Peptide Sequence	Description	Full Length				
		Length	E-value	% Identity	% Similarity	Amino acid overlap
RWTSSTVWRLDKYDESTGQYF	Allergen pollen allergen	263	0.62	36	60	25
	Putative group I pollen allergen Po	263	0.62	42.1	63.2	19
	Allergen proteinase inhibitor	221	0.65	46.7	66.7	15
	Allergen pollen allergen	252	0.75	42.1	68.4	19
	Allergen pollen allergen Lolium per	263	0.79	42.1	68.4	19
	Allergen Pollen allergen Lol p 1 prec	263	0.79	42.1	68.4	19
	Allergen Chain A, Crystal Structure	241	0.92	36.8	68.4	19
	Putative protein with incomplete sig	248	0.95	42.1	68.4	19
STVWRLDKYDESTGQYF	Allergen pollen allergen	252	0.44	50	71.4	14
	Allergen pollen allergen	263	0.46	50	71.4	14
	Allergen Pollen allergen Lol p 1 prec	263	0.46	50	71.4	14
	Putative group I pollen allergen Po	263	0.46	40	73.3	15
	Allergen pollen allergen Lolium per	263	0.46	50	71.4	14
	Allergen Chain A, Crystal Structure	241	0.55	42.9	71.4	14
	Putative protein with incomplete sig	248	0.56	50	71.4	14
	Allergen Phl p I allergen Phleum pra	263	0.6	42.9	71.4	14
	Putative major group I allergen Hol	263	0.6	50	71.4	14
	Putative allergen Hol-II Holcus lana	265	0.6	50	71.4	14
	Putative Major pollen allergen Pha a	269	0.61	42.9	71.4	14
	Allergen proteinase inhibitor	221	0.66	46.2	69.2	13
	Allergen group 1 allergen Dac g 1.0	240	0.71	42.9	71.4	14
	Allergen Group 1 allergen-like Dac	264	0.78	42.9	71.4	14
TVWRLDKYDESTGQYF	Allergen pollen allergen	252	0.44	50	71.4	14
	Allergen pollen allergen	263	0.46	50	71.4	14
	Allergen Pollen allergen Lol p 1 prec	263	0.46	50	71.4	14
	Putative group I pollen allergen Po	263	0.46	40	73.3	15
	Allergen pollen allergen Lolium per	263	0.46	50	71.4	14
	Allergen Chain A, Crystal Structure	241	0.55	42.9	71.4	14
	Putative protein with incomplete sig	248	0.56	50	71.4	14

Table S1 Search Results of AllergenOnline Database Version 19 with Peptides Following Pepsin Digest Mapping

Peptide Sequence	Description	Full Length				
		Length	E-value	% Identity	% Similarity	Amino acid overlap
	Allergen Phl p I allergen Phleum pra	263	0.6	42.9	71.4	14
	Putative major group I allergen Hol	263	0.6	50	71.4	14
	Putative allergen Hol-II Holcus lana	265	0.6	50	71.4	14
	Putative Major pollen allergen Pha a	269	0.61	42.9	71.4	14
	Allergen proteinase inhibitor	221	0.65	46.2	69.2	13
	Allergen group 1 allergen Dac g 1.0	240	0.71	42.9	71.4	14
	Allergen Group 1 allergen-like Dac	264	0.78	42.9	71.4	14
WRLDKYDESTGQYF	No sequences with E(< 1.000000)					
LDKYDESTGQYFV	No sequences with E(< 1.000000)					
FVTIGGVKGNPGPET	No sequences with E(< 1.000000)					
FVTIGGVKGNPGPETISSW	No sequences with E(< 1.000000)					
FVTIGGVKGNPGPETISSWF	No sequences with E(< 1.000000)					
TIGGVKGNPGPETISSWF	No sequences with E(< 1.000000)					
YFVTIGGVKGNPGPET	No sequences with E(< 1.000000)					
YFVTIGGVKGNPGPETI	No sequences with E(< 1.000000)					
YFVTIGGVKGNPGPETISS	No sequences with E(< 1.000000)					
YFVTIGGVKGNPGPETISSW	No sequences with E(< 1.000000)					
YFVTIGGVKGNPGPETISSWF	No sequences with E(< 1.000000)					
YFVTIGGVKGNPGPETISSWFK	No sequences with E(< 1.000000)					
YFVTIGGVKGNPGPETISSWFKIEEF	No sequences with E(< 1.000000)					
SWFKIEEF	No sequences with E(< 1.000000)					
SWFKIEEFC	No sequences with E(< 1.000000)					
GIYIDQKGRRRLALS	Putative Kunitz trypsin inhibitor KTi	203	0.047	66.7	80	15
	Putative Kunitz trypsin inhibitor KTi	204	0.047	66.7	80	15
	Allergen Aspartic protease inhibitor	188	0.25	64.3	85.7	14
	Allergen proteinase inhibitor	221	0.4	63.6	90.9	11
GIYIDQKGRRRLALSDKPFAF	Putative Kunitz trypsin inhibitor KTi	203	0.0089	60	80	20
	Putative Kunitz trypsin inhibitor KTi	204	0.0089	60	80	20
	Allergen proteinase inhibitor	221	0.079	52.9	88.2	17

Table S1 Search Results of AllergenOnline Database Version 19 with Peptides Following Pepsin Digest Mapping

Peptide Sequence	Description	Full Length				
		Length	E-value	% Identity	% Similarity	Amino acid overlap
	Allergen Aspartic protease inhibitor	188	0.21	50	85	20
	Putative Kunitz trypsin inhibitor; KT	216	0.45	52.4	76.2	21
	Putative trypsin inhibitor subtype B	217	0.46	52.4	76.2	21
	Putative trypsin inhibitor subtype A	217	0.46	52.4	76.2	21
	Putative allergen Pen m 2 Fennero	53	0.68	43.8	87.5	16
	Putative allergen Pen m 2 Fennero	53	0.68	43.8	87.5	16
GIYIDQKGRRRLALSDKPFAFE	Putative Kunitz trypsin inhibitor KTi	204	0.0011	56.5	78.3	23
	Putative Kunitz trypsin inhibitor KTi	203	0.0038	52.2	78.3	23
	Allergen proteinase inhibitor	221	0.051	52.9	88.2	17
	Allergen Aspartic protease inhibitor	188	0.14	50	85	20
	Putative Kunitz trypsin inhibitor; KT	216	0.24	45.8	75	24
	Putative trypsin inhibitor subtype B	217	0.24	45.8	75	24
	Putative trypsin inhibitor subtype A	217	0.24	45.8	75	24
	Putative allergen Pen m 2 Fennero	53	0.55	43.8	87.5	16
	Putative allergen Pen m 2 Fennero	53	0.55	43.8	87.5	16
YIDQKGRRRLALSDKPFAFE	Putative Kunitz trypsin inhibitor KTi	204	0.057	55	80	20
	Allergen proteinase inhibitor	221	0.062	52.9	88.2	17
	Putative Kunitz trypsin inhibitor KTi	203	0.19	50	80	20
	Allergen Aspartic protease inhibitor	188	0.6	53.3	93.3	15
LSDKPFAFE	No sequences with E(< 1.000000)					
EFNKTVYF	No sequences with E(< 1.000000)					