



OPEN ACCESS

EDITED BY

Weicheng Hu,
Yangzhou University, China

REVIEWED BY

Darren Nicholas Nesbeth,
University College London, United Kingdom
Gisele Monteiro,
University of São Paulo, Brazil
Brian Effer Roldan,
University of La Frontera, Chile

*CORRESPONDENCE

Carrie-Anne Malinczak
✉ carrie@myhelaina.com

RECEIVED 31 January 2024

ACCEPTED 03 July 2024

PUBLISHED 24 July 2024

CITATION

Anaya Y, Rosario Martinez R, Goodman RE, Johnson P, Vajpeyi S, Lu X, Peterson R, Weyers SM, Breen B, Newsham K, Scottoline B, Clark AJ and Malinczak C-A (2024) Evaluation of the potential food allergy risks of human lactoferrin expressed in *Komagataella phaffii*. *Front. Immunol.* 15:1380028. doi: 10.3389/fimmu.2024.1380028

COPYRIGHT

© 2024 Anaya, Rosario Martinez, Goodman, Johnson, Vajpeyi, Lu, Peterson, Weyers, Breen, Newsham, Scottoline, Clark and Malinczak. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Evaluation of the potential food allergy risks of human lactoferrin expressed in *Komagataella phaffii*

Yanisa Anaya¹, Raysa Rosario Martinez¹, Richard E. Goodman², Philip Johnson³, Shashwat Vajpeyi⁴, Xiaoning Lu¹, Ross Peterson⁵, Sarah M. Weyers⁶, Bella Breen⁴, Kahler Newsham⁴, Brian Scottoline⁷, Anthony J. Clark^{4,5} and Carrie-Anne Malinczak^{1*}

¹Nutritional Biology & Safety, Helaina, Inc, New York, NY, United States, ²RE Goodman Consulting, Lincoln, NE, United States, ³University of Nebraska-Lincoln, Lincoln, NE, United States, ⁴Late Stage R&D, Helaina, Inc, New York, NY, United States, ⁵Regulatory Affairs, Helaina, Inc, New York, NY, United States, ⁶Independent Scholar, San Diego, CA, United States, ⁷Department of Pediatrics, Oregon Health & Science University, Portland, OR, United States

Introduction: Prior to the introduction of novel food ingredients into the food supply, safety risk assessments are required, and numerous prediction models have been developed and validated to evaluate safety.

Methods: The allergenic risk potential of Helaina recombinant human lactoferrin (rhLF, EfferTM), produced in *Komagataella phaffii* (*K. phaffii*) was assessed by literature search, bioinformatics sequence comparisons to known allergens, glycan allergenicity assessment, and a simulated pepsin digestion model.

Results: The literature search identified no allergenic risk for Helaina rhLF, *K. phaffii*, or its glycans. Bioinformatics search strategies showed no significant risk for cross-reactivity or allergenicity between rhLF or the 36 residual host proteins and known human allergens. Helaina rhLF was also rapidly digested in simulated gastric fluid and its digestibility profile was comparable to human milk lactoferrin (hmLF), further demonstrating a low allergenic risk and similarity to the hmLF protein.

Conclusion: Collectively, these results demonstrate a low allergenic risk potential of Helaina rhLF and do not indicate the need for further clinical testing or serum IgE binding to evaluate Helaina rhLF for risk of food allergy prior to introduction into the food supply.

KEYWORDS

Komagataella phaffii, *Pichia pastoris*, human lactoferrin, allergenicity, recombinant protein, codex, glycosylation

1 Introduction

Lactoferrin (LF) is an 80-kDa, iron-binding glycoprotein found in exocrine secretions of humans and many other mammals, predominantly milk and colostrum in humans and in colostrum in ruminants. The roles of this bioactive protein are multifunctional, including linking innate and adaptive immune responses, iron homeostasis, and many other functional benefits (1). Lactoferrin is produced in the epithelial cells of mammary, salivary, and lacrimal glands as well as in neutrophils. Composed of 691 amino acids, human milk lactoferrin (hmLF) plays several critical roles, including but not limited to the neonatal immune system and iron homeostasis (1–3). Bovine lactoferrin (bLF) is a 689-amino acid polypeptide sharing 69% sequence homology with hmLF and shares a number of beneficial properties with hmLF. Bovine LF is generally recognized as safe (GRAS) (US FDA, GRN 000067, 2001; US FDA, GRN 000077, 2001c; US FDA, GRN 130, 2003; US FDA, GRN 000423, 2012; US FDA, GRN 000464, 2013a; US FDA, GRN 000465, 2013b; US FDA, GRN 000669, 2016) (4–10) for use in cow's milk-based term infant formula, sports and functional foods, ice cream, powdered milk, yogurt chewing gum, and as an antimicrobial agent (11). In addition to the intended uses described above, bLF is used in cosmetics and dietary supplements. While bLF is effective as a dietary ingredient, its bioavailability and subsequent functions may differ from human LF (hLF)—for example, *in vitro* studies have shown that hLF digests more slowly and has a stronger affinity for the human LF receptor in the small intestine, indicating that it may be more bioavailable than bLF (12–14). Thus, utilizing hLF as a substitute for bLF could be more appropriate. However, isolating LF from human milk is impractical for this purpose because it is found in relatively low concentrations, resulting in high production costs as well as the ethical challenges of using human milk to isolate LF rather than to feed infants. Strategies to produce cost-effective, sustainable, and safe hLF are warranted to develop successful commercial production.

Recombinant human lactoferrin (rhLF) has been produced in several biological expression systems, including but not limited to rice (15), transgenic cows (16), Chinese hamster ovary (CHO) cells (17), and yeast, namely, *Komagataella phaffii* (*K. phaffii*) (previously known as *Pichia pastoris*) (18). Introduced as a commercial biological expression system in the 1980s (19), *K. phaffii* is a methylotrophic yeast that remains one of the most popular fermentation systems for recombinant protein synthesis in molecular biology due to its ease of use, standardized protocols, scalability, and capabilities to appropriately fold and secrete desired proteins (20). Food ingredients available for specific intended uses in the US food supply that use *K. phaffii* for protein production include β -lactoglobulin (21) (Remilk), egg white protein (22) (The EVERY company), myoglobin (23) (Motif FoodWorks), and soy leghemoglobin (24) (Impossible Foods). Furthermore, *K. phaffii* is an approved protein source for broiler chickens at up to 10% of the total feed (25).

Helaina Inc. (New York City, New York, USA) successfully developed rhLF derived from a glycoengineered yeast that is substantively similar in structure to hmLF (26). Helaina rhLF (EfferTM) was engineered according to the UniProt protein ID

for human LF, TRFL_HUMAN (P02788). A proprietary technology, involving the disruption of an endogenous glycosyltransferase gene (OCH1) (27) and a stepwise introduction of heterologous glycosylation enzymes, enables a modified strain of *K. phaffii* (GS-115) to produce rhLF (28). It has been well established that hmLF is a glycoprotein mainly having complex N-linked glycans on three sites, including Asn-156, Asn-497, and Asn-642 (29). Helaina rhLF possesses glycans on the exact same three sites with primarily oligomannose N-glycans having five to 14 mannose residues. A thorough comparison of Helaina rhLF glycosylation profile compared to hmLF is described in detail elsewhere (26).

Evaluating the food allergy risk of novel proteins prior to their introduction into the food supply is a best practice. While there is no definitive, single factor that serves as an indicator of protein safety, the Codex Alimentarius Commission's multistep approach provides a framework by which to evaluate potential allergenic risk and the safety profile of novel food proteins in genetically modified microorganisms (30). However, it is important to recognize that Codex was not intended to evaluate novel foods that can have tens to hundreds of new proteins that may be evolutionarily conserved in multiple taxa and often of low abundance in food products. Codex guidelines recommend reviewing the safety of the protein gene source and history of exposure, comparing the amino acid sequence similarity of the novel protein to known human allergens, and testing the stability of the protein to digestion by pepsin to provide an overall assessment on allergenic potential. Although other methods (e.g., animal models and targeted sera screening) are being developed to test for food allergy, such approaches have not been thoroughly evaluated or validated for predicting protein allergenicity (31–33).

Codex guidelines have been applied to evaluate the potential allergenicity of other novel proteins expressed in a recombinant host and residual host proteins that remain in food fractions, such as rhLF produced in transgenic cows (34) and soybean leghemoglobin (35). However, the guidelines have yet to be applied to rhLF expressed in *K. phaffii*; therefore, the main objective of this work was to apply Codex criteria to evaluate the potential allergenic risk of Helaina rhLF. Additionally, an *in vitro* pepsin digestion model developed by Astwood et al. (36) and refined by Thomas et al. (37) and Ofori-Anti et al. (38), which is widely used to simulate the digestion of proteins to give an indication of risk predictions (e.g., allergenicity), was performed. Although not part of the Codex weight-of-evidence approach, allergenicity based on the glycan signature of Helaina rhLF was also assessed due to the glycoprotein nature of hLF.

2 Materials and methods

2.1 Scientific literature review—evidence of allergenicity

The National Center for Biotechnology Information (NCBI) PubMed[®] literature database was searched on October 26, 2023 for

peer-reviewed evidence of allergenicity to hLF as well as proteins from *K. phaffii* and from its previous name, *Pichia pastoris*, using keyword limits. The searches used combinations of species names “*Komagataella phaffii*” and “*Pichia pastoris*” with either “clinical allergy” or “food allergy” or “IgE binding”.

The potential for cross-reactivity against specific glycans introduced by the yeast expression system (i.e., mannose) was also assessed to determine if a risk of allergenicity against asparagine-linked glycans exists. This risk includes identifying the presence or absence of cross-reactive carbohydrate determinants (CCD) that are protein-linked carbohydrate structures (i.e., glycans) involved in the phenomenon of cross-reactivity of sera from allergic patients toward a wide range of allergens typically found in plants and insects.

2.2 Allergenicity assessment of relatively abundant residual proteins from *Komagataella phaffii*

2.2.1 Sequence of rhLF and residual *K. phaffii* host proteins

Three representative production batches of Helaina rhLF (UniProt P02788) were analyzed by liquid chromatography–tandem mass spectrometry (LC–MS/MS). Three replicate analyses of each production batch were performed.

The LF samples were solubilized in a 50-mL Tris HCl buffer (pH 8.0). The samples were diluted in 25 mM ammonium carbonate and 10 mM dithiothreitol (DTT) at 60°C, followed by the addition of 50 mM iodoacetamide (IAA) at room temperature. The diluted samples (10 µg) were digested with trypsin enzyme (Promega, Madison, WI, USA) at 37°C for 18 h, quenched with formic acid, and desalted using SPE.

Peptide digests were analyzed by nano-flow LC–MS/MS with a Waters M-class high-pressure liquid chromatography (HPLC) system interfaced to an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). Peptides were loaded onto a trapping column and eluted over a 75-µm analytical column at 350 nL/min; both columns were packed with Luna C18 resin (Phenomenex, Torrance, CA, USA). A 4-h gradient was applied. The mass spectrometer was operated in data-dependent mode, with MS and MS/MS performed at resolutions of 60,000 full-width at half-maximum (FWHM) and 15,000 FWHM, respectively. Advanced peak determination was enabled. The instrument was run using a 3-s cycle for MS and MS/MS. Only the peptides having two to seven charges were selected for MS/MS and were excluded for 25 s from repeated MS/MS. The fragments were generated by higher-energy collisional dissociation (HCD) with a fixed normalized collision energy at 30. Orbitrap was used as the detector with a resolution of 15,000, maximum injection time of 22 m, and automatic gain control target of 50,000, respectively.

Data-dependent acquisition mass-spectral data was analyzed using PEAKS version 8.5 (39). The sequence dataset used for the analysis was the publicly available *K. phaffii* set from UniProt (Taxon ID 4922, containing 5,257 entries) appended with the amino acid sequence of hmLF (UniProt P02788) and mannosidase from *Trichoderma reesei* (*T.*

reesei) (UniProt G0RBB5). The following data analysis parameters were used: no mis-cleavages, carbamidomethylation as a fixed modification, oxidation of methionine as a variable modification, parent mass error tolerance of 5 ppm, fragment mass error tolerance of 0.02 Da, false discovery rate (FDR) of 1%, and charge states between 2 and 4 were accepted. Only proteins identified via two or more peptides were considered for further analysis. The resultant lists of protein IDs (one for each replicate analysis for each batch) were combined for use in comparison to allergenic proteins in the www.AllergenOnline.org version 22 database, as described below.

After protein identification, label-free quantitation was performed using peptide ion (MS1) abundance to allow for the determination of the quantity of *K. phaffii* proteins relative to the quantity of rhLF and determination of which residual *K. phaffii* proteins from strain GS-115 were most abundant. Abundance data was not used to select proteins for comparison to allergenic proteins.

Additional methods employed in the analysis utilized the R programming language, leveraging several key packages for data manipulation and visualization: (1) filtering specific proteins: using the *dplyr* package, proteins of humans, pigs, sheep, and bovine origin were excluded. This step was critical to eliminate potential contaminants and to focus solely on relevant proteins; (2) normalization of protein intensities: protein intensities were normalized to their relative percentages of the total measurable protein intensities across samples. This normalization was performed using functions from the *dplyr* package, which facilitated the standardization of data across different samples or batches; (3) batch consistency analysis: data was aggregated by batch, and average intensities along with standard deviations were calculated using *dplyr*. This analysis was essential to ensure batch-to-batch consistency in our processes; (4) visualizations: The *ggplot2* package was employed to create intuitive visual representations of the data, including bar charts and pie charts. These visualizations effectively communicate significant trends and variations in protein abundance across batches, providing clear insights into the dataset.

2.2.2 AllergenOnline, version 22

The www.AllergenOnline.org database (version 22, released on May 25, 2023) is a publicly available tool that provides a peer-reviewed allergen list and sequence database to identify proteins that may pose risk of allergenic cross-reactivity. Updated annually, the website is designed to help assess the safety of novel proteins developed through genetic engineering or food processing prior to their introduction into the food supply (40).

Identified proteins from UniProt that were matched by LC–MS/MS were queried against allergens in AllergenOnline to determine significant alignment with known allergens. The full amino acid sequence of human lactoferrin and each identified *K. phaffii* protein (strain GS-115) and *T. reesei* mannosidase were entered into the sequence search window of AllergenOnline using the full-length FASTA, version 36 search. A separate search was performed for 80 amino acid alignments using FASTA, which uses a sliding window of 80 amino acid segments of each protein to find identities of

greater than 35%. Matches to allergenic proteins were evaluated further using publications in the AllergenOnline.org database associated with the matched allergens as well as consideration of the breadths of taxonomic matches to the same protein by BLASTP.

2.2.3 BLASTP in NCBI entrez protein database

The Basic Local Alignment Search Tool (BLAST®) is a publicly available tool on the NCBI website (<http://www.ncbi.nlm.nih.gov/BLAST/>) that identifies regions of local similarity between gene or protein sequences. Protein Blast (BLASTP) identifies a protein sequence against the NCBI Entrez Protein Database, a search and retrieval system that allows users to access information from many health science databases to provide information on protein structure, function, and evolution.

Given the frequency by which BLASTP is updated compared to AllergenOnline (daily or weekly *versus* annually), this tool was utilized to identify newly discovered allergens. The identified proteins were queried in BLASTP, version 2.14.01, on September 26, 2023 to check for allergens that may have been missed in AllergenOnline. Two searches were conducted: first, without any keyword limit looking for the best identities and, second, with a keyword limit of “Homo sapiens species taxonomic ID 9606” to consider likely tolerance limits (41). The primary criterion to judge the significance for any alignment was whether the BLASTP result for “no keyword” was higher than BLASTP of the protein to human proteins. In addition, high identity matches across diverse taxa are highly unlikely to act as clinically relevant allergens based on restricted clinical mandates from the allergy literature.

2.3 *In vitro* pepsin digestibility study

A simulated *in vitro* digestion study was performed to determine the stability of Helaina rhLF in simulated gastric conditions, compared to hmLF and bLF, in the presence of pepsin enzyme per Codex guidelines. Experimental procedures were modeled after those developed by Astwood et al. (36) and refined by Thomas et al. (37) and Ofori-Anti et al. (38).

2.3.1 Preparation of LF (rhLF, hmLF, and bLF), simulated gastric fluid, and controls

Recombinant human lactoferrin (Helaina rhLF), α -isoform, was produced by the fermentation of *K. phaffii* (GS-115) recombinant system at the industrial scale according to established procedures (42–44). At the end of fermentation, cells were separated from the protein-containing broth, and recombinant human lactoferrin was produced by fermentation and expressed from a *K. phaffii* (GS-115) recombinant system, purified by microfiltration/diafiltration and cation exchange chromatography and then spray-dried to a powder. The purity was measured by high-performance liquid chromatography (HPLC) and sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and is typically greater than 95%. Iron saturation was determined to be 58% saturated (26). Helaina

rhLF was reconstituted and diluted into phosphate-buffered saline (PBS) prior to use.

Lactoferrin was isolated from human milk that was prepared by the Northwest Mothers Milk Bank and provided by Brian Scottoline (MD, PhD) at Oregon Health and Science University (OHSU). The milk from six donors was pooled and frozen prior to use. hmLF was then isolated from the milk using the following procedure: the milk samples were centrifuged at $10,000 \times g$ for 30 min, and the liquid fraction removed. The resultant liquid fraction was adjusted to pH 4.7 before incubation at 40°C for 30 min for casein precipitation. The solution was then centrifuged at $10,000 \times g$ for 30 min at 4°C to separate liquid and solid fractions. The supernatant was decanted and purified via ultrafiltration/diafiltration (UFDF) and chromatography as described above for rhLF. The purified product was then buffer-exchanged into $1 \times$ PBS, pH 7.4 (no spray drying), and concentrated against a 30-kDa membrane. This procedure yielded approximately 1 g of LF per liter of human milk with a purity level greater than 95% as measured by HPLC and SDS-PAGE. The measured iron saturation was ~30%. This work was not determined to be human research by the OHSU Institutional Review Board.

Bovine lactoferrin (bLF) isolated from bovine milk was purchased from Lactoferrin Co, Australia (product 11683) and reconstituted into PBS prior to use. The purity provided by the supplier was greater than 95%, and iron saturation was 9.9%.

Lyophilized β -lactoglobulin (Sigma, #L3908) was reconstituted with PBS, and the concentration (mg/mL) was verified via Nanodrop® (model: NanoDrop™ One/OneC, Thermo Fisher Scientific).

Simulated gastric fluid (SGF) was comprised of 0.084 N HCl, 35 mM NaCl, pH 2.0, and 3,200 units (U) porcine pepsin (Sigma, P6887, 4,048 U/mg activity) to obtain 10 U pepsin activity per microgram of protein.

The protein-positive (no pepsin) and protein-negative (no protein) controls were SGF containing 0.2 mg/mL LF (or β -lactoglobulin) without pepsin and SGF containing 3,200 U of pepsin with no LF, respectively. Positive controls were prepared for each experimental sample.

2.3.2 Experimental procedures

Each experimental tube contained a total volume of 1.6 mL of SGF, 0.2 mg/mL of LF (rhLF, hmLF, or bLF), and 3,200 U of pepsin to create a ratio of 10 units of pepsin to 1 μ g protein. β -Lactoglobulin (0.2 mg/mL) was used as a non-digestible control. Pepsin was added immediately to each tube and mixed. The tubes were incubated at 37°C, and 100- μ L aliquots were removed after specified sample times: 0.5, 2, 5, 10, 20, 30, and 60 minutes. Pepsin activity was quenched using 35 μ L of 1 M NaHCO₃, pH 11, for every 100 μ L of the experimental digestion aliquot. For the 0 time-point control, pepsin activity was quenched prior to the addition of protein.

Samples from each digestion time-point, including all controls, were run on a tris-glycine gel to visualize the degradation of rhLF in comparison to hmLF and bLF. Additionally, all proteins were visually compared to a corresponding 10% protein control (0.02 mg/mL) to ensure $\geq 90\%$ digestion profile of each protein. To

determine the detection limit, a serial dilution of protein samples was prepared with PBS covering the range of 200% total protein (0.4 mg/mL) per well to 5% (0.01 mg/mL) per well. Samples were separated by SDS-PAGE at a constant 185 V for 45 min, followed by staining with Imperial Protein Stain.

2.3.3 Automated western blot

The JessTM Simple Western System (ProteinSimple, San Jose, CA, USA) is an automated protein separation and immunodetection system. To visualize and quantify the digestion of proteins in SGF, manufacturer's instructions for 12–230 kDa JessTM separation module (ref. no. SM-FL004) were followed. The experimental samples were diluted to 0.004 mg/mL with distilled water, and then one part of 5× Fluorescent Master Mix was combined with four parts of the diluted sample to achieve a final concentration of 0.0032 mg/mL. The samples were denatured for 10 min at 90°C and centrifuged for 5 min at 2,500 rpm (~1,000 × g). The fluorescent molecular weight marker, (12–230 kDa, biotinylated ladder), and sample proteins were separated in capillaries as they moved through a separation matrix at 375 V. JessTM immobilized the protein samples to the capillary walls using UV light. In addition, it automatically incubated the samples in blocking solution (antibody diluent 2, ProteinSimple, cat. no. 042-203), primary antibody (anti-human lactoferrin antibody, Sigma, cat. no. L3262), and secondary antibody (goat-anti-rabbit secondary antibody, ProteinSimple, cat. no. 040-656), and we performed the necessary washing steps between incubations. Chemiluminescence was established with a luminol-peroxide mix (luminol-S, ProteinSimple, cat. no. 043-31; peroxide, ProteinSimple, cat. no. 043-379). Digital imaging was captured with Compass Simple WesternTM software (version 6.2.0, ProteinSimple) to produce gel images and histograms. Relative quantification was measured by calculating the percent (%) area of the chromatogram peaks.

2.3.4 Conventional western blot

Samples were separated in an SDS-PAGE at a constant 185 V for 45 min, the gels were removed from the cassette, and the iBlot 2 Gel Transfer was used to transfer the proteins to polyvinylidene difluoride (PVDF) membranes. The membranes were placed in an opaque gel box and incubated with blocking buffer [PBST + 10% fish gelatin (PBS 10×, pH 7.2, Thermo Fisher Scientific, cat. no. 70-013-032; biotium 10× fish gelatin blocking agent, Thermo Fisher Scientific, cat. no. NC0382999) at room temperature for 1 h. The samples were washed with PBST [PBS + 0.1% Tween 20 (Tween 20, Sigma-Aldrich, cat. no. P9416-100ML)] and then incubated with the primary antibody (rabbit anti-human lactoferrin antibody; Sigma, cat. no. L3262) for 1 h at room temperature. The samples were washed and incubated with the goat anti-rabbit Alexa-Fluor secondary antibody (Thermo Fisher Scientific, cat. no. A-11034) at room temperature for 30 min. The samples were washed five times and then rinsed for 2 min in distilled water to stop the reaction. The results were visualized using the iBright imager (Thermo Fisher Scientific) and the following settings: Alexa Fluor 488, no false colors, exposure time: 00:00 (min/s).

3 Results

3.1 Literature searches identify no allergenic risk for rhLF, *K. phaffii*, or its glycans

Literature searches in PubMed for the primary protein of interest, hLF, did not identify studies showing allergy to this protein, even though the amino acid sequence is moderately identical to bovine lactoferrin (70% full-length identity) and ovotransferrin (52% overall identity), known potential allergens. The literature results included a study on rhLF from transgenic cows that underwent an allergenicity assessment using the Codex guidelines and included a bioinformatics analysis, stability testing of rhLF in pepsin, and serum reactivity tests to evaluate potential allergenicity (34). Considering the rapid digestion by pepsin and no specific binding of IgE using serum from patients with milk and egg allergy, the authors concluded that the allergenicity potential of this protein is very low (34). A single paper that suggested a possible risk from human lactoferrin was for protein expressed in rice that has cross-reactive carbohydrate determinants (CCD) on asparagine-linked glycosylation sites which may be linked to allergy (45). Importantly, the rice-expressed protein did not elicit basophil activity due to IgE binding to CCD, and there was no proof of allergy to that product, even after the clinical oral challenge. These results indicate that the *in vitro* methods for evaluating cross-reactivity are not sufficient to predict clinical outcomes of allergenic risk. In addition, the IgE binding described in the paper on rice was to complex carbohydrates with $\alpha(1-3)$ -fucose and/or $\beta(1-2)$ -xylose on the stem loop, which Helaina rhLF does not contain.

Literature searches were also conducted for peer-reviewed evidence of allergenic risk to proteins from *K. phaffii* and *P. pastoris*. The searches used combinations of species names “*Komagataella phaffii*” and “*Pichia pastoris*” with either “clinical allergy” or “food allergy” or “IgE binding”. Keyword limits “allergy” and “*K. phaffii*” identified one publication (45). “*P. pastoris*” identified 156 publications, including Jin et al. on the production of soybean leghemoglobin in the same yeast expression system (35). The other 155 publications describe the cloning and expression of various plant and fungal proteins for laboratory testing, including major allergens such as the 2S albumin Ara h 6 from peanuts. Using the full Latin name “*Komagataella phaffii*” and “clinical allergy” did not identify any papers. Using the search terms “*Pichia pastoris*” and “clinical allergy” identified 34 papers, including one with a direct administration of recombinant human serum albumin made in this yeast, injected three times over 3 days, in 423 cirrhosis patients with ascites or edema (46). There were 96 adverse reactions from the treatment, but no allergic reactions or IgE binding to yeast proteins were reported. Furthermore, the injection administration route of recombinant human serum albumin is not applicable to Helaina rhLF that would be taken orally as a dietary protein and digested through the gastrointestinal tract.

Information relative to the expression system, *K. phaffii*, was considered, including the potential addition of glycans. While *K. phaffii* produces oligomannose glycans that can be bound by mannose receptors on dendritic cells (DC), Kreer et al. found no proof for the DC triggering immunogenic and/or subsequent allergic reactions (47). Glycans that have been implicated in IgE binding are CCD that are common on many plant-produced glycoproteins or $\alpha(1,3)$ -galactose (α -gal) sugars added to proteins in tick saliva and on some red meat proteins (45, 48). Importantly, Helaina rhLF does not contain any of these CCD, and the literature does not support the allergenicity of the oligomannose N-glycans that make up Helaina rhLF (26). Furthermore, this study revealed that Helaina rhLF and hmLF have identical amino acid sequences, and despite differences in glycans, both proteins folded nearly identically; rhLF underwent thermo-denaturation at the same temperatures as apo- and holo-hmLF, and both bound iron, suggesting comparable structure and potential functionality (26). These data, along with the literature search results, suggest that rhLF and hmLF would have a similar allergenic risk potential, regardless of glycan profile. However, it is acknowledged that *K. phaffii* has the ability to add high-level mannose structures which could have immunogenic potential; this has been mitigated in Helaina rhLF through the inactivation of OCH1 to limit the number of mannose residues (27, 49). Nonetheless, as Helaina rhLF does contain slightly higher mannose structures than hmLF, it is important to characterize the potential for immunogenicity. The authors have identified testing strategies to ensure the immunogenic safety of Helaina rhLF (EfferTM) that were vetted by experts in the field of immunology/allergy, immunotoxicology, lactoferrin biology and nutrition, and food safety assessment, the details of which were recently published (50).

3.2 Bioinformatics search strategies show no significant risk for cross-reactivity and allergenicity

Protein identification on three individual batches of Helaina rhLF was performed using LC-MS/MS. This analysis confirmed the expected sequence for rhLF with full protein coverage, excepting signal sequence and those peptides unlikely to be detected due to the size of peptides (m/z cutoff). Label-free quantitation indicated that $\geq 99\%$ of the identifiable protein content was, in fact, the target protein, human lactoferrin (Table 1, Figure 1A, Supplementary Table 1). A comparison of hLF sequence to AllergenOnline identified protein matches of 70% identity to *Bos taurus* (bovine) lactotransferrin (lactoferrin), which is a minor milk allergen, and 52% identity to ovo-transferrin, a minor food allergen in chicken eggs, by overall FASTA (Table 2). However, by sliding 80 AA window, the identities were higher; bovine lactotransferrin was 83.8% with rhLF, and there were 631 matched 80mers compared to 69.7% overall identity to rhLF. The best 80mer match was to ovo-transferrin at 67.6% ID with 631 matches of 80 AA compared to 51.8% overall ID by FASTA. The remaining <1% protein content included 36 host cell proteins (Table 1), with no single protein having an average relative abundance greater than 0.05%. Some

variability in host cell proteins was observed between the three batches (Figure 1B); however, no single protein had greater than 0.075% relative abundance in any single batch (Table 1).

A bioinformatics search strategy was used to determine the amino acid sequence homology and structural similarity of the identified *K. phaffii* proteins (strain GS-115) to known human allergens. The amino acid sequences of 36 unique proteins were identified by LC-MS/MS (Table 1, Supplementary Data), verified by UniProt accession numbers, and searched using AllergenOnline and BLASTP. The workflow for protein identification and subsequent risk evaluation by homology searching is shown in Figure 1C. The overall results of the bioinformatics comparisons to AllergenOnline by FASTA and to NCBI Protein database by BLASTP are defined in Supplementary Table S1. Importantly, of these 36 identified proteins, none was linked with a significant risk for cross-reactivity and allergenicity.

The most abundant host cell protein identified was phosphoglycerate kinase at 0.045% of the protein content (Table 1) and no allergen match by sliding 80 AA window (Table 3). It had the best overall FASTA alignment to α -s1 casein at 23% overall identity. The BLASTP was to itself, and the best human alignment was 65% ID to human phosphoglycerate kinase. Proteins with high identity to human proteins are likely to be tolerated, and this identity match is much higher than to the milk allergen α -s1 casein and presents a low risk of allergic cross-reactivity. The second most abundant host cell protein was endoplasmic reticulum chaperone BiP protein (Table 1), found to have 84%–36% ID to heat shock protein (HSP) of mosquito, mites, fungi, and grass with BLASTP to itself and 65.6% ID to human HSP, indicating likely tolerance due to close identity to human (Table 3). Mannosidase from *T. reesei*, which is a gene added to *K. phaffii* to limit the potential for hypermannosylation (51), was the next most abundant protein (Table 1) and was determined to be of low allergenic risk due to low abundance (0.033%), no match by sliding 80 AA window, and 31% ID to the human homologue (Table 3). Of the remaining host cell proteins (Table 1), five were found to have >50% 80mer match to a known or putative allergen (Table 3, Supplementary Table S1). Catalase (0.014%) was found to have a match to a potential allergen, *Penicillium citrinum* catalase, at 60% identity; however, this protein was at low abundance of 0.03% and had high identity (51%) to human catalase, indicating likely tolerance. 5-Methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase, which showed 77.5% ID to Sal k 3 pollen allergen, was also identified, with 48% ID overall FASTA to that protein and 100% ID to *K. phaffii* but no matches to human protein. By BLASTP, this protein had 88% ID to 100 enzymes of a wide variety of plants, trees, and weeds; there is only one laboratory that has published data on IgE binding to this pollen protein as an allergen (52). It is highly unlikely that there is any clinically identifiable co-reactivity as it is a low-abundance enzyme (<0.011%), therefore with limited risk. Cytosolic superoxide dismutase was also identified (0.010%) with 62% identity to Ambrosia and olive superoxide dismutase and 100% identity to *K. phaffii* superoxide dismutase. However, this protein also had 59% identity to human superoxide dismutase, suggesting tolerance due to human recognition. Cytochrome c isoform 1 (0.008%), with one

TABLE 1 Proteins in Helaina rHLF product.

Identified proteins	Accession number	Batch 1 relative %	Batch 2 relative %	Batch 3 relative %	Batch average relative %
Lactoferrin	Lactoferrin	99.76877025	99.66679381	99.92559697	99.78705368
Phosphoglycerate kinase	C4QY07	0.071008065	0.055368055	0.010088388	0.045488169
Endoplasmic reticulum chaperone BiP	C4QZS3	0.03425931	0.060051426	0.034198497	0.042836411
Mannosidase TR	Mannosidase TR	0.020111374	0.074992832	0.003290942	0.032798383
SCP domain-containing protein	C4R3H3	0.026875124	0.043759254	0.014924343	0.028519573
Catalase	C4R2S1	0.008502079	0.024912314	0.008422246	0.013945546
5-Methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase	C4QZU2	0.009931174	0.012304305	0	0.011117739
Superoxide dismutase (Cu-Zn)	C4R8X7	0.005425298	0.023314325	0.001151436	0.009963687
Ribonuclease T2	C4QZY8	0.012172773	0.008977466	0.003208633	0.008119624
Cytochrome c isoform1	C4R6L9	0.015564249	0.000428016	0	0.007996132
Uncharacterized protein	C4R0Z8	0.002923402	0.007970453	0.000780232	0.003891362
L-type lectin-like domain-containing protein	C4R2L8	0.008521138	0.00173005	0.001235001	0.00382873
Metalloprotease with similarity to the zinc carboxypeptidase family	C4R4I2	0	0.005215868	0.002293202	0.003754535
Endo-beta-13-glucanase major protein of the cell wall involved in cell wall maintenance	C4QYF3	0	0.003487018	0	0.003487018
Uncharacterized protein	C4R0V6	0.002923402	0	0	0.002923402
13-Beta-glucanosyltransferase	C4QVL4	0	0.002356878	0	0.002356878
AP-1 accessory protein	C4R325	0	0	0.002168491	0.002168491
Heat shock protein that cooperates with Ydj1p (Hsp40) and Ssa1p (Hsp70)	C4QV89	0.003942372	0.001490769	0.000340725	0.001924622
Major exo-13-beta-glucanase of the cell wall involved in cell wall beta-glucan assembly	C4R0Q7	0	0.001824648	0	0.001824648
Aminotransferase class I/class II domain-containing protein	C4R862	0.003218988	0.001592512	0.000512162	0.001774554
Molecular chaperone	C4QVC4	0.001764937	0	0	0.001764937
Superoxide dismutase copper/zinc-binding domain-containing protein	C4QW48	0.00228556	0	0.00108123	0.001683395
Protein of the SUN family (Sim1pUth1pNca3pSun4p) that may participate in DNA replication	C4R2Z5	0	0.001646576	0	0.001646576
Type II HSP40 co-chaperone that interacts with the HSP70 protein Ssa1p	C4R2Q1	0.001411637	0.00045896	0	0.000935299
Sedoheptulose 17-bisphosphatase	C4R2M0	0.000546974	0.000771788	0	0.000659381
Fructose-bisphosphatase	C4R5T8	0	0.000450901	0	0.000450901
Xylose and arabinose reductase	C4R135	0.000386269	0.000487704	0	0.000436987
Dihydrolipoyl dehydrogenase	C4R312	0	0.000389153	0	0.000389153
Serine/threonine-protein phosphatase	C4R1A4	0	0.000328839	0	0.000328839
FACT complex subunit	C4QYQ8	0.000319168	0	0	0.000319168
Cellulase	C4R8H7	0	0.000284482	0	0.000284482
Alanine-glyoxylate transaminase	C4R7U0	0	0.000257802	0	0.000257802
Transketolase similar to Tkl2p	C4R5P8	0.000250801	0	0	0.000250801
Fructose-bisphosphate aldolase	C4QW09	0.000246246	0	0	0.000246246

(Continued)

TABLE 1 Continued

Identified proteins	Accession number	Batch 1 relative %	Batch 2 relative %	Batch 3 relative %	Batch average relative %
ATPase involved in protein folding and the response to stress	C4R3X8	0	0.000214762	0	0.000214762
Lectin-like protein with similarity to Flo1p thought to be expressed and involved in flocculation	C4QYW7	0	0.000184489	0	0.000184489
Plasma membrane Mg ⁽²⁺⁾ transporter expression and turnover are regulated by Mg ⁽²⁺⁾ concentration	C4QXF3	0	0	0	0

Proteins were identified by mass spectrometry. Human lactoferrin is shown for comparison. Accessions (Uniprot) and protein name (Uniprot) are shown. % abundance was estimated by label-free quantitation calculated as a percentage of the total MS abundance of the total identifiable protein abundance.

80mer match of 74% ID to *Curvularia lunata* cytochrome c, full ID to *Curvularia lunata* cytochrome c, 100% ID to *K. phaffii*, and 65% ID to human cytochrome c as well as many other proteins, was also present. An ATPase involved in protein folding (0.0002%) was also identified with 91% identity to *Davidiella* HSP and 14 other proteins at lower ID, 100% identity to *K. phaffii* ATPase, and 75% identity to human HSP70, suggesting tolerance due to human recognition. The remaining host cell proteins identified (Table 1) had <50% 80mer match to any allergen and therefore have low allergenic risk (Supplementary Table S1).

3.3 Helaina rhLF digests in a comparable manner to human milk LF and shows low allergenic risk potential in the pepsin digestion model

Previous studies have shown that pepsin digestion of proteins in a simulated gastric fluid (pH 2.0) can be used to correlate allergenic risk potential (53, 54). In order to assess the digestibility of Helaina rhLF, an *in vitro* digestion study was performed in SGF, pH 2.0, with a pepsin/protein ratio of 10 units pepsin per 1 µg of protein. Successful digestion to ≤10% of the initial starting concentration indicates low allergenic risk potential (38). Initially, to evaluate the staining characteristics of each test protein and to ensure detection down to 10% of the target protein, a preliminary SDS-PAGE-stained gel with samples of serially diluted target protein was performed and showed similar detection levels of all proteins tested (Supplementary Figures S1A, B). Next, the pepsin digestion study was performed as previously described to test the digestibility of rhLF compared to hmLF and bLF. These results indicate that Helaina rhLF digests in a comparable manner to LF isolated from human milk and digests >90% in the presence of pepsin, as no detectable band is visible compared to the 0.02 mg/mL, 10% protein control (Figures 2A, B). Furthermore, the gel images show that the rhLF is highly pure (>99%), and no visible detection of host proteins is observed (Figure 2B). Bovine LF shows a similar trend (Supplementary Figure S1C). β-Lactoglobulin, a known allergen previously shown to not digest in the presence of pepsin (55), appears completely intact, indicating that it was not digested in these studies (Supplementary Figure S1D), thus confirming the accuracy of the assay. We further verified the digestion of the proteins using an automated western blot technique (JessTM, ProteinSimple, Bio-Techne). Importantly, this test further

confirmed that Helaina rhLF digests in a comparable manner to hmLF, with complete digestion of intact protein observed within 20 min of digestion (Figures 2C, D) and significant overlap of digestion products indicating similar peptide identities between rhLF and hmLF (Figure 2D). These data also corroborate the high purity of rhLF and a lack of host protein presence. Interestingly, one peptide survived pepsin digestion with an approximate molecular weight of 2 to 3 kDa (Figure 2D) which correlates with published literature as lactoferricin (56) and remains intact throughout the entire 60-min exposure (Supplementary Figure S1E). Additionally, quantification of intact proteins shows similar amounts between rhLF and hmLF throughout the digestion procedure, with no intact protein remaining by 10 min for rhLF and 20 min for hmLF (Figure 2E). A pepsin-only control (no LF) was used to distinguish LF-specific peaks in the chromatogram from non-specific peaks (Supplementary Figure S1F). Here pepsin can be detected and is shown at ~41 kDa as well as an unknown artifact at ~27 kDa (Supplementary Figure S1F). Finally, to further corroborate these findings, a standard western blot analysis was performed and showed similar results (Supplementary Figure S1G). Overall, these data indicate that Helaina rhLF digests comparably to hmLF, both showing low allergenic risk potential, and further suggest similar digestibility and potential bioavailability between the proteins.

4 Discussion

Novel foods require thorough safety evaluation prior to their introduction into the food supply, and the Codex Alimentarius Commission's weight-of-evidence approach to evaluate the allergenicity risk potential provides a framework by which to evaluate this aspect of the safety profile of novel proteins (30, 31). To this end, we conducted a rigorous literature search, several bioinformatics exercises using AllergenOnline and BLASTP, and an *in vitro* digestion study using simulated gastric conditions to evaluate the potential allergenic risk of Helaina rhLF, including its residual proteins expressed in *K. phaffii*. Neither the literature search nor the bioinformatics exercises and pepsin digestion study suggest Helaina rhLF to pose allergenic risk to consumers based on allergens known in 2023.

Literature searches for the primary protein of interest, human lactoferrin, did not identify studies showing allergy to this protein, even though the amino acid sequence is moderately identical to known potential allergens, bovine lactoferrin and ovo-transferrin—

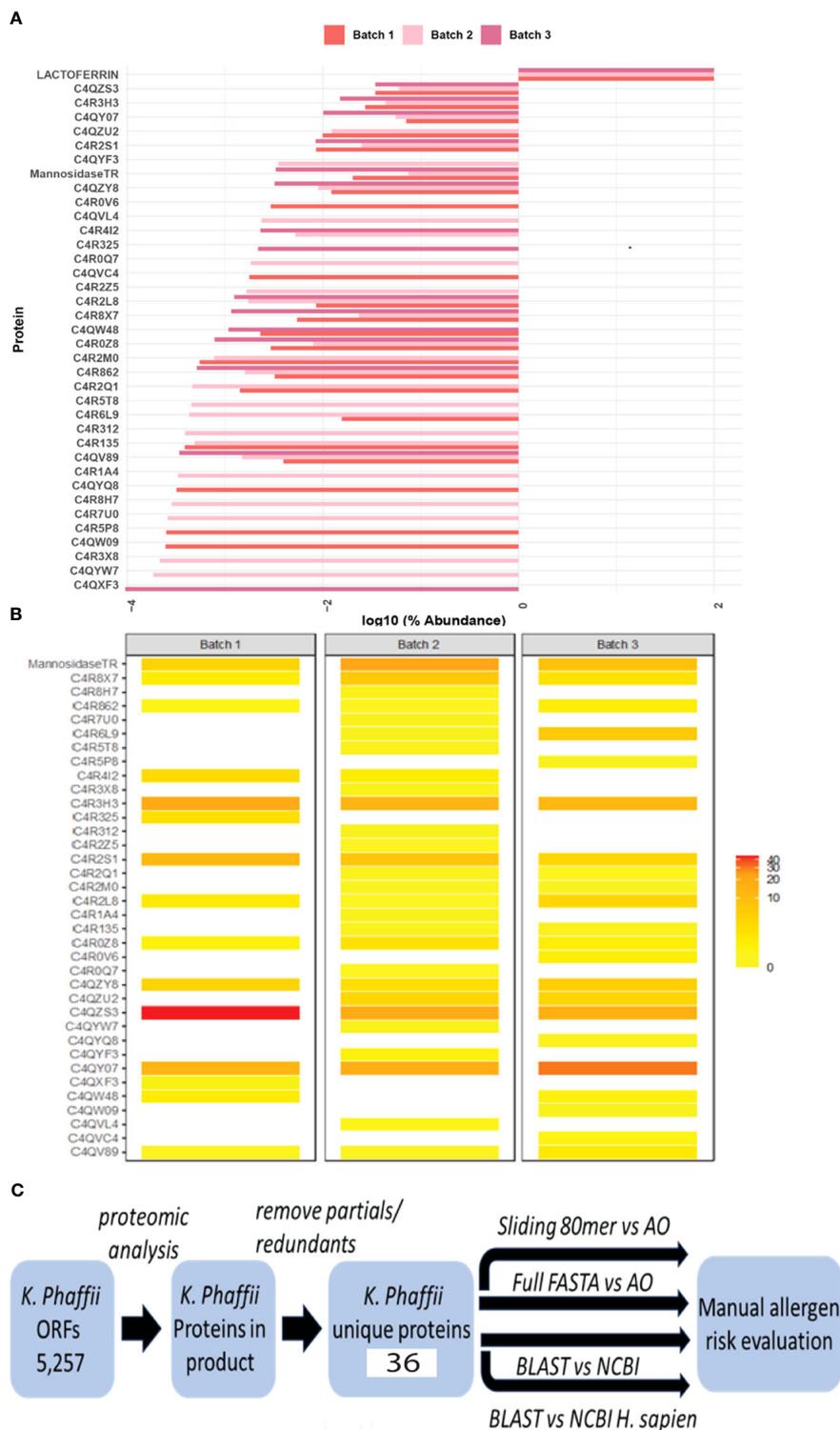


FIGURE 1

Bioinformatics search strategies show no significant risk for cross-reactivity and allergenicity. (A) Relative abundance of rhLF compared to *K. phaffii* host proteins on a log₁₀ scale in three representative batches. Lactoferrin comprises more than 99% of the total proteins in these batches. (B) Heatmap showing the abundance of host cell proteins. (C) Workflow to evaluate the allergenic risk of residual *K. phaffii* proteins. Mass spectrometry was used to determine the presence of proteins in rhLF product from a database of *K. phaffii* open reading frames. After removing redundant sequences and proteins not specific to *K. phaffii* GS-115 or *T. reesei* (mannosidase), 36 protein sequences were used to query AllergenOnline (AO) using identity in a sliding 80mer search and by full-length FASTA identity search. The 36 protein sequences were also used in a BLASTP search of NCBI proteins and in a separate NCBI search limited to *H. sapiens* proteins. Evaluation of risk is manual and is informed by the results of these four sequence identity comparisons.

TABLE 2 Recombinant human lactoferrin allergen assessment.

Accession number	FASTA of NCBI protein	80mer match	Best overall FASTA vs. v22	BLASTP vs. <i>Homo sapiens</i>	Is there a risk of allergy?
P02788	Lactoferrin	83.8% ID to <i>Bos taurus</i> lactotransferrin	70% ID to <i>Bos taurus</i> lactotransferrin	100% ID human lactoferrin	No tolerance to human
		67.6% ID to ovo-transferrin	52% ID to ovo-transferrin	100% ID human lactoferrin	No tolerance to human

for example, there are no publications showing IgE cross-reactivity to hLF from sera of those allergic to bovine milk, but in rare cases, bLF is the target of IgE binding from those allergic to cow's milk outright (57) or due to α -gal carbohydrate on the asparagine-linked sites on bLF (58). The other suggested cross-reactive allergen based on bioinformatics searches is ovo-transferrin, or Gal d 3, a major allergen naturally found in egg white that binds IgE (59, 60). Importantly, a study by Zhou et al. evaluated potential serum IgE binding between human LF (rhLF from transgenic cows or hLF) or bovine LF and sera from individuals allergic to milk or egg and found no visible cross-reactivity of rhLF, hLF, or bLF to the test serum samples (34). The paper that suggested a possible risk from human lactoferrin was for protein expressed in rice that has cross-reactive carbohydrate determinants on some asparagine-linked glycosylation sites (45). Interestingly, the rice-expressed protein did not elicit basophil activity, and there was no proof of allergy to that product. Literature searches also did not identify studies of clinical allergy to the host source of *Komagataella phaffii* or using the previous name, *Pichia pastoris*.

Asparagine-linked glycans in eukaryotes can influence laboratory results in evaluating the potential risks of allergy, such as the introduction of CCD in rice-produced hLF as described above (45). In the case of α -gal, which can occur in mammalian hosts, studies have shown this allergenic glycan in bovine milk, γ -globulin, lactoferrin, and lactoperoxidase (58). In the study evaluating rhLF produced in transgenic cows (34) which would likely contain α -gal, no increased allergenic risk was observed, including no specific IgE binding in serum. Helaina-evaluated glycans detect neither CCD nor α -gal and contain only N-linked glycans that are oligomannose. While some studies have linked mannose binding the mannose receptor on DC to the increased expression of Th2 cytokines, no definitive allergic testing was performed or identified (61, 62). Furthermore, Kreer et al. did a thorough analysis of N-glycans from *K. phaffii* and found no proof for mannose receptor on DC triggering allergic reactions (47). Therefore, Helaina rhLF glycans are unlikely to pose a significant risk of actual (α -gal) or perceived (CCD, mannose) allergen epitopes.

The current study identified that Helaina rhLF is a very pure protein with $\geq 99\%$ of the protein content being the target protein, hLF, which is not linked with allergy. Due to low abundance of host proteins (<1% total) and the most abundant protein being only 0.045% of the total protein content, allergic responses against host proteins are unlikely. Nevertheless, a bioinformatics approach was utilized to determine the allergenic potential of *K. phaffii* host proteins. The Codex Alimentarius criterion to consider the potential risks of IgE cross-reactivity and clinical co-reactivity is

35% identity over 80 amino acids. As presented by Abdelmoteleb et al., these criteria are too low to evaluate the potential risks of novel foods (63). There are many proteins conserved in evolution, and allergy across species is much more restricted to proteins that are not loose homologues. For most protein types, sequences of much greater than 50% identity are needed over most of the length of proteins for probable cross-reactivity (63). This risk is further reduced with AA sequence identity linked to human proteins, as these are likely to induce tolerance. These were the criteria considered when evaluating cross-reactivity and allergenic risk for the *K. phaffii* proteins identified in Helaina rhLF. The BLASTP matches of the host proteins to NCBI Protein did not uncover any matches to proteins that are likely allergens and that are not in the AllergenOnline version 22 database. There were no high identity matches to allergens that would suggest potential clinical relevance or risks based on either the AllergenOnline sequence alignments or the NCBI Protein BLASTP searches. There are some interesting alignments to evolutionarily conserved proteins, yet those are unlikely to be allergens as human consumers would react to a wide variety of food sources that contain these proteins, which is rarely, if ever, documented.

Several studies have demonstrated pepsin resistance in food allergens, therefore showing a correlation between allergenic potential and resistance to pepsin digestion (37, 38, 64)—for example, these studies have shown that allergenic proteins, such as β -lactoglobulin, are resistant to pepsin digestion, whereas non-allergenic proteins, such as horseradish peroxidase, concanavalin A, and soybean leghemoglobin, are susceptible to pepsin digestion. Previous work by Almond et al. (65) and Ofori-Anti et al. (38) showed that rhLF (expressed in rice) and/or human LF are readily digested in the presence of pepsin. In our current studies, pepsin digestion modeling showed that Helaina rhLF was also rapidly digested in simulated gastric fluid. Importantly, its digestibility profile was comparable to hmLF, thus demonstrating a low allergenic risk. Interestingly, the digested products (Helaina rhLF and hmLF) contained a small peptide that survived the entire digestion procedure, which corresponds in molecular weight to lactoferricin, a known peptide of lactoferrin (56).

Importantly, the use of these prediction models and safety risks has proven successful over the years for protein evaluation prior to introduction to the food supply (21, 24, 66, 67). This includes bovine milk basic protein (66), which was determined to be GRAS in 2006 (GRN196) (67). Additionally, two other recombinant proteins produced in *K. phaffii*, recently added to the food supply, were assessed following Codex procedures as described in the current work. Soy leghemoglobin (24) showed a similar low allergenic risk potential based on the protein sequence (although

TABLE 3 Allergenic assessment of select host cell proteins.

Accession number	FASTA of NCBI protein	Relative abundance	80mer match	Best overall FASTA vs. v22	BLASTP vs. protein	BLASTP vs. <i>Homo sapiens</i>	Is there a risk of allergy?
C4QY07	Phosphoglycerate kinase	0.045%	No match	Alpha s1 casein 23% ID	>CCA37205.1 3-phosphoglycerate kinase (<i>Komagataella phaffii</i>)	65% ID human phosphoglycerate kinase	No; <50% ID and high ID to human protein
C4QZS3	Endoplasmic reticulum chaperone BiP	0.043%	84% down to 36% to HSP of mosquitoes, mites, fungi, and grass	65% down to 46% to HSP of mosquito, mites, fungi, and grass	100% <i>Komagataella phaffii</i> and down to 75% other fungi	65.6% HSP	No; tolerance to human
GORBB5	Mannosidase (<i>T. reesei</i>)	0.033%	No matches	No matches by FASTA	100% to <i>Trichoderma reesei</i> full-length down to 58% ID to other fungal homologous proteins	31% ID to human protein	No; tolerance to human
C4R2S1	Catalase	0.014%	60% ID to <i>Penicillium citrinum</i> catalase over 80 AA and 37% ID of that protein by full-FASTA	37% ID <i>Penicillium</i> catalase and less for three species	100% to <i>Komagataella phaffii</i> down to 65% ID for 100 fungal proteins	51% down to 41% to 9 human proteins catalase	No; tolerance to human
C4QZU2	5-Methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase	0.011%	77.5% ID to the same two proteins as by overall FASTAS, Sal k 3 and other	49% ID to Sal k 3 and 48% to another methyltetrahydropteroyltriglutamate protein	100% ID <i>Komagataella</i> cobalamin down to 77% ID to many fungal proteins	No match	No; low abundance
C4R8X7	Cytosolic superoxide dismutase	0.010%	62% to <i>Ambrosia</i> superoxide dismutase and to 57% to 23 olive superoxides	59% ID to <i>Ambrosia</i> superoxide dismutase and 52% or higher to olive	100% ID to <i>Komagataella phaffii</i> down to 75% ID 100 fungi	59% ID to human superoxide dismutase down to 44% ID to other human superoxides	No; tolerance to human
C4R6L9	Cytochrome c isoform 1	0.008%	74% ID to <i>Curvularia</i> cytochrome c	71% ID to <i>Curvularia lunata</i> cytochrome c	100% ID to <i>Komagataella</i> cytochrome c and 80% ID to many other fungal cytochromes	65 ID to human cytochrome c multiple proteins (genes) and heat shock protein with more than 55, 14, >52% ID	No; tolerance to human
C4R3X8	ATPase involved in protein folding	0.0002%	91% ID to <i>Davidiella</i> HSP and 14 more to lower identities	Heat shock proteins <i>Tyrophagus</i> 76% ID down to 46% ID pollen HSP	100% ID <i>Komagataella</i> ATPase and HSP 85% <i>Ascoidea</i> down to different fungi ~100 with down to 83% identity fungal proteins	75% ID to heat shock protein 70 <i>Homo sapiens</i> and 100 to 45% ID	No; tolerance to human

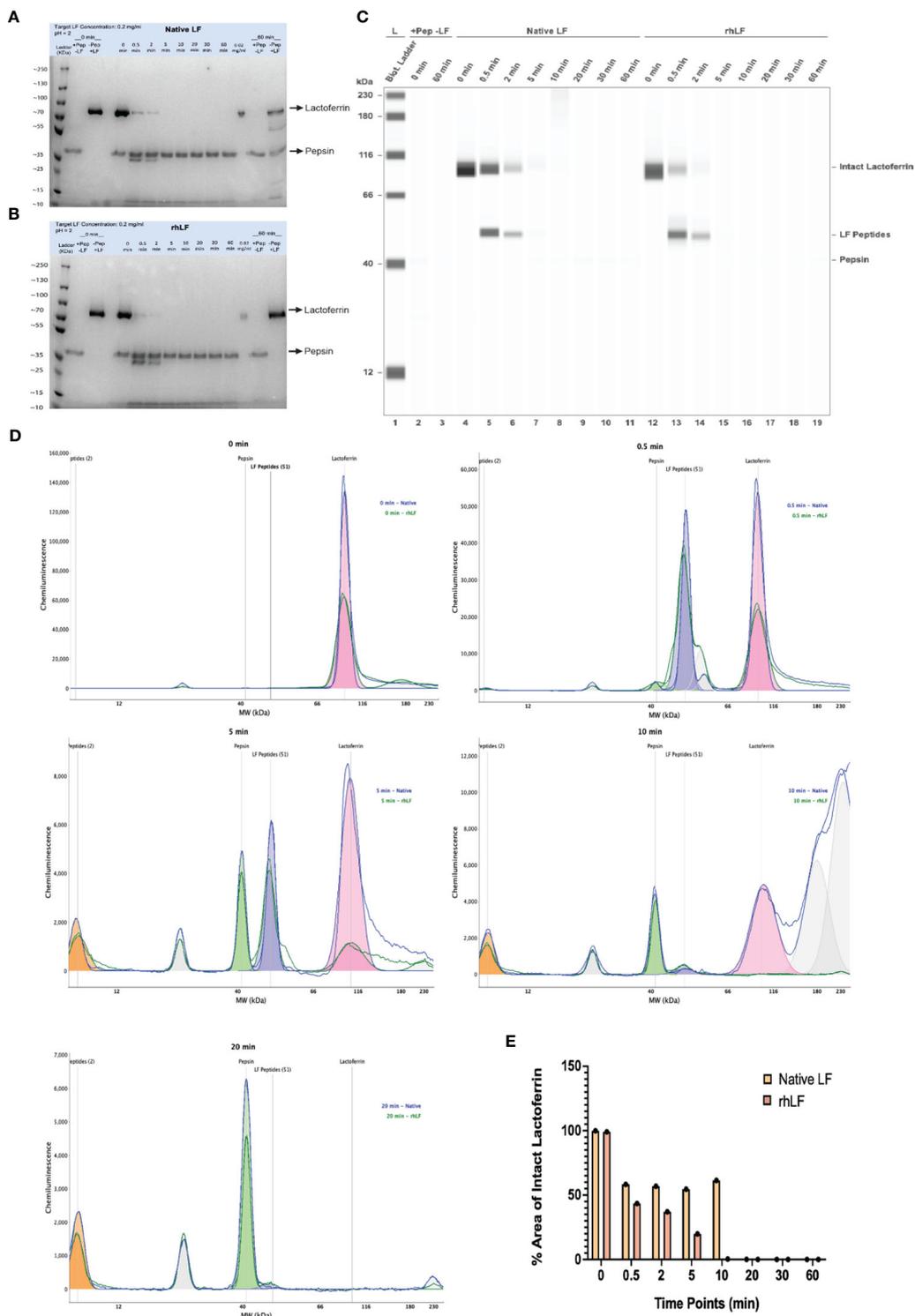


FIGURE 2

Helaina rhLF digests in a comparable manner to native human LF and shows low allergenic risk potential in the pepsin digestion model. (A, B) Coomassie Brilliant Blue-stained SDS-PAGE gel showing the digestion of hLf (A) or rhLF (B) in SGF at 10 units of pepsin per microgram of test protein at pH 2.0. Lane 1, protein ladder; lane 2, pepsin control (no LF) at 0 min; lane 3, protein control without pepsin at 0 min; lanes 4–12, digestion at 0, 0.5, 2, 5, 10, 20, 30, and 60 min; lane 13, protein at 10% concentration (0.02 mg/mL); lane 14, pepsin only control (no LF) at 60 min; lane 15, protein control without pepsin at 60 min. (C) Jess™ gel blot showing the digestion results for native hLf and rhLF. Lane 1, protein ladder; lane 2, pepsin control (no LF) at 0 min; lane 3, pepsin control (no LF) at 60 min; lanes 4–11, hLf digestion at 0, 0.5, 2, 5, 10, 20, 30, and 60 min; lanes 12–19, rhLF digestion at 0, 0.5, 2, 5, 10, 20, 30, and 60 min. (D, E) Jess™ histogram of rhLF and hLf following digestion at 0, 0.5, 5, 10, and 20 min. (F) Quantification of intact proteins of hLf and rhLF following digestion at 0, 0.5, 2, 5, 10, 20, 30, and 60 min from the Jess™ method.

foods with soy leghemoglobin are labeled as a soy allergen due to regulatory requirements) and residual host cell proteins and was determined to be GRAS in 2018 (GRN737). Additionally, β -lactoglobulin was evaluated for allergenic risk regarding the host cell proteins, showed no increased risk, and was determined to be GRAS in 2023 (GRN1056) (21). All three proteins—milk basic protein, soy leghemoglobin, and β -lactoglobulin—received “no questions” letters from FDA for use in various food products. This supports that the allergenicity assessment conducted on rhLF is sufficient to assess the food allergy risk of a novel protein intended for use as a food ingredient. To further corroborate this work and modeling, a recently published safety roadmap has been developed by the authors and vetted by experts in the fields of immunology/allergy, immunotoxicology, food safety, and lactoferrin biology and nutrition to thoroughly characterize rhLF for safety to provide support for repeated ingestion and long-term use (50).

Although the existing evidence suggests a low allergenic risk potential of Helaina rhLF, EfferTM, it cannot prove a nonexistent allergy risk. Nonetheless, these results do not support the need for further clinical testing or serum IgE binding to evaluate Helaina rhLF for risk of food allergy.

Data availability statement

The data presented in the study are deposited in the ProteomeXchange repository, accession number PXD053985.

Author contributions

YA: Conceptualization, Formal analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing. RRM: Formal analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing. RG: Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. PJ: Conceptualization, Formal analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing. SV: Formal analysis, Methodology, Visualization, Writing – review & editing. XL: Conceptualization, Methodology, Writing – review & editing. RP: Conceptualization, Writing – original draft, Writing – review & editing. SW: Visualization, Writing – original draft, Writing –

review & editing. BB: Formal analysis, Methodology, Writing – review & editing. KN: Formal analysis, Methodology, Writing – review & editing. BS: Methodology, Resources, Writing – review & editing. AC: Conceptualization, Resources, Writing – review & editing. CM: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Privately funded by Helaina, Inc. The funder had no involvement in the study design, collection, analysis, interpretation of the data, the writing of the article or the decision to submit for publication.

Conflict of interest

Authors YA, RR, SV, XL, BB, KN, RP, AC, and C-AM are all employees of Helaina, Inc. RG is an employee of RE Goodman Consulting. Corresponding author C-AM is a review editor for *Frontiers in Immunology*.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2024.1380028/full#supplementary-material>

References

- Lönnerdal B. Nutritional roles of lactoferrin. *Curr Opin Clin Nutr Metab Care*. (2009) 12:293–7. doi: 10.1097/MCO.0b013e328328d13e
- Yang Z, Jiang R, Chen Q, Wang J, Duan Y, Pang X, et al. Concentration of lactoferrin in human milk and its variation during lactation in different Chinese populations. *Nutrients*. (2018) 10:1235. doi: 10.3390/nu10091235
- Sill C, Biehl R, Hoffmann B, Radulescu A, Appavou M-S, Farago B, et al. Structure and domain dynamics of human lactoferrin in solution and the influence of Fe(III)-ion ligand binding. *BMC Biophys*. (2016) 9:7. doi: 10.1186/s13628-016-0032-3
- US FDA. *GRN No. 67. Milk-Derived Lactoferrin*. (2001). Available online at: https://www.cfsanappsexternal.fda.gov/scripts/fdcc/?set=GRASNotices&id=67&sort=GRN_No&order=DESC&startrow=1&type=basic&search=lactoferrin.
- US FDA. *GRN No. 77. Milk-Derived Lactoferrin* (2001). Available online at: https://www.cfsanappsexternal.fda.gov/scripts/fdcc/?set=GRASNotices&id=77&sort=GRN_No&order=DESC&startrow=1&type=basic&search=lactoferrin.
- US FDA. *GRN No. 130. Bovine Milk-Derived Lactoferrin*. (2004). Available online at: https://www.cfsanappsexternal.fda.gov/scripts/fdcc/?set=GRASNotices&id=130&sort=GRN_No&order=DESC&startrow=1&type=basic&search=lactoferrin.

7. US FDA. GRN No. 423. *Cow's Milk-Derived Lactoferrin*. (2012). Available online at: https://www.cfsanappsexternal.fda.gov/scripts/fdcc/?set=GRASNotices&id=423&sort=GRN_No&order=DESC&startrow=1&type=basic&search=lactoferrin.
8. US FDA. GRN No. 464. *Cow's Milk-Derived Lactoferrin*. (2014). Available online at: https://www.cfsanappsexternal.fda.gov/scripts/fdcc/?set=GRASNotices&id=464&sort=GRN_No&order=DESC&startrow=1&type=basic&search=lactoferrin.
9. US FDA. GRN No. 465. *Cow's Milk-Derived Lactoferrin*. (2014). Available online at: https://www.cfsanappsexternal.fda.gov/scripts/fdcc/?set=GRASNotices&id=465&sort=GRN_No&order=DESC&startrow=1&type=basic&search=lactoferrin.
10. US FDA. GRN No. 669. *Cow's Milk-Derived Lactoferrin*. (2017). Available online at: https://www.cfsanappsexternal.fda.gov/scripts/fdcc/?set=GRASNotices&id=669&sort=GRN_No&order=DESC&startrow=1&type=basic&search=lactoferrin.
11. Superti F. Lactoferrin from bovine milk: A protective companion for life. *Nutrients*. (2020) 12:2562. doi: 10.3390/nu12092562
12. Lönnnerdal B, Jiang R, Du X. Bovine lactoferrin can be taken up by the human intestinal lactoferrin receptor and exert bioactivities. *J Pediatr Gastroenterol Nutr*. (2011) 53:606–14. doi: 10.1097/MPG.0b013e318230a419
13. Ma Y, Hou Y, Xie K, Zhang L, Zhou P. Digestive differences in immunoglobulin G and lactoferrin among human, bovine, and caprine milk following in vitro digestion. *Int Dairy J*. (2021) 120. doi: 10.1016/j.idairyj.2021.105081
14. Furlund CB, Ulleberg EK, Devold TG, Flengsrud R, Jacobsen M, Sekse C, et al. Identification of lactoferrin peptides generated by digestion with human gastrointestinal enzymes. *J Dairy Sci*. (2013) 96:75–88. doi: 10.3168/jds.2012-5946
15. Huang N, Bethell D, Card C, Cornish J, Marchbank T, Wyatt D, et al. Bioactive recombinant human lactoferrin, derived from rice, stimulates mammalian cell growth. *In Vitro Cell Dev Biol Anim*. (2008) 44:464–71. doi: 10.1007/s11626-008-9136-7
16. Wang M, Sun Z, Yu T, Ding F, Li L, Wang X, et al. Large-scale production of recombinant human lactoferrin from high-expression, marker-free transgenic cloned cows. *Sci Rep*. (2017) 7:10733. doi: 10.1038/s41598-017-11462-z
17. Kruzel ML, Actor JK, Zimecki M, Wise J, Ploszaj P, Mirza S, et al. Novel recombinant human lactoferrin: differential activation of oxidative stress related gene expression. *J Biotechnol*. (2013) 168:666–75. doi: 10.1016/j.jbiotec.2013.09.011
18. Kurtzman CP. Description of *Komagataella phaffii* sp. nov. and the transfer of *Pichia pseudopastoris* to the methylotrophic yeast genus *Komagataella*. *Int J Syst Evol Microbiol*. (2005) 55:973–6. doi: 10.1099/ijs.0.63491-0
19. Ahmad M, Hirz M, Pichler H, Schwab H. Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol*. (2014) 98:5301–17. doi: 10.1007/s00253-014-5732-5
20. Karbalaei M, Rezaee SA, Farsiani H. *Pichia pastoris*: A highly successful expression system for optimal synthesis of heterologous proteins. *J Cell Physiol*. (2020) 235:5867–81. doi: 10.1002/jcp.29583
21. US FDA. GRN No. 1056. *β-Lactoglobulin Produced by Komagataella Phaffii Strain "yRMK-66"*. (2023). Available online at: <https://www.cfsanappsexternal.fda.gov/scripts/fdcc/index.cfm?set=GRASNotices&id=1056> (Accessed November 29, 2023).
22. US FDA. GRN No. 967. *Soluble Egg-White Protein Produced by Komagataella Phaffii Strain GSD-1209*. (2021). Available online at: https://www.cfsanappsexternal.fda.gov/scripts/fdcc/?set=GRASNotices&id=967&sort=GRN_No&order=DESC&startrow=1&type=basic&search=clara (Accessed November 29, 2023).
23. US FDA. GRN No. 1001. *Myoglobin preparation from a strain of Pichia pastoris expressing the myoglobin gene from Bos taurus*. (2021). Available online at: <https://www.cfsanappsexternal.fda.gov/scripts/fdcc/index.cfm?set=GRASNotices&id=1001> (Accessed November 29, 2023).
24. US FDA. GRN No. 737. *Soy Leghemoglobin Preparation from a Strain of Pichia Pastoris*. (2018). Available online at: https://www.cfsanappsexternal.fda.gov/scripts/fdcc/index.cfm?set=GRASNotices&id=737&sort=GRN_No&order=DESC&startrow=1&type=basic&search=impossible (Accessed December 4, 2023).
25. Department of Health and Human Services, Food and Drug Administration. *Food Additives Permitted in Feed and Drinking Water of Animals* (2023). Available online at: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=573.750#:-:text=The%20food%20additive%20Pichia%20pastoris,weight%20of%20the%20total%20formulation> (Accessed December 4, 2023).
26. Lu X, Cummings C, Osuala UA, et al. Characterization of recombinant human lactoferrin expressed in *Komagataella phaffii*. (2024). doi: 10.1101/2024.01.09.574900
27. Jacobs PP, Geysens S, Vervecken W, Contreras R, Callewaert N. Engineering complex-type N-glycosylation in *Pichia pastoris* using GlycoSwitch technology. *Nat Protoc*. (2009) 4:58–70. doi: 10.1038/nprot.2008.213
28. Katz L, Botero Besada-Lombada P. *Methods and compositions for protein synthesis and secretion*. Available online at: [https://patents.google.com/patent/CA3187918A1/en?q=\(botero+besada\)&inventor=lombana&assignee=pamela&dq=pamela+botero+besada+lombana](https://patents.google.com/patent/CA3187918A1/en?q=(botero+besada)&inventor=lombana&assignee=pamela&dq=pamela+botero+besada+lombana) (Accessed November 12, 2023).
29. Zlatina K, Galuska SP. The N-glycans of lactoferrin: more than just a sweet decoration. *Biochem Cell Biol*. (2021) 99:117–27. doi: 10.1139/bcb-2020-0106
30. Codex Alimentarius Commission. *Guideline For The Conduct of Food Safety Assessment of Foods Produced Using Recombinant-DNA Microorganisms* (2003). Available online at: https://www.fao.org/fao-who-codexalimentarius/sh-proxy/en/?lnk=1&url=https%253A%252F%252Fworkspace.fao.org%252Fsites%252Fcodex%252FStandards%252FCXG%2B46-2003%252FCXG_046e.pdf (Accessed December 6, 2023).
31. Ladics GS. Current codex guidelines for assessment of potential protein allergenicity. *Food Chem Toxicol Int J Publ Br Ind Biol Res Assoc*. (2008) 46 Suppl 10:S20–23. doi: 10.1016/j.fct.2008.07.021
32. Goodman RE, Hefle SL, Taylor SL, van Ree R. Assessing genetically modified crops to minimize the risk of increased food allergy: a review. *Int Arch Allergy Immunol*. (2005) 137:153–66. doi: 10.1159/000086314
33. Goodman RE, Vieths S, Sampson HA, Hill D, Ebisawa M, Taylor SL, van Ree R. Allergenicity assessment of genetically modified crops—what makes sense? *Nat Biotechnol*. (2008) 26:73–81. doi: 10.1038/nbt1343
34. Zhou C, Sun N. Allergenicity assessment of a genetically modified protein-recombinant human lactoferrin. *J Allergy Ther*. (2013) S3. doi: 10.4172/2155-6121.S3-002
35. Jin Y, He X, Andoh-Kumi K, Fraser RZ, Lu M, Goodman RE. Evaluating potential risks of food allergy and toxicity of soy leghemoglobin expressed in *pichia pastoris*. *Mol Nutr Food Res*. (2018) 62:1700297. doi: 10.1002/mnfr.201700297
36. Astwood JD, Leach JN, Fuchs RL. Stability of food allergens to digestion *in vitro*. *Nat Biotechnol*. (1996) 14:1269–73. doi: 10.1038/nbt1096-1269
37. Thomas K, Aalbers M, Bannon GA, Bartels M, Dearman RJ, Esdaile DJ, et al. A multi-laboratory evaluation of a common *in vitro* pepsin digestion assay protocol used in assessing the safety of novel proteins. *Regul Toxicol Pharmacol RTP*. (2004) 39:87–98. doi: 10.1016/j.yrtph.2003.11.003
38. Ofori-Anti AO, Ariyaratna H, Chen L, Lee HL, Pramod SN, Goodman RE. Establishing objective detection limits for the pepsin digestion assay used in the assessment of genetically modified foods. *Regul Toxicol Pharmacol RTP*. (2008) 52:94–103. doi: 10.1016/j.yrtph.2008.06.006
39. Zhang J, Xin L, Shan B, Chen W, Xie M, Yuen D, et al. PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol Cell Proteomics MCP*. (2012) 11:M111.010587. doi: 10.1074/mcp.M111.010587
40. Goodman RE, Ebisawa M, Ferreira F, Sampson HA, van Ree R, Vieths S, et al. AllergenOnline: A peer-reviewed, curated allergen database to assess novel food proteins for potential cross-reactivity. *Mol Nutr Food Res*. (2016) 60:1183–98. doi: 10.1002/mnfr.201500769
41. Jenkins JA, Breiteneder H, Mills ENC. Evolutionary distance from human homologs reflects allergenicity of animal food proteins. *J Allergy Clin Immunol*. (2007) 120:1399–405. doi: 10.1016/j.jaci.2007.08.019
42. Liu WC, Inwood S, Gong T, Sharma A, Yu LY, Zhu P. Fed-batch high-cell-density fermentation strategies for *Pichia pastoris* growth and production. *Crit Rev Biotechnol*. (2019) 39:258–71. doi: 10.1080/07388551.2018.1554620
43. Kobayashi K, Kuwae S, Ohya T, Ohda T, Ohyama M, Tomomitsu K. High level secretion of recombinant human serum albumin by fed-batch fermentation of the methylotrophic yeast, *Pichia pastoris*, based on optimal methanol feeding strategy. *J Biosci Bioeng*. (2000) 90:280–8. doi: 10.1016/S1389-1723(00)80082-1
44. Kale A, Patnaik R. *Recombinant protein recovery methods* (2023). Available online at: [https://patents.google.com/patent/WO2023192957A1/en?q=\(egg+white+protein+clara+foods+pichia\)&dq=egg+white+protein+clara+foods+pichia&page=2](https://patents.google.com/patent/WO2023192957A1/en?q=(egg+white+protein+clara+foods+pichia)&dq=egg+white+protein+clara+foods+pichia&page=2) (Accessed April 23, 2024).
45. Mari A, Ooievaar-de Heer P, Scala E, Giani M, Pirrotta L, Zuidmeer L, et al. Evaluation by double-blind placebo-controlled oral challenge of the clinical relevance of IgE antibodies against plant glycans. *Allergy*. (2008) 63:891–6. doi: 10.1111/j.1398-9995.2008.01703.x
46. Kasahara A, Kita K, Tomita E, Toyota J, Imai Y, Kumada H. Repeated administration of recombinant human serum albumin caused no serious allergic reactions in patients with liver cirrhosis: a multicenter clinical study. *J Gastroenterol*. (2008) 43:464–72. doi: 10.1007/s00535-008-2178-5
47. Kreer C, Kuepper JM, Zehner M, Quast T, Kolanus W, Schumak B, et al. N-glycosylation converts non-glycoproteins into mannose receptor ligands and reveals antigen-specific T cell responses *in vivo*. *Oncotarget*. (2017) 8:6857–72. doi: 10.18632/oncotarget.v8i4
48. Platts-Mills TAE, Li RC, Keshavarz B, Smith AR, Wilson JM. Diagnosis and management of patients with the α-gal syndrome. *J Allergy Clin Immunol Pract*. (2020) 8:15–23.e1. doi: 10.1016/j.jaip.2019.09.017
49. Vervecken W, Callewaert N, Kaigorodov V, Geysens S, Contreras R. Modification of the N-glycosylation pathway to produce homogeneous, human-like glycans using glycoSwitch plasmids. In: Cregg JM, editor. *Pichia Protocols*. Humana Press (2007). p. 119–38. doi: 10.1007/978-1-59745-456-8_9
50. Malinczak CA, Burns Naas LA, Clark A, Conze D, DiNovi M, Kaminski N, et al. Workshop report: A study roadmap to evaluate the safety of recombinant human lactoferrin expressed in *komagataella phaffii* intended as an ingredient in conventional foods - recommendations of a scientific expert panel. *Food Chem Toxicol Int J Publ Br Ind Biol Res Assoc*. (2024) 190:114817. doi: 10.1016/j.fct.2024.114817
51. Van Petegem F, Contreras H, Contreras R, Van Beeumen J. Trichoderma reesei α-1,2-mannosidase: structural basis for the cleavage of four consecutive mannose residues. *J Mol Biol*. (2001) 312:157–65. doi: 10.1006/jmbi.2001.4946
52. Assarehzadegan MA, Sankian M, Jabbari F, Tehrani M, Falak R, Varasteh A. Identification of methionine synthase (Sal k 3), as a novel allergen of *Salsola kali* pollen. *Mol Biol Rep*. (2011) 38:65–73. doi: 10.1007/s11033-010-0078-2

53. Dupont D, Mandalari G, Molle D, Jardin J, Léonil J, Faulks RM, et al. Comparative resistance of food proteins to adult and infant *in vitro* digestion models. *Mol Nutr Food Res.* (2010) 54:767–80. doi: 10.1002/mnfr.200900142
54. Akkerdaas JH, Cianferoni A, Islamovic E, Kough J, Ladics GS, McClain S, et al. Impact of food matrices on digestibility of allergens and poorly allergenic homologs. *Front Allergy.* (2022) 3:909410. doi: 10.3389/falgy.2022.909410
55. Mandalari G, Adel-Patient K, Barkholt V, Baro C, Bennett L, Bublin M, et al. *In vitro* digestibility of beta-casein and beta-lactoglobulin under simulated human gastric and duodenal conditions: a multi-laboratory evaluation. *Regul Toxicol Pharmacol RTP.* (2009) 55:372–81. doi: 10.1016/j.yrtph.2009.08.010
56. Wang L, Wang YL, Lv ZL, Zhang E-p, Guo A-z. Design of bovine lactoferricin-derived peptide and its expression and activity in *Pichia pastoris*. *Biochem Biophys Res Commun.* (2021) 534:822–9. doi: 10.1016/j.bbrc.2020.10.098
57. Gaudin J-C, Rabesona H, Choiset Y, Yeretsian G, Chobert J -M., Sakanyan V, et al. Assessment of the immunoglobulin E-mediated immune response to milk-specific proteins in allergic patients using microarrays. *Clin Exp Allergy.* (2008) 38:686–93. doi: 10.1111/j.1365-2222.2008.02952.x
58. Perusko M, Apostolovic D, Kiewiet MBG, Grundström J, Hamsten C, Starkhammar M, et al. Bovine γ -globulin, lactoferrin, and lactoperoxidase are relevant bovine milk allergens in patients with α -Gal syndrome. *Allergy.* (2021) 76:3766–75. doi: 10.1111/all.14889
59. Mine Y, Zhang JW. Comparative studies on antigenicity and allergenicity of native and denatured egg white proteins. *J Agric Food Chem.* (2002) 50:2679–83. doi: 10.1021/jf0112264
60. Jacobsen B, Hoffmann-Sommergruber K, Have TT, Foss N, Briza P, Oberhuber C, et al. The panel of egg allergens, Gal d 1-Gal d 5: Their improved purification and characterization. *Mol Nutr Food Res.* (2008) 52 Suppl 2:S176–185. doi: 10.1002/mnfr.200700414
61. Al-Ghoulh A, Johal R, Sharquie IK, Emara M, Harrington H, Shakib F, et al. The glycosylation pattern of common allergens: the recognition and uptake of der p 1 by epithelial and dendritic cells is carbohydrate dependent. *PLoS One.* (2012) 7:e33929. doi: 10.1371/journal.pone.0033929
62. Royer PJ, Emara M, Yang C, Al-Ghoulh A, Tighe P, Jones N, et al. The mannose receptor mediates the uptake of diverse native allergens by dendritic cells and determines allergen-induced T cell polarization through modulation of IDO activity. *J Immunol.* (2010) 185:1522–31. doi: 10.4049/jimmunol.1000774
63. Abdelmoteleb M, Zhang C, Furey B, Kozubal M, Griffiths H, Champeaud M, et al. Evaluating potential risks of food allergy of novel food sources based on comparison of proteins predicted from genomes and compared to www.AllergenOnline.org. *Food Chem Toxicol Int J Publ Br Ind Biol Res Assoc.* (2021) 147:111888. doi: 10.1016/j.fct.2020.111888
64. Foster ES, Kimber I, Dearman RJ. Relationship between protein digestibility and allergenicity: comparisons of pepsin and cathepsin. *Toxicology.* (2013) 309:30–8. doi: 10.1016/j.tox.2013.04.011
65. Almond RJ, Flanagan BF, Kimber I, Dearman RJ. Influence of protein expression system on elicitation of IgE antibody responses: Experience with lactoferrin. *Toxicology.* (2012) 301:50–7. doi: 10.1016/j.tox.2012.06.018
66. Goodman RE, Taylor SL, Yamamura J, Kobayashi T, Kawakami H, Kruger CL, et al. Assessment of the potential allergenicity of a Milk Basic Protein fraction. *Food Chem Toxicol Int J Publ Br Ind Biol Res Assoc.* (2007) 45:1787–94. doi: 10.1016/j.fct.2007.03.014
67. US FDA. GRN No. 196. *Bovine Milk Basic Protein Fraction.* (2006). Available online at: https://www.cfsanappsexternal.fda.gov/scripts/fdcc/?set=GRASNotices&id=196&sort=GRN_No&order=DESC&startrow=1&type=basic&search=196.