Exponent
1150 Connecticut Ave., NW
Suite 1100
Washington, DC 20036
telephone 202-772-4900
facsimile 202-772-4979
www.exponent.com

February 28, 2024

Susan J. Carlson, Ph.D.
Director, Division of Food Ingredients
Office of Food Additive Safety (HFS-200)
Center for Food Safety and Applied Nutrition
U.S. Food and Drug Administration
5001 Campus Drive
College Park, MD 20740

Subject: GRAS Notice for serendipity berry sweet protein containing monellin expressed in *Komagataella phaffii*

Dear Dr. Carlson:

In accordance with 21 CFR part 170, subpart E, Oobli, Inc. hereby provides a notice of a claim that the food ingredient described in this submission is excluded from the premarket approval requirement of the Federal Food, Drug, and Cosmetic Act because the notifier has concluded such use to be generally recognized as safe (GRAS), based on scientific procedures.

The materials in this submission include Form 3667 and one complete electronic copy of the GRAS notice. If you have any questions or require additional information, please do not hesitate to contact me at 202-772-4953 or mmurphy@exponent.com.

Sincerely,

Mary M. Murphy, MS, RD
Principal Scientist

February 22, 2024

Jessica Bernhardt
Electronic Submissions Gateway
U.S. Food and Drug Administration
3WFN, Room 7C34
12225 Wilkins Avenue
Rockville, MD 20852

Re: Authorization Letter

To whom it may concern:

Pursuant to Section 11.100 of Title 21 of the Code of Federal Regulations, please accept this Authorization Letter. This letter is to certify that Oobli, Inc. authorizes Exponent, Inc. to submit in the Electronic Submissions Gateway on behalf of Oobli, Inc.

Sincerely yours,

DocuSigned by:
[redacted]
D6AD848049D24FB...

Dr. Jason Ryder, PhD.
Chief Technology Officer
Oobli, Inc.

GRAS Notice for serendipity berry sweet protein containing monellin expressed in *Komagataella phaffii*

PREPARED FOR:

Oobli, Inc.
202 Cousteau Place
Suite 210
Davis, CA 95618

SUBMITTED TO:

U.S. Food and Drug Administration
Center for Food Safety and Applied Nutrition
Office of Food Additive Safety
5001 Campus Drive
College Park, MD 20740

PREPARED BY AND CONTACT FOR TECHNICAL OR OTHER INFORMATION:

Exponent, Inc.
1150 Connecticut Avenue, NW
Washington, DC 20036

February 28, 2024

Table of Contents

List of Tables

List of Figures

List of Acronyms

%	percent
°C	degrees Celsius
µg	microgram
ATCC	American Type Culture Collection
BSA	bovine serum albumin
CFR	Code of Federal Regulations
CFU	colony forming unit
cGMP	current good manufacturing practice
DM	dry matter
EC	European Commission
EDI	estimated daily intake
EFSA	European Food Safety Authority
EU	European Union
FDA	U.S. Food and Drug Administration
FNDDS	Food and Nutrient Database for Dietary Studies
g	gram
GMP	good manufacturing practices
GRAS	Generally Recognized As Safe
GRN	GRAS Notification
h	hour
JECFA	Joint FAO/WHO Expert Committee on Food Additives
kcal	kilocalorie
kg	kilogram
kJ	kilojoule
L	liter
LOD	limit of detection
m	meter
mg	milligram
mL	milliliter
MON	monellin
MWt	molecular weight
NCBI	National Center for Biotechnology Information
NOAEL	no observed adverse effect level
PCR	polymerase chain reaction
QPS	Qualified Presumption of Safety
SBSP	serendipity berry sweet protein
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
U.S.	United States

USDA	United States Department of Agriculture
WHO/IUIS	World Health Organization and International Union of Immunological Societies
WT	wild-type
YPD	yeast extract peptone dextrose

Part 1. Signed Statements and Certification

1.1. Introduction

Oobli, Inc. submits to the U.S. Food and Drug Administration (FDA) this generally recognized as safe (GRAS) notice in accordance with 21 CFR part 170, subpart E.

1.2. Name and Address of Notifier

Oobli, Inc.
202 Cousteau Place
Suite 210
Davis, CA 95618

1.3. Name of GRAS Substance

The substance that is the subject of this GRAS notice is serendipity berry sweet protein (SBSP) containing single polypeptide chain monellin (MON) expressed in *Komagataella phaffii* (*K. phaffii*) and produced via precision fermentation.

1.4. Intended Conditions of Use

The intended use of SBSP containing MON expressed in *K. phaffii* and produced via precision fermentation is use as a general-purpose sweetener in foods at levels determined by good manufacturing practices (GMP). It is not intended for use in infant formula and meat and poultry products.

1.5. Basis for Conclusion of GRAS Status

Oobli, Inc.'s conclusion of GRAS status for the intended use of SBSP as a general-purpose sweetener in foods is based on scientific procedures in accordance with 21 CFR §170.30(a) and (b).

1.6. Pre-Market Approval Exclusion Claim

The intended use of SBSP containing MON expressed in *K. phaffii* and produced via precision fermentation is not subject to the pre-market approval requirements of the Federal Food, Drug, and Cosmetic Act because Oobli, Inc. has concluded that such use is GRAS through scientific procedures.

1.7. Availability of Information

The data and information that serve as the basis for this GRAS conclusion will be sent to the FDA upon request, or are available for the FDA's review and copying during customary business hours at the office of Mary Murphy, Principal Scientist, Exponent, Inc., located at 1150 Connecticut Ave, NW, Washington, DC 20036.

1.8. Exemptions from Disclosure

Our view is that none of the data and information in Parts 2 through 7 of the GRAS notice are exempt from disclosure under the Freedom of Information Act (FOIA).

1.9. Certification Statement

On behalf of Oobli, Inc., I hereby certify that, to the best of my knowledge, this GRAS notice is a complete, representative, and balanced submission that includes unfavorable, as well as favorable, information known to me and pertinent to the evaluation of the safety and GRAS status of the intended use of SBSP containing MON expressed in *K. phaffii* and produced via precision fermentation.

DocuSigned by:
[redacted]
D6AD848049D24FB...

2/28/2024

Jason Ryder, Ph.D.
Chief Technology Officer
Oobli, Inc.

Date

Part 2. Identity, Method of Manufacture, and Specifications

2.1. Common or Usual Name

The substance that is the subject of this dossier is serendipity berry sweet protein (SBSP). This ingredient is produced via precision fermentation from a genetically engineered strain of *Komagataella phaffii (K. phaffii)* to express a gene encoding for single polypeptide chain monellin (MON), which is the active constituent of SBSP.

2.2. Identity

Monellin is the name used to refer to a sweet tasting protein originally isolated from the fruit of *Dioscoreophyllum cumminsii (Stapf) Diels*. This plant is in the class Magnoliopsida, order Ranunculales and family *Menispermaceae* and is native to West Tropical Africa. Known locally as serendipity berry (Govaerts, 2021), *Dioscoreophyllum cumminsii* (serendipity berry) is a sweet, red and fleshy edible fruit. It is used traditionally in West African countries, both eaten raw and to sweeten food.

There were five sweet proteins originally named as monellins isolated from the West African serendipity berry (Fry, 2012). The main protein found in serendipity berry now known as 'monellin' is monellin 4, a protein of 10.7 kD, found at levels of 0.3-0.5% w/w in the fruit. Native monellin is composed of two, non-covalently associated polypeptide chains: monellin chain A and chain B with 44 and 50 amino acid residues, respectively (Kinghorn & Compardre, 2012; Agboola *et al*., 2014). Monellin is related to the cystatin family of thiol proteinase inhibitors. The cystatin superfamily of proteins is a family of reversible competitive inhibitors that bind reversibly to the widely distributed Cys proteases (Murzin, 1993). The three dimensional structure of monellin is characterized by a single alpha helix packed against five antiparallel β-strands in a β-grasp fold ('cystatin-fold') (Esposito & Temussi, 2011). Monellin is reported to provide 1500 to 3000 times the sweetness of sucrose on a weight basis (Kant, 2005; Fry, 2012). When heated above 50°C, monellin dissociates into two chains and loses sweetness (Kinghorn & Compardre, 2012; Agboola *et al*., 2014; Bilal *et al*., 2022; Saraiva *et al*., 2023). Single chain constructs, in which the two chains are linked, have considerable thermal stability while retaining the sweetening power of native monellin. MON produced as a component of SBSP links two monellin chains A and B together by a glycine with a molecular weight of 11.27 kD.

2.3. Production Process

SBSP is produced using precision fermentation of a genetically engineered production strain of *K. phaffii*. The 2-step production process includes a fermentation step to produce the protein, followed by a downstream processing step to isolate and purify the soluble protein. Upon reaching the end of fermentation, a solid-liquid separation process separates wet microbial biomass from the fermentation supernatant containing MON using a sequence of diafiltration and ultrafiltration steps. The resulting concentrated SBSP solution is then dried into a powder. SBSP consists of approximately 70% w/w total protein, of which >60% w/w is MON with the

remaining balance consisting of host cell proteins, 5% w/w ash, 15% w/w high molecular weight carbohydrates, and 10% w/w moisture. MON accounts for at least 50% w/w of the total dried powder as confirmed by high performance liquid chromatography (HPLC).

2.3.1. Host (Parental) Organism

The parental organism, *K. phaffii,* is an obligate aerobic yeast, which was previously recognized as *Pichia pastoris* (*P. pastoris*). All *P. pastoris* strains were moved to a new genus *Komagataella spp.* in 1995. Throughout this dossier, the organism is referenced as *K. phaffii.* The taxonomic identity of the yeast is shown in Table 1. *K. phaffii* is a nonpathogenic, non-toxigenic, and well-characterized yeast with a history of safe use in the food industry and a widely used candidate for the production of recombinant proteins (Barone *et al.,* 2023). Current laboratory strains of *K. phaffii* are from lineages isolated from oak and chestnut trees and were deposited in the culture collection at the Northern Regional Research Laboratories (NRRL).

Table 1. Taxonomic identity of *Komagataella phaffii*

Taxonomic Classification	Identity
Kingdom	Fungi
Phylum	Ascomycota
Class	Saccharomycetes
Order	Saccharomycetales
Family	Phaffomycetaceae
Genus	*Komagataella*
Species	*phaffii*
Strain	NRRL Y-11430 BG10

The parental strain is *K. phaffii* BG10 (produced by BioGrammatics, Inc.), which is derived from the well characterized host organism, *K. phaffii* NRRL Y-11430.

The Y-11430 strain was used by the company BioGrammatics (Carlsbad, CA, USA) to develop strain BG08 that was further engineered to create BG10 through the loss of endogenous plasmids. This work by BioGrammatics is described, along with the genome sequence for BG10 (Sturmberger, *et al*., 2016). *K. phaffii* BG10 has been demonstrated to be genomically similar to strain Y-11430 and does not contain any native plasmids or antibiotic resistance genes.

2.3.2. Construction of the Production Strain

P-MON-040 (hereinafter “the production strain”) was created through a series of transformations with different expression constructs to enable the biosynthesis of MON. The production strain contains copies of MON using a hybrid signal sequence for secretion under the control of a *K. phaffii* native promoter and a terminator. The MON gene sequence was synthesized *de novo* and codon optimized for *K. phaffii.* The MON amino acid sequence in SBSP is identical to the engineered single polypeptide chain monellin sequence published in the National Center for Biotechnology Information (NCBI) database as NCBI 1IV9_A (expressed in *E. coli*) and is

formed of monellin chain A (https://www.ncbi.nlm.nih.gov/protein/P02881.1) and monellin chain B (https://www.ncbi.nlm.nih.gov/protein/P02882.2) linked by a single glycine amino acid.

All sequences were synthesized and assembled into a cloning plasmid, where the expression cassette was transformed into *K. phaffii*. The MON cassettes were stably integrated into the genome and there were no plasmids or antibiotic resistance genes present in the production strain. The production strain has been confirmed for the absence of plasmids and antibiotic resistance genes. A glycerol stock of the production and wild-type strain (WT) was streaked on yeast extract peptone dextrose (YPD) agar plates with the antibiotic selection markers geneticin, zeocin, and hygromycin, and on YPD agar plates without antibiotics as a positive control. On the plates with the selection markers, there was no indication of growth of the production and WT strain, whereas growth was detected with the production and WT strain on YPD without selection markers (Figure 1). This indicated that the production strain does not contain any plasmids with selective genes. The absence of plasmids was also confirmed with colony polymerase chain reaction (PCR) using primers to amplify the selective genes used for plasmid selection (i.e., encoding for resistance to hygromycin, geneticin, or zeocin; Figure 2). Therefore, no plasmids or antibiotic resistance genes are expected to be transferred to non-related microorganisms or the final product.

Figure 1. Streaked YPD plates with hygromycin, zeocin, and geneticin selection markers with the production (P-MON-040) and wildtype (WT) *K. phaffii* strain

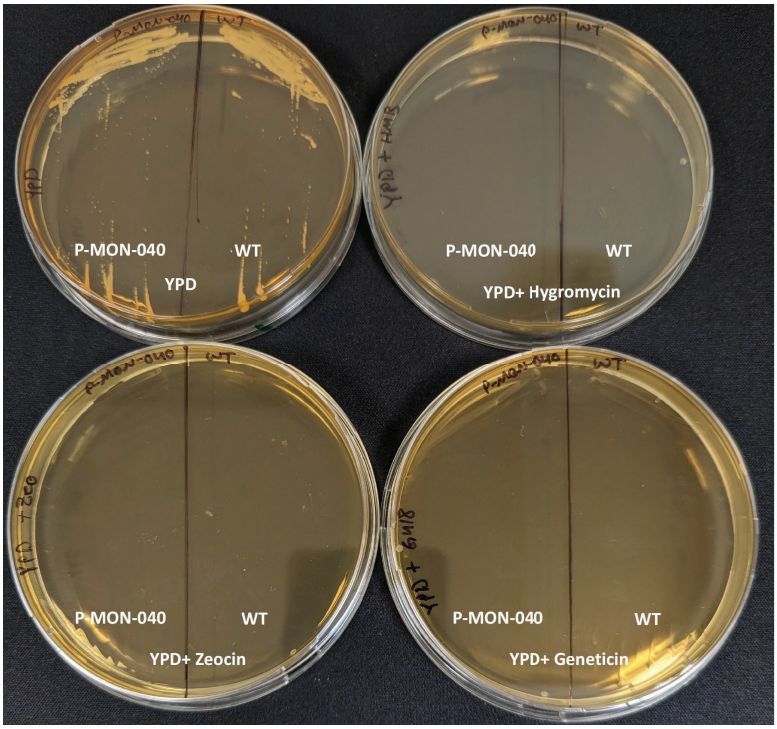
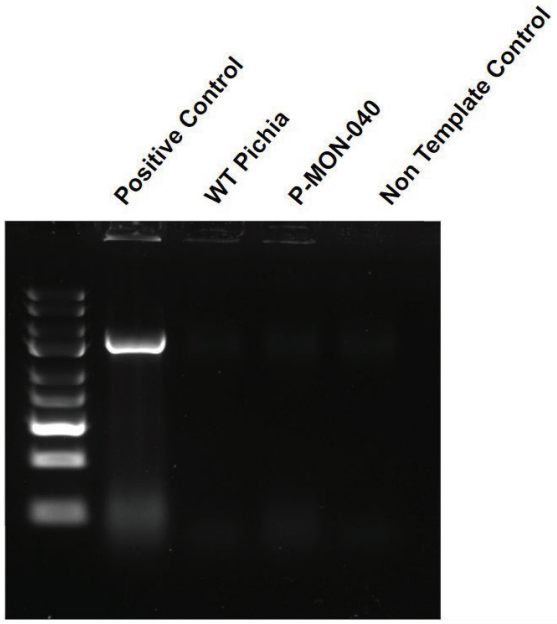


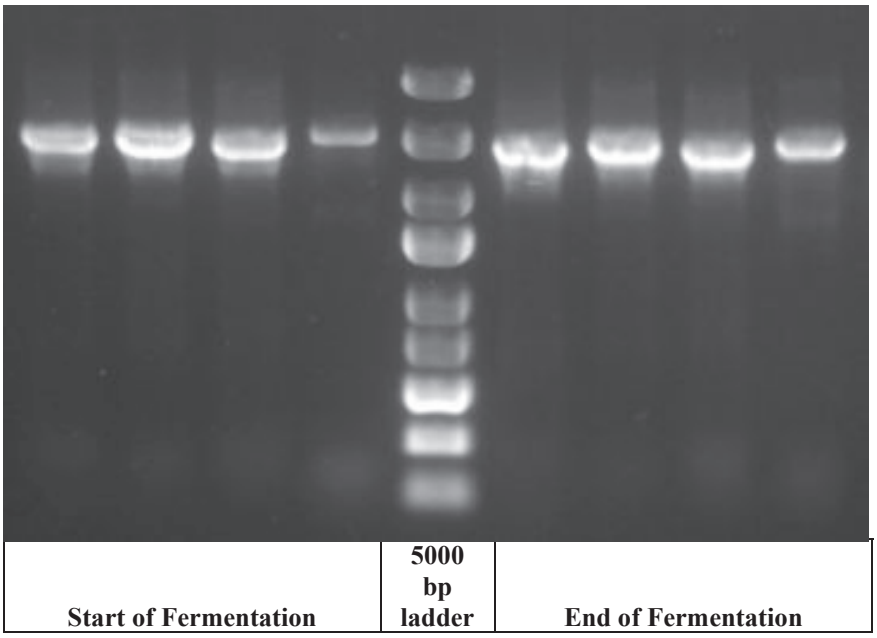
Figure 2. Evidence of absence of plasmids in the production strain (P-MON-040) via colony polymerase chain reaction (PCR)



2.3.3. Production Strain Stability

The production strain was evaluated via Sanger sequencing to ensure genomic stability throughout the fermentation process. The results produced by Sanger sequencing of these expression cassettes showed no point mutations for all copies of expression cassettes, thus demonstrating stability of the genome throughout the fermentation process. Samples were also taken at the start and end of fermentation and sent for colony PCR with primers flanking regions of representative expression cassettes for MON in the *K. phaffii* genome. The same expression cassettes were found to be present prior to and after fermentation (Figure 3).

Figure 3. Evidence of genetic stability of the production strain (P-MON-040) at the beginning and end of the fermentation process



2.3.4. Quality and Stability of Cell-line Stocks

Working cell banks are maintained for the production strain cell line at -80°C in 25% v/v glycerol as the source inoculum for production of SBSP containing MON. The working cell banks are tested for microbial purity, specific growth rate, and protein yield prior to the main fermentation. Continued cell line stability was demonstrated by using primary and secondary cell banks and comparing productivities. Extended seed trains were routinely tested to ensure retention of phenotype over generations of the production strain. Furthermore, the production strain was consistently tested for microbial contamination and strain performance according to reference standard operation procedures.

2.3.5. Production of SBSP

SBSP is produced via precision fermentation following current good manufacturing practices (cGMP) The fermentation and downstream processing steps are described below and shown in Figure 4.

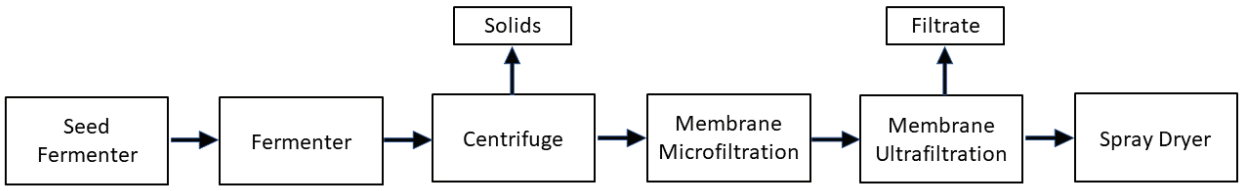
Step 1 - Fermentation

As a first step, inoculum is prepared from working cell banks using multiple seed propagation steps. Next, the resulting bolus of cells is used to inoculate the production fermentation process. Both the seed propagation and production fermentation steps are monitored for cell growth, culture purity, protein titer, glycerol, glucose, sulfate, ammonia, and ethanol. Process variables including pH, temperature, and dissolved oxygen are monitored and controlled throughout the precision fermentation process. Microbial contamination and other process deviations throughout the fermentation process that impact the safety and/or quality of the product are mitigated by incorporating critical control points involving sterilization by autoclave and discarding spent fermentation media to ensure the quality of the protein.

Step 2 - Downstream Processing

The process for recovery and purification of SBSP begins immediately after fermentation. First, a solid-liquid separation process separates the wet biomass from the fermentation supernatant containing MON. This supernatant is then concentrated, pH-adjusted, and further purified using a sequence of microfiltration and ultrafiltration steps. The liquid-concentrate SBSP solution is spray-dried into a powder consisting of >50% w/w MON, with the balance consisting principally of residual *K. phaffii* host cell proteins, carbohydrate, moisture, and ash.

Figure 4. Flow diagram for production of SBSP



2.3.6. Raw Materials and Processing Aids

All ingredients used in the production of SBSP are of food-grade quality or otherwise fit for their intended use and comply with appropriate specifications. The regulatory status of the ingredients and processing aids used in the production SBSP are provided in Table 2 below.

Table 2. Raw materials and processing aids used in the fermentation stage of SBSP production

Raw material / processing aid	Function / purpose	Regulatory Citation
Glucose (dextrose)	Nutrient source	21 CFR§184.1857
Glycerol (glycerin)	Nutrient source	21 CFR§182.1320
Biotin	Nutrient source	USP41 grade
Yeast extract 'Hy-yest 412'	Nutrient source	21 CFR§184.1983
Ammonium sulfate, anhydrous	Nutrient source	21 CFR§184.1143
Potassium phosphate, dibasic	Nutrient source	21 CFR§182.6285
Citric acid, anhydrous	Nutrient source	21 CFR§184.1033
Calcium sulfate, dihydrate	Nutrient source	21 CFR§184.1230
Magnesium sulfate, anhydrous	Nutrient source	21 CFR§184.1443
Potassium hydroxide	pH regulator	21 CFR§184.1631
Ammonium hydroxide solution	pH regulator	21 CFR§184.1139
Anti-foam	Processing aid	21 CFR§173.340
Boric acid	Trace element	NF grade
Sodium iodide	Trace element	USP grade
Sodium molybdate, dihydrate	Trace element	USP/NF grade
Manganese sulfate, monohydrate	Trace element	21 CFR§184.1461
Copper (II) sulfate, pentahydrate	Trace element	21 CFR§184.1261
Zinc chloride, anhydrous	Trace element	21 CFR§182.8985
Iron (II) sulfate, heptahydrate	Trace element	21 CFR§184.1315
Sulfuric acid (pH adjustment)	Processing aid	21 CFR§184.1095
Sodium acetate (buffering)	Processing aid	21 CFR§184.1721
Sodium chloride (buffering)	Processing aid	FCC, USP grade

2.4. Characterization of SBSP Produced by P-MON-040

SBSP produced by the production strain is composed of MON, residual *K. phaffii* host cell proteins remaining from the manufacturing process, high molecular weight carbohydrate, moisture, and ash. Investigations were performed to map MON produced by the production strain against the NCBI peptide sequence reference for single-chain monellin and to identify residual *K. phaffii* host cell proteins against the *K. phaffii* proteome from the manufacturing process of SBSP.

2.4.1. MON Produced by P-MON-040 vs Reference Monellin

The amino acid sequence of MON in SBSP is shown in Figure 5. The glycine linking the two polypeptide chains together is identified in red.

Figure 5. MON produced by P-MON-040

GEWEIIDIGPFTQNLGKFAVDEENKIGQYGRLTFNKVIRPCMKKTIYEENG
FREIKGYEYQLYVYASDKLFRADISEDYKTRGRKLLRFNGPVPPP

MON protein contains amino acid sequences identical to native monellin chain A (UniProt/Swiss-Prot P02881.1) and native monellin chain B (UniProt/Swiss-Prot P02882.2) coupled by an additional glycine residue (Table 3). The reference single-chain monellin NCBI 1IV9_A is a biosynthetic monellin derived from an *E.coli* system expression. In Table 3, peptide mapping showed there was a 100% sequence coverage, with 28 exclusive unique peptides identified matching the sequence for single-chain monellin NCBI 1IV9_A. The liquid chromatography-mass spectrometry-measured monoisotopic mass of Oobli's MON is 11,264.7939 Da and can be matched to the calculated mass of 11,264.790 Da for single-chain monellin NCBI 1IV9_A. These data indicate that the 96 amino acid MON produced from the production strain is identical to single-chain monellin NCBI 1IV9_A.

Table 3. Amino acid sequences for MON, native monellin chain A (P02881.1), native monellin chain B (P02882.2), and single-chain monellin NCBI 1IV_A

Protein	No. of Amino Acids	Amino acid sequence
MON	96	GEWEIIDIGPFTQNLGKFAVDEENKIGQYGRLTFNKVIRPCMKKTIYEENG FREIKGYEYQLYVYASDKLFRADISEDYKTRGRKLLRFNGPVPPP
Monellin chain A P02881.1	45	FREIKGYEYQLYVYASDKLFRADISEDYKTRGRKLLRFNGPVPPP
Monellin chain B P02882.2	50	GEWEIIDIGPFTQNLGKFAVDEENKIGQYGRLTFNKVIRPCMKKTIYEEN
Monellin NCBI 1IV9_A	96	GEWEIIDIGPFTQNLGKFAVDEENKIGQYGRLTFNKVIRPCMKKTIYEENG FREIKGYEYQLYVYASDKLFRADISEDYKTRGRKLLRFNGPVPPP

2.4.2. Proteomics Assessment of SBSP

In addition to MON protein, SBSP contains other residual proteins derived from *K. phaffii,* which are also expressed into the supernatant and co-purified along with MON. A proteomics assessment of SBSP was performed to investigate the identity of the host cell proteins originating from the production organism *K. phaffii*. Protein solutions (1 mg/mL) were prepared from three non-consecutive production batches of SBSP (referenced herein as Batch 1, Batch 2, and Batch 3). The samples were digested by trypsin followed by high resolution LC-MS/MS, where peptides were reassembled back into the complete protein sequence. Data were processed using MSFragger, and relative abundance of identified proteins was compared using total spectral counts. Of the 27 total proteins found present in three non-consecutive batches of SBSP through spectral analysis, 100% have been identified with UniProt accession numbers assigned and originate from *K. phaffii* (see data in Appendix A, Table 1). MON and six of the host cell

proteins found to be present in the final SBSP product at the cut-off of >1% account for 95.45% of total protein in the final SBSP product. The remaining 20 proteins account for <5% of total protein in the final SBSP product.

K. phaffii is recognized as a safe and suitable organism for production of food ingredients based on use in the production of ingredients recognized as GRAS (e.g., GRNs 1104, 1056, 1001, 697, and 737) and Qualified Presumption of Safety (QPS) status in the EU for use in enzyme production (EFSA, 2018); the organism is not capable of producing toxic metabolites when used for food protein production. While the SBSP consists predominately of MON and residual *K. phaffii* host cell proteins, other substances produced by the *K. phaffii* host strain may be present, though it is reasonable to assume that these substances would not be unique to the production of MON, and thus would be present in other ingredients produced by modified *K. phaffii*.

2.5. Product Specifications

Specifications and methods of analysis for SBSP produced from the production strain are presented in Table 4. All test methods are validated and fit for the intended purpose.

Table 4. Product specifications of SBSP

Parameter	Specification	Method of Analysis
Proximates		
Moisture, % w/w	<10	AOAC 950.46
Total protein, as is, % w/w	>70	AOAC 991.20
MON, % w/w	>50	Internal HPLC method
Fat by fatty acid profile, g/100g	<1	AOAC 996.06
Ash, % w/w	<5	AOAC 945.46
Carbohydrate, % w/w	<15	By calculation
Microbiological Analysis		
Aerobic plate count, CFU/g	<10,000	AOAC 990.12
Total Coliforms, CFU/g	<10	AOAC 991.14
E. coli, CFU/g	<10	AOAC 991.14
Yeast, CFU/g	<10	AOAC 2014.05
Mold, CFU/g	<10	AOAC 2014.05
Salmonella spp., absence/10g	Negative	AOAC 2011.03
Listeria monocytogenes, absence/10g	Negative	AOAC 2004.02
Heavy Metals		
Arsenic, ppm	<0.01 (LOQ 0.01)	AOAC 2015.01 (ICP MS)
Cadmium, ppm	<0.01 (LOQ 0.001)	AOAC 2015.01 (ICP MS)
Lead, ppm	<0.1 (LOQ 0.01)	AOAC 2015.01 (ICP MS)
Mercury, ppm	<0.01 (LOQ 0.005)	AOAC 2015.01 (ICP MS)

Abbreviations: AOAC = Association of Official Analytical Collaboration; CFU = colony-forming units; ICP MS = Inductively coupled plasma mass spectrometry; LOQ = limit of quantification; ppm = parts per million

Analytical data from three non-consecutive batches of SBSP are presented in Table 5. The SBSP product is primarily protein, with the balance composed of carbohydrate, moisture, ash, and fat. The ash content of the ingredient is composed of primarily sodium and lesser amounts of other minerals. The analytical data demonstrate that SBSP can consistently be manufactured so as to meet the established specifications.

Table 5. Analytical data from non-consecutive batches of SBSP

Parameter	Specification[a]	Batch 1	Batch 2	Batch 3
Proximates				
Moisture, % w/w	<10	4.52	4.66	6.04
Total protein, as is, % w/w	>70	89.53	93.87	86.57
MON, % w/w	>50	63	65	60
Fat by fatty acid profile, g/100g	<1	0.04[b]	0.02	0.1
Ash, % w/w	<5	1.53	0.04	3.02
Carbohydrate, % w/w	<15	4.38	1.41	4.27
Microbiological Analysis				
Aerobic plate count, CFU/g	<10,000	550	80	420
Total coliforms, CFU/g	<10	<10	<10	<10
E. coli, CFU/g	<10	<10	<10	<10
Yeast, CFU/g	<10	<10	<10	<10
Mold, CFU/g	<10	<10	<10	<10
Salmonella spp., absence/10g	Negative	Negative	Negative	Negative
Listeria monocytogenes, absence/10g	Negative	Negative	Negative	Negative
Heavy Metals				
Arsenic, ppm	<0.01 (LOQ 0.01)	<0.01	<0.01	<0.01
Cadmium, ppm	<0.01 (LOQ 0.001)	<0.001	0.001	<0.001
Lead, ppm	<0.1 (LOQ 0.01)	<0.01	0.06	0.06
Mercury, ppm	<0.01 (LOQ 0.005)	<0.005	0.008	<0.005

[a] LOQ – limit of quantification.
[b] Calculated value.

2.6. Stability

Physicochemical properties of sweeteners such as water solubility and stability across a range of temperatures and pH conditions are important for their usefulness in food technology applications. Stability of the final SBSP preparation is also important for food safety. Monellin is reported to be sensitive to denaturation at temperatures above 50°C (Kim *et al.,* 1989; Kinghorn and Compardre, 2012; Agboola *et al*., 2014; Bilal *et al*., 2022). Studies were undertaken to observe the impact of pH and temperature on protein denaturation and precipitation as well as retained sweetness. Additionally, a shelf-life study was undertaken to assess stability of the material over typical anticipated shelf life.

2.6.1. Effect of Temperature

Investigations were performed to evaluate the thermostability of SBSP by measuring protein concentration (MON) and relative sweetness under various temperature conditions. Aqueous SBSP solutions were held at temperatures ranging from 40°C to 85°C, and solid dry powder SBSP was held at 60°C and 120°C with assessments at 1, 3 and 6 h. Results from the testing detailed below demonstrate that MON protein in aqueous solution is sensitive to elevated temperature with loss of intact protein and lowered relative sweetness (Table 6). In contrast, the MON protein component of SBSP in solid powder remains intact with no impact on relative sweetness at the tested temperature ranges.

Table 6. Relative concentration of MON protein at different temperature timepoints in solid and aqueous phases

Heating time	Percent Protein (w/w with Respect to Untreated) in Aqueous Solution of SBSP						Percent Protein (w/w with Respect to Untreated) in Solid SBSP Powder	
	40°C	50°C	60°C	70°C	80°C	85°C	60°C	120°C
1 h	99%	93%	89%	86%	15%	0	92%	84%
3 h	92%	92%	91%	62%	0	0	97%	91%
6 h	94%	90%	88%	0	0	0	107%	85%

2.6.2. Effect of pH

The stability of MON in SBSP was tested in aqueous solutions with pH of 3, 5, 7 and 10 at ambient temperature (Table 7). Samples were taken to determine any significant differences in MON protein activity via HPLC protein quantification. Results show the average of 3 replicates per pH condition. MON in SBSP is stable from pH 3 to pH 7, while stability decreases by 50% at pH 10.

Table 7. Effect of pH on concentration of MON protein in SBSP

pH	MON Concentration (g/L)
Untreated (pH 6)	0.225
3	0.201
5	0.216
7	0.232
10	0.112

2.6.3. Shelf-Life Stability

An accelerated shelf-life stability test was conducted to assess shelf life of the product over a period of 12 months. The product was in finished product packaging and held at accelerated temperature conditions throughout the duration of the shelf-life study.

Following the baseline test point, samples were held at a temperature of 35ºC ± 2ºC and relative humidity of 65% ± 5%. At 0, 2, 4, 6, and 8 months, the SBSP samples were tested for aerobic plate count, yeasts and molds, total coliforms, *E. coli*, *Salmonella* spp, *Listeria monocytogenes*, pH, protein, moisture, and water activity. Results are presented below (Table 8). The study will continue, with tests at months 10 and 12.

Table 8. Shelf-life stability testing of SBSP

Parameter	Months of Testing				
	T=0	**T=2**	**T= 4**	**T=6**	**T= 8**
Proximates					
Moisture (loss on drying), % w/w	4.1	4.1	4.1	4.2	5.0
pH	4.02	4.01	4.21	4.01	4.24
Water activity	0.23	0.22	0.26	0.25	0.25
Total protein, as is, % w/w	96	95.7	94.4	95	94.7
MON, % w/w	65	65	65	69	65
Microbiological Analysis					
Aerobic plate count, CFU/g	120	50	50	<10	20
Total coliforms, CFU/g	<10	<10	<10	<10	<10
E. coli, CFU/g	<10	<10	<10	<10	<10
Yeast, CFU/g	<10	<10	<10	<10	<10
Mold, CFU/g	<10	<10	<10	<10	<10
Salmonella spp., absence/10g	Negative	Negative	Negative	Negative	Negative
Listeria monocytogenes, absence/10g	Negative	Negative	Negative	Negative	Negative
Sensory					
Sweetness potency (MON)	2898X	3072X	2686X	2750X	2405X

Note: Methods of analysis for proximates and microbiological analysis are provided in Table 4.

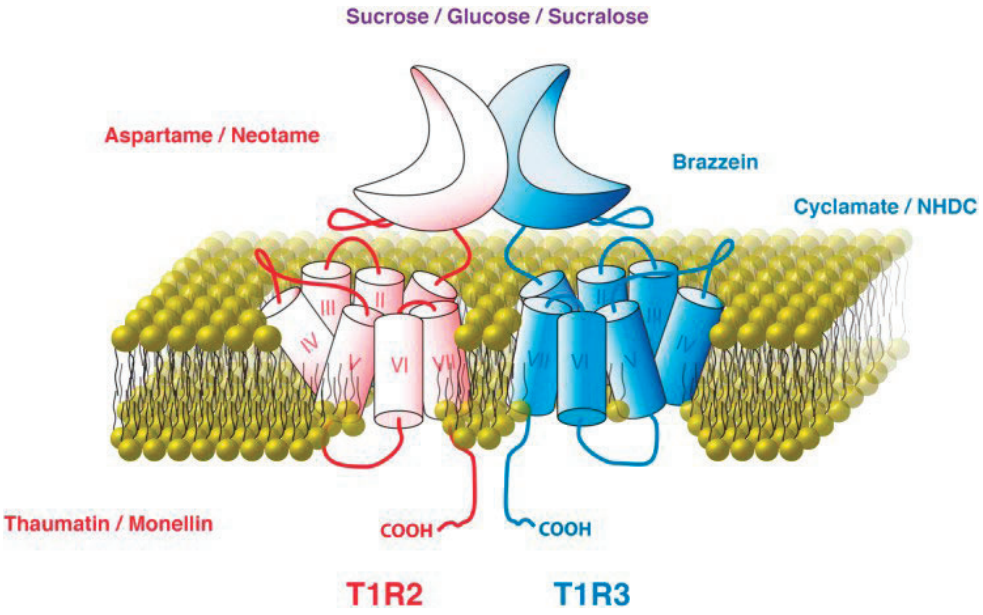
2.7. Technical Effect

2.7.1. Mechanism of Action

The receptor that mediates all sweet taste sensation is the T1R2/T1R3 receptor, a heterodimer composed of monomer units from the T1R family of G-protein coupled receptors (Smith *et al*., 2021; Fernstrom *et al.*, 2012). The overall architecture of heterodimer T1R2/T1R3 displays asymmetry, which accommodates sweeteners with widely different chemical structures, though not all sweeteners bind to the same sites on the receptor (Temussi, 2002; Smith *et al*., 2021;

Fernstrom *et al*., 2012). As shown in Figure 6, mono- and disaccharides including sucrose, glucose, and sucralose bind to areas of both the T1R2 and T1R3 receptors, whereas dipeptide sweeteners bind only to the T1R2 domain. Sweet proteins, including monellin, interact across a larger binding surface that can include both subunits and the cysteine-rich linkers (Temussi, 2002; Picone and Temussi, 2012; Fernstrom *et al*., 2012). No reports of sweet proteins, including monellin, resulting in taste desensitization or other effects on taste perception, were reported in the literature. Additionally, no reports of sweet proteins, including monellin, exerting effects on the physiology of gustatory cells were identified in the literature.

Figure 6. T1R2 and T1R3 and the compounds that can activate them



Source: Fernstrom *et al*., 2012. Font color indicates sweet compounds that bind T1R2 (red), T1R3 (blue), or both subunits (purple).

2.7.2. Sensory Studies on SBSP

The relative sweetness of aqueous and solid SBSP solution preparations was measured in sensory tests. Relative sweetness of SBSP was determined on a scale of 0 to 100, where a score of 100 indicates 100% sweetness equivalence with 5% w/w sucrose, and a score of 50 indicates 50% sweetness equivalence with 5% w/w sucrose. A threshold score of 35 was pre-defined as value at which SBSP would be considered inactive with respect to sweetness. Relative sweetness scores were obtained via sensory tests obtained from male and female adults from an internal testing team. Results from this organoleptic testing (n=10 subjects) shown in Table 9 demonstrate that the relative sweetness of SBSP remains organoleptically active with discernable sweet taste in aqueous solution up to 6 h at 70°C and in solid solution up to 6 h at 120°C.

Table 9. Relative sweetness of SBSP at different temperature timepoints in aqueous and solid phase

Heating time	Relative Sweetness in Aqueous Solution of SBSP						Relative Sweetness of Solid SBSP Powder	
	40°C	50°C	60°C	70°C	80°C	85°C	60°C	120°C
1 h	100.0	105.2	102.9	97.9	62.2	0	111.7	106.6
3 h	101.4	95.7	97.5	86.5	14.4	0	92.5	98.0
6 h	90.7	111.0	95.4	67.7	11.3	0	100.5	89.4

Relative sweetness on a scale of 0 to 100, where 100 = equivalent sweetness as 5% w/w sucrose, 50 = half as sweet as 5% w/w sucrose.

Part 3. Intended Use and Estimated Daily Intake

3.1. Proposed Use and Level

SBSP is intended for use as a general-purpose sweetener in accordance with cGMP, excluding infant formula and meat and poultry products.

3.2. Sweetness Intensity of SBSP

SBSP contains MON, which is a sweet protein. The sweetness intensity of SBSP has been tested using a sweet potency protocol designed to measure the sweetness response of an unknown sample compared to sucrose. Results from this testing demonstrate that MON is approximately 2700-fold sweeter than sucrose on a weight basis, which is consistent with reported sweetness potency for monellin in the literature (Kant, 2005; Fry, 2012). Given that SBSP is approximately 63% MON by weight, SBSP is approximately 1700-fold sweeter than sucrose.

3.3. Estimated Daily Intake

3.3.1. Approach

The estimated daily intake of SBSP from the intended use as a general-purpose sweetener in foods was derived using the replacement approach described by Renwick (2008). The approach detailed by Renwick provided estimated intakes of rebaudioside A, a new high-intensity sweetener at the time. Estimates of rebaudioside A intake were based on published data for intake of other high-intensity sweeteners among populations of children and adults with and without diabetes and the assumption that intake of the new sweetener would fully replace intake of existing high-intensity sweeteners. Replacement of high-intensity sweeteners was calculated based on sweetness intensities relative to sucrose.

The approach reported by Renwick was first used to support the exposure assessment for rebaudioside A in the GRAS conclusion submitted to FDA (GRN 253), to which FDA responded with a “no questions” letter. Since that time, Renwick’s replacement methodology has been used to develop estimated daily intakes (EDIs) to support numerous GRAS conclusions for use of rebaudioside A and related sweeteners providing steviol glycosides (e.g., most recently GRN 1106), as well as other high-intensity sweeteners derived from sources such as Luo Han Guo (e.g., GRN 359). FDA issued “no questions” letters to these GRAS notices.

To provide estimates of SBSP intake assuming use as a general-purpose sweetener, the Renwick (2008) approach was applied. The EDls of the MON component were calculated by dividing the estimated sucrose equivalent intakes by the relative sweetness of MON. The EDIs of SBSP were calculated by multiplying the EDI values for MON by the mean concentration of MON in SBSP based on representative data (i.e., 63%).

Consumption of serendipity berry could provide an additional source of exposure to monellin. The USDA FoodData Central repository (USDA, 2023) was searched using the term

"serendipity berry" and no hits were returned, indicating that serendipity berry is not represented as a food or an ingredient in foods contained within the database. Based on this review, there is no indication that serendipity berry is a component of the U.S. food supply at this time. Estimates of intake for this safety assessment thus reflect only the intended use of SBSP.

3.3.2. Intake of SBSP

Table 10 presents the Renwick estimates of high-intensity sweetener intake by population group along with the estimated intake of SBSP. The EDIs of SBSP for "high" consumers in populations of non-diabetic adults, diabetic adults, non-diabetic children, and diabetic children are up to 0.40 mg/kg bw/day, 0.53 mg/kg bw/day, 0.58 mg/kg bw/day, and 0.53 mg/kg bw/day, respectively.

Table 10. Estimated intake of SBSP for populations in the United States using an intense sweetener intake assessment methodology

Population Group	Intake of Intense Sweeteners (expressed as sucrose equivalents) (mg/kg bw/day) [a]		Estimated Intake of SBSP (mg/kg bw/day) [b]	
	Average Consumer	"High" Consumer ≥90th percentile	Average Consumer	"High" Consumer ≥90th percentile
Non-diabetic adults	255	675	0.15	0.40
Diabetic adults	280	897	0.16	0.53
Non-diabetic children	425	990	0.25	0.58
Diabetic children	672	908	0.40	0.53

[a] Estimates as reported in Table 7 in Renwick, 2008.
[b] Calculated assuming relative sweetness potency of 2700 for MON and assuming 63% MON in SBSP.

3.3.3. Intake of MON

The estimated intake of MON in SBSP is summarized in Table 11. The EDIs of MON for "high" consumers in populations of non-diabetic adults, diabetic adults, non-diabetic children, and diabetic children are up to 0.25 mg/kg bw/day, 0.33 mg/kg bw/day, 0.37 mg/kg bw/day, and 0.34 mg/kg bw/day, respectively.

Table 11. Estimated intake of MON for populations in the United States using an intense sweetener intake assessment methodology

Population Group	Intake of Intense Sweeteners (expressed as sucrose equivalents) (mg/kg bw/day) [a]		Estimated Intake of MON (mg/kg bw/day) [b]	
	Average Consumer	"High" Consumer ≥90th percentile	Average Consumer	"High" Consumer ≥90th percentile
Non-diabetic adults	255	675	0.09	0.25
Diabetic adults	280	897	0.10	0.33
Non-diabetic children	425	990	0.16	0.37
Diabetic children	672	908	0.25	0.34

[a] Estimates as reported in Table 7 in Renwick, 2008, assuming relative sweetness potency of 2700 for MON (see Table 10).
[b] Calculated assuming relative sweetness potency of 2700X for MON, and assuming 63% MON in SBSP (see Table 10).

Part 4. Self-Limiting Levels of Use

The use of MON in SBSP as a general-purpose sweetener is self-limiting due to increasingly unacceptable levels of sweetness and overall organoleptic profile.

Part 5. Experience Based on Common Use in Food Before 1958

General recognition of safety of the use of SBSP has been established through scientific procedures. Therefore, information regarding experience based on common use of the notified substance in food prior 1958 is not applicable.

Part 6. Narrative

6.1. Approach for Evaluating the Safety of SBSP Produced from Modified *K. phaffii*

The safety of SBSP produced through precision fermentation from a strain of *K. phaffii* genetically engineered to express a gene encoding for MON under the intended conditions of use can be evaluated through scientific procedures by considering safety of the native plant-derived monellin, safety of the host organism (*K. phaffii*) used to produce MON, safety of the production strain expressing MON, and safety of the SBSP product.

Literature searches as detailed in Appendix B were conducted to support the safety evaluation. Papers were screened at the title level. Abstracts of potentially relevant papers were reviewed and the relevant articles relating to history of consumption, physiological effects, and safety including allergenicity or hypersensitivity potential, were further evaluated. Bioinformatics analyses were performed to assess the likelihood of potential allergenicity and toxigenicity using sequence homology comparison.

6.2. Historical Exposure to Native Plant-Derived Serendipity Berry and Native Monellin

Dioscoreophyllum cumminsii (serendipity berry) is an edible sweet, red and fleshy fruit. Serendipity berries have been consumed by indigenous communities in West Africa for years, with the berries consumed both as a traditional food source and to sweeten cooked and raw food and beverages (Terashima & Ichikawa, 2003). The tubers of this plant can also be eaten cooked (roasted or boiled) like potato tubers and are often used to thicken soup (Agboola *et al.,* 2014; Govaerts *et al*., 2021).

Most articles in the published literature on monellin address monellin protein structure characterization, structural relationship with sweetness, studies on sweet taste intensity in various species, and the biotechnology developments of monellin manufacturing and production. Several articles reported studies with aqueous extracts from *Dioscoreophyllum cumminsii* leaves, but not from the berry itself in which monellin is naturally found. No estimates of daily consumption of serendipity berry or monellin were reported in in the literature, and no pre-clinical or clinical trials or other information on intake of serendipity berry or monellin in animals or humans were located.

6.3. Safety of the Production Organism

The safety of the production strain used to produce SBSP was assessed using the same principles applied for assessing the safety of microbially-derived enzymes for use in food production where a “safe food enzyme production strain” is defined as a non-pathogenic, non-toxigenic microbial strain with a demonstrated history of safe use in the production of food enzymes. Consideration

is therefore given to production strain history of safe use, pathogenicity, toxigenicity, and stability (Pariza & Johnson, 2001; FAO/WHO, 2020).

6.3.1. Production Strain

The production strain created to produce SBSP containing MON is an engineered strain of *K. phaffii,* an obligate aerobic yeast which was previously recognized as *Pichia pastoris*. All *P. pastoris* strains were moved to a new genus *Komagataella spp.* in 1995. The safety of the *K. phaffii* strain of yeast as an industrial enzyme production organism has been reviewed in many publications. *K. phaffii* is a nonpathogenic, non-toxigenic, and well-characterized yeast with a long history of safe use in industrial scale food enzyme production (Bernauer *et al*., 2021) and a widely used candidate for the production of recombinant proteins (Barone *et al.,* 2023). *K. phaffii* has QPS status in the European Union for production purposes only, where the qualification 'for production purpose only' implies the absence of viable cells of the production organism in the final product (EFSA, 2018). The American Type Culture Collection (ATCC) classified *K. phaffii* as Biosafety Level 1, indicating that it is well-characterized with no known pathogenic effect in healthy human adults and minimal safety precautions are required for handling and storage (ATCC, 2023). This classification offers an additional general indicator of lack of concern for health. *K. phaffii* does not produce any pathogens, toxins, or antibiotics under submerged fermentation conditions.

The World Health Organization and International Union of Immunological Societies (WHO/IUIS) allergen nomenclature database[1] for recognized allergens, and the allergen databases AllergenOnline,[2] Allergome,[3] or COMPARE[4] were searched for the terms *Komagataella phaffii* and *Pichia pastoris* as sources of allergens. No entries were found for allergens deriving from the source *K. phaffii* or *P. pastoris*. The only reference found were records where *P. pastoris* had been used as an expression vector for allergens from other sources for laboratory research purposes only. In addition, *K. phaffii* proteins specifically originating from the NRRL Y-11430 (also known as CBS 7435) organism have been demonstrated to lack allergenicity and toxigenicity potential (Jin *et al*., 2018; Reyes *et al*., 2021).

The production strain was created through a series of transformations with different expression constructs to enable the biosynthesis of MON. The MON gene sequence was synthesized *de novo*, codon optimized for *K. phaffii,* assembled into a cloning plasmid, and the resulting expression cassette was transformed into *K. phaffii*. Antibiotic selection media growth studies were performed to confirm no growth of the host strain or wild-type strain in the presence of antibiotics, and colony PCR confirmed the absence of plasmids or antibiotic resistance genes present in the production. No plasmids or antibiotic resistance genes are expected to be transferred to non-related microorganisms or the final product. The production strain's genomic stability was confirmed via Sanger sequencing which showed that the same expression cassettes were found present prior to and after fermentation and a lack of point mutations were found for all copies of expression cassettes. SBSP was produced by the production strain during the

[1] WHO/IUIS allergen database
[2] AllergenOnline
[3] Allergome
[4] COMPARE

fermentation processes. In line with safety guidelines, the materials added to support the fermentation process are all food grade and meet applicable regulatory standards.

The production strain is a *K. phaffii* strain that has been genetically engineered to express MON. The strain was derived from *K. phaffii* BG10 (produced by BioGrammatics, Inc.), which is derived from the well characterized host organism, *K. phaffii* NRRL Y-11430. The *K. phaffii* variant Y-11430 served as the host organism for the production strain used to produce soy leghemoglobin that is recognized as GRAS for its intended uses under GRN 737 (U.S. FDA, 2018b). In addition, the production strain was constructed in a manner similar to other strains of *K. phaffii* used to produce proteins for human consumption including soluble egg-white protein (GRN 967), myoglobin (GRN 1001), and β-lactoglobulin (GRN 1056). These proteins and enzyme preparations produced from *K. phaffii* have been concluded to be safe for the intended use and FDA had no questions regarding the safety conclusions (Table 12), which supports the safety of the intended use of the Y-11430 derived production strain to produce SBSP.

Table 12. Food ingredients produced with modified *Komagataella phaffii (Pichia pastoris)* strains concluded to be GRAS for the intended use

GRN No.	Substance	Date of closure	FDA's Letter
1104	Egg-white protein produced by *Komagataella phaffii* ATCC GSD-1235	Oct 17, 2023	FDA has no questions
1056	β-lactoglobulin produced by *Komagataella phaffii* strain "yRMK-66"	Feb 15, 2023	FDA has no questions
1025	Pepsin A enzyme preparation produced by Pichia pastoris to overexpress the gene encoding pepsin A	Jan 6, 2023	FDA has no questions
1001	Myoglobin preparation from a strain of *Pichia pastoris* expressing the myoglobin gene from *Bos taurus*	Dec 3, 2021	FDA has no questions
967	Soluble egg-white protein produced by *Komagataella phaffii* strain GSD-1209	Sep 9, 2021	FDA has no questions
737	Soy leghemoglobin preparation from a strain of *Pichia pastoris*	Jul 23, 2018	FDA has no questions
204	Phospholipase C enzyme preparation from *Pichia pastoris* expressing a heterologous phospholipase C gene	Dec 5, 2006	FDA has no questions

While the safety of *K. phaffii* for use in the production of food ingredients has been established in numerous previous reviews to support safety assessments as reviewed in Table 12, an updated literature search was performed to ensure the current safety assessment considers all relevant data and information. As detailed in Appendix B, the updated literature search for *K. phaffii* spanned the time period from October 1, 2022, to the present to augment the literature review in the most recent GRN addressing *K phaffii* with a posted "no questions letter" at the time of the search (GRN 1056); the literature update in GRN 1056 was conducted on October 3, 2022. One

paper published since GRN 1056 was relevant for the safety of food ingredients produced using *K. phaffii (*Reyes *et al*., 2023), which is summarized below. The literature search was again updated January 29, 2024, and no new information was identified pertinent to the safety assessment.

Reyes *et al*., (2023) evaluated the safety of a product containing a soy hemoglobin preparation derived from *P. pastoris* expressing a soy hemoglobin gene from *Glycine max*. The soy hemoglobin preparation was found to be non-mutagenic in a bacterial reverse mutation assay and non-clastogenic in an *in vitro* micronucleus assay in human lymphocytes. Systemic toxicity was evaluated in a 90-day dietary study in male and female Sprague-Dawley rats that included a 28-day recovery period. No animal deaths were associated with the test article at the highest dose tested with no observed adverse effects on either male or female rats over the course of the 28-day recovery period. The study established a no observed adverse effect level (NOAEL) of 4798 and 5762 mg/kg bw/day, the highest doses tested in male and female rats, respectively. Seven residual host proteins were identified that represented ≥1% of the total protein fraction in at least one batch of the soy leghemoglobin preparations; identification of proteins that account for less than 1% of the total protein fraction were not reported. The study authors concluded that the soy hemoglobin preparation derived from *P. pastoris* is not toxic for consumption under the conditions tested.

6.3.2. Proteins from the Production Strain

As noted in the product specifications, the production of SBSP results in the presence of host proteins produced from the modified *K. phaffii* in the SBSP ingredient (i.e., non-MON proteins). The EDI of non-MON protein in SBSP can be calculated by multiplying the EDI for SBSP by the concentration of non-MON in SBSP, where the average concentration of non-MON is represented as the average difference between total protein and MON (i.e., 90% - 63% = 27%; see Table 10). For a consumer of SBSP under the intended use as a general-purpose sweetener, the “high” intake of non-MON proteins produced from *K. phaffii* for children without diabetes is up to 0.16 mg/kg bw/day (0.58 mg SBSP/kg bw/day x 27% non-MON protein in SBSP). Assuming a body weight of 20 kg, this intake corresponds to 3.2 mg/day of non-MON proteins produced from *K. phaffii*. A review of previous GRAS conclusions in which modified *K. phaffii* has been used as a production organism demonstrates that many of the ingredients also contain residual host cell proteins (GRN 1104, GRN 1056, GRN 1001, GRN 967, GRN 737) with an estimated intake in the range of 56 to 118 mg/day (Table 13). In each GRN, intake of the residual host proteins was concluded to present no safety concerns. The reviews are incorporated by reference herein. Likewise, it is reasonable to conclude that the residual proteins from *K. phaffii* in SBSP present no safety concern under the intended conditions of use.

Table 13. Host proteins in ingredients produced from modified *Komagataella phaffii (Pichia pastoris)* strains concluded to be GRAS for the intended use

GRN No.	Substance	Intake of *K. phaffii* protein		Identification of *K. phaffii* proteins	
		Data	Source	Data	Source
1104	Egg-white protein produced by *Komagataella phaffii* strain ATCC GSD-1235	*K. phaffii* protein intake not specified	-	2 proteins, (Gal d 2, the target protein) and thioredoxin, identified post-purification of the NEWP powder; 12 proteins identified total pre-processing	GRN 1104 p 26-27 of 65
1056	β-lactoglobulin produced by *Komagataella phaffii* strain "yRMK-66"	Estimated per user 90th percentile intake of *K. phaffii* protein is 59 mg/day -Based on 90th percentile intake of β-lactoglobulin ingredient of 56.4 g/day and assuming residual *K. phaffii* proteins are ≤0.15% total protein and ingredient is 70% protein (w/w)	GRN 1056 with amendments; Response to Q 11, p174 of 181; LC-MS/MS report Appendix 2 p57 of 181; Agency response letter p4 of 6	Residual *K. phaffii* proteins identified by the presence of at least two peptides with >0.001% of total protein content are listed in LC-MS/MS report, include 34, 25, and 54 proteins in 3 samples	GRN 1056 with amendments; Response to Q 11, p174 of 181; LC-MS/MS report Appendix 2 p57-61 of 181
1001	Myoglobin preparation from a strain of *Pichia pastoris* expressing the myoglobin gene from *Bos taurus*	Estimated per user 90th percentile intake of *K. phaffii* protein is 118 mg/day -Assuming 2% use rate - *K. phaffii* intake (g) = total protein (g) – total myoglobin protein (g). Estimated intake was presented in GRN 1001	GRN 1001 with amendments; Revised Table 1, p71 of 73	Notifiers of GRN 1001 state: As the Pichia proteins evaluated by Jin *et al*., (2018) and Reyes *et al*., (2021) were derived from the same host strain lineage as that used by Motif, conclusions that residual Pichia proteins are of low toxicity risk can be extended to Motif's Myoglobin Preparation.	GRN 1001 with amendments; Text on p27 of 73
967	Soluble egg-white protein	Estimated per user 90th percentile intake	GRN 967 Amendment	26 residual *K. phaffii* proteins	GRN 967 Amendment

GRN No.	Substance	Intake of *K. phaffii* protein		Identification of *K. phaffii* proteins	
		Data	Source	Data	Source
	produced by *Komagataella phaffii* strain GSD-1209	of *K. phaffii* protein is 56-96 mg/day -Based on 0.8 g/d egg white protein intake 90th %ile -Assumes max 7-12% *K. phaffii* residual host protein	Response to Q 9, p12 of 30; Agency response letter p3 of 5	identified in referenced analysis of same material	p11-15 of 30; considers GRN 204 to be most closely similar despite no ID of host proteins
737	Soy leghemoglobin preparation from a strain of *Pichia pastoris*	Estimated per user 90th percentile intake of *K. phaffii* protein is 111 mg/day -Based on the 90th percentile intake of 533 mg/3day intake of LegH prep dry solids and 20.8% *K. phaffii* protein in LegH prep dry solids -20.8% *K. phaffii* in solids (i.e., 5/24) derived as follows: 5 g = 24 g LegH prep dry solids - 9 g soy LegH protein - 2 g fat - 4 g CHO - 4 g ash)	GRN 737 with amendments p9 of 526, Table 3, p18 of 526	17 proteins identified at >1% of the total protein fraction	GRN 737 with amendments p36-38; Agency response letter p3 of 5; No Concerns letter p4 of 16

6.3.3. Application of the Pariza and Johnson Decision Tree

Pariza and Johnson (2001) developed guidelines that can be used to evaluate the safety of microbial enzyme preparations for use in food processing which have been structured into a decision tree to support safety evaluations (Pariza & Johnson, 2001). This guidance and decision tree has been applied to assess the safety of the production strain used to produce SBSP. Application of the Pariza and Johnson decision tree (Pariza & Johnson, 2001) concludes that *K. phaffii* is accepted as a safe and suitable production strain for a food ingredient, and that SBSP produced by *K. phaffii* is accepted as a safe and suitable food ingredient. The decision tree, in a question-and-answer format is included below in Table 14.

Table 14. Decision tree for evaluating the safety of microbial cultures applied to SBSP

1. Is the production strain genetically modified? If YES, go to 2. If NO, go to 6. ***YES***
2. Is the production strain modified using rDNA techniques? If YES, go to 3. If NO, go to 5. ***YES***
3. Issues relating to the introduced DNA are addressed in 3a-3e.
3a. Do the expressed enzyme products(s) which are encoded by the introduced DNA have a history of safe use in food? If YES, go to 3c. If NO, go to 3b. ***NO***
3b. Is the NOAEL for the test article in appropriate short-term oral studies sufficiently high to ensure safety? If YES, go to 3c. If NO, go to 12. ***YES***
3c. Is the test article free of transferable antibiotic resistance gene DNA? If YES, go to 3e. If NO, go to 3d. ***YES***
3e. Is all other introduced DNA well characterized and free of attributes that would render it unsafe for constructing microorganisms to be used to produce food-grade products? If YES, go to 4. If NO, go to 12. ***YES***
4. Is the introduced DNA randomly integrated into the chromosome? If YES, go to 5. If NO, go to 6. ***YES***
5. Is the production strain sufficiently well characterized so that one may reasonably conclude that unintended pleiotropic effects which may result in the synthesis of toxins or other unsafe metabolites will not arise due to the genetic modification method that was employed? If YES, go to 6. If NO, go to 7. ***YES***
6. Is the production strain derived from a safe lineage, as previously demonstrated by repeated assessment via this evaluation procedure? If YES, the test article is ACCEPTED. If NO, go to 8a or 8b. ***YES - ACCEPTED***

6.4. Safety of the SBSP Product

6.4.1. Absorption, Distribution, Metabolism, and Excretion of SBSP

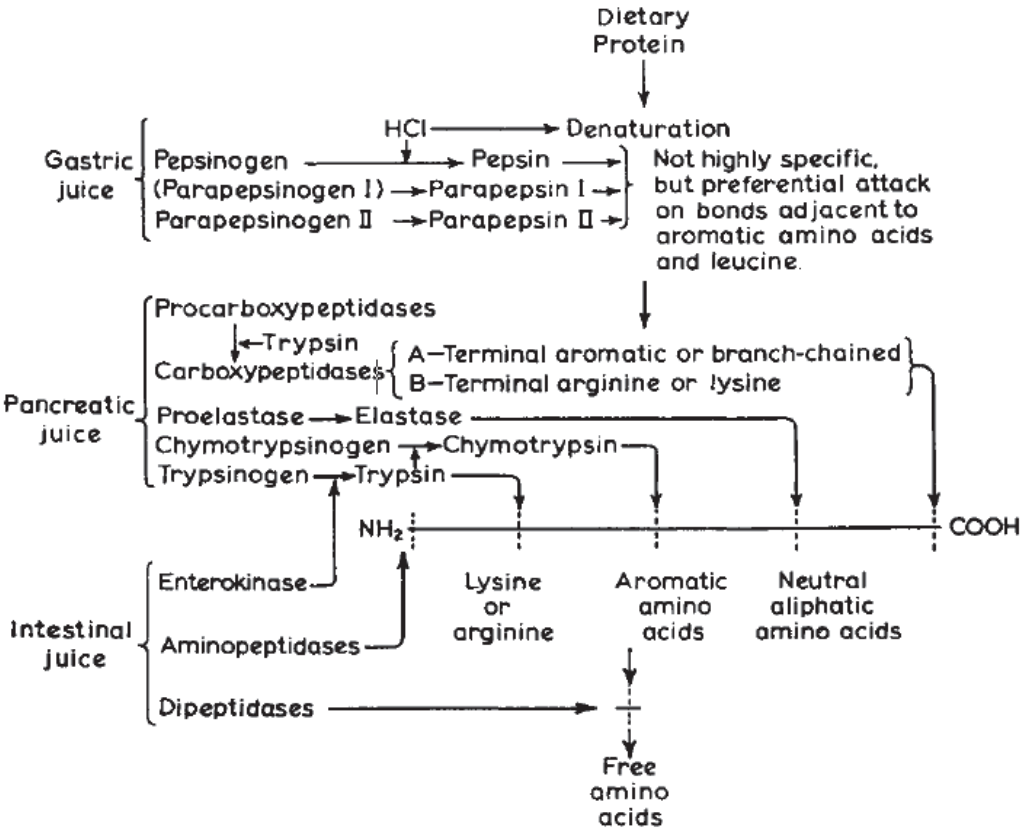
Information on the absorption, distribution, metabolism, and excretion (ADME) specifically of monellin and SBSP was not identified. However, ADME information on proteins is available and applicable to monellin and SBSP since proteins are a necessary component of the diet of humans and other mammals. Oral bioavailability below 1-2% is often recorded for proteins and peptides because of limiting factors including inactivation due to stomach acid or digestive

proteases. Biological availability is limited based on the hydrophilic nature of these proteins, their tendency to undergo denaturation and aggregation, and first-pass metabolism (Renukyuntla *et al*., 2013).

It is generally recognized that the normal fate of dietary proteins is digestion into small peptides and individual amino acids that are absorbed into the systemic circulation *via* transcellular or paracellular routes (EFSA, 2021). In the native state, the peptide chain of the protein, its primary structure, is held relatively firmly in a three-dimensional pattern involving secondary and tertiary levels of structure, and the peptide bonds sensitive to a given proteolytic enzyme may be placed as to be inaccessible to the active center of the enzyme. Upon denaturation, the peptide chain assumes a more random arrangement and the sensitive bonds become exposed to the digestive enzyme. The acidity encountered in the stomach serves to cause instability of the secondary and tertiary protein structure and renders the dietary proteins more susceptible to dietary enzymes. The secondary and tertiary structures of proteins are denatured by the low pH, mechanical action, and protease as well as other enzyme action in the stomach resulting in the exposure of the primary structure of protein (Gitler *et al*., 1964; Steinhardt *et al*., 1955; Callahan *et al*., 2022). The primary structure of the protein is then broken down by proteolytic enzymes in the small intestine. In the large intestine, where the microbiota concentration is much higher and the transit time is longer than in the small intestine, the remaining protein is broken down to peptides and amino acids via extracellular bacterial proteases and peptidases (Macfarlane *et al*., 1986). The combined action of the proteolytic enzymes of pancreatic juice and intestinal mucosa results in rapid further digestion. The loss of tertiary structure to yield the polypeptide chain thereby impairs the protein's function (Callahan *et al*., 2022). Accordingly, the protein structure and biological activity of most proteins are usually lost following ingestion (Metcalfe *et al*., 1996). Proteins that are not fully digested or are partially digested travel to the large intestine where they are ultimately fermented by the gut microbiota (Portune *et al*., 2016; Joye, 2019). Figure 7 summarizes the gastrointestinal digestion of proteins and the enzymes involved.

Proteins which are resistant to protease digestion, and which have thermal stability, have been proposed to be more effective at stimulating immune responses, including allergic sensitization, and be more easily absorbed in forms that could be potentially toxic (Naegeli *et al*., 2017; FAO/WHO, 2020). *In vitro* digestibility is therefore considered as additional information to support a weight of evidence approach for evaluating the safety of newly expressed proteins (FAO/WHO, 2009). *In silico* digestibility prediction models are complementary tools which can be applied alongside *in vitro* digestibility experiments to ascertain likelihood of gastric digestion. These models were employed to examine SBSP containing MON.

Figure 7. Overview of protein digestion in the gastrointestinal tract

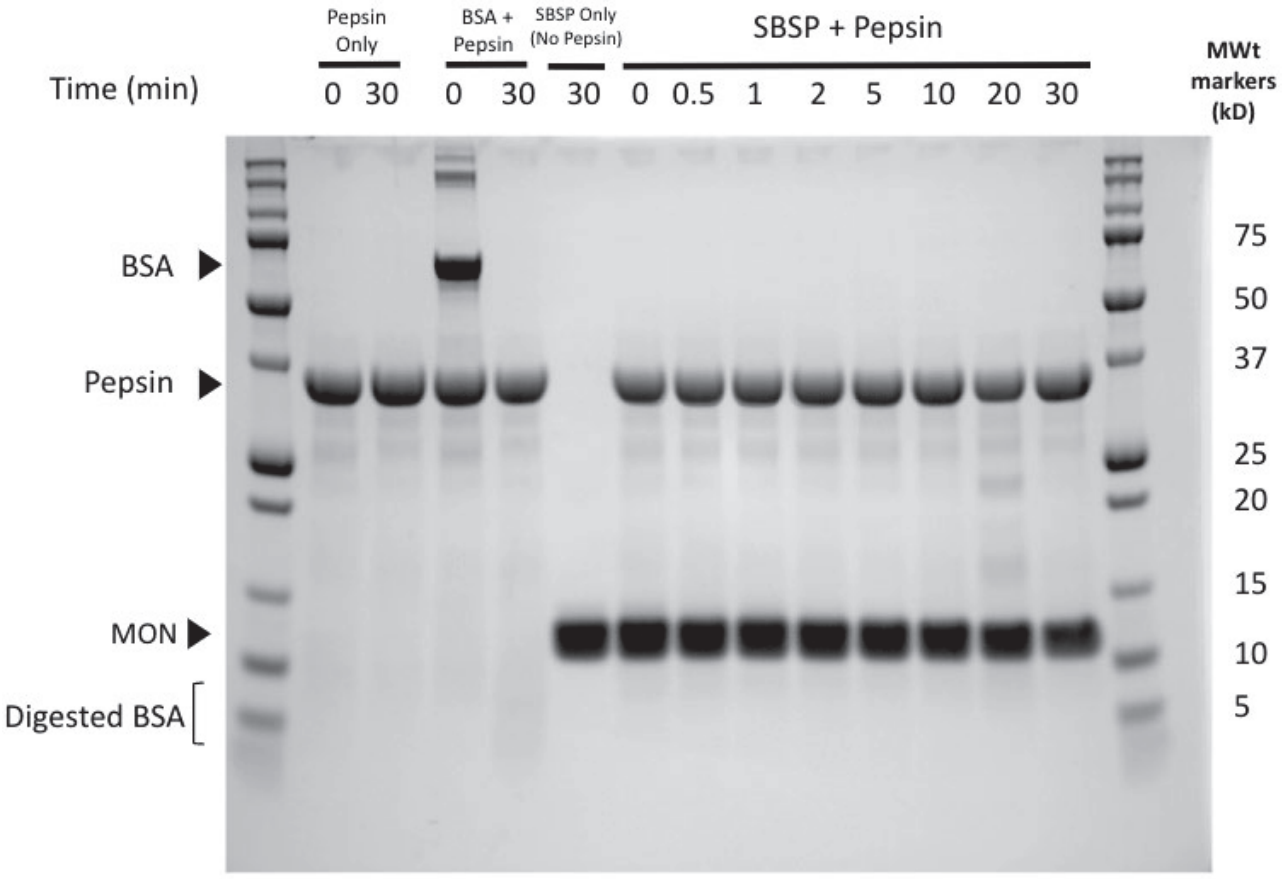


Source: Gilter, 1964

6.4.1.1. In Vitro Digestibility of MON in SBSP

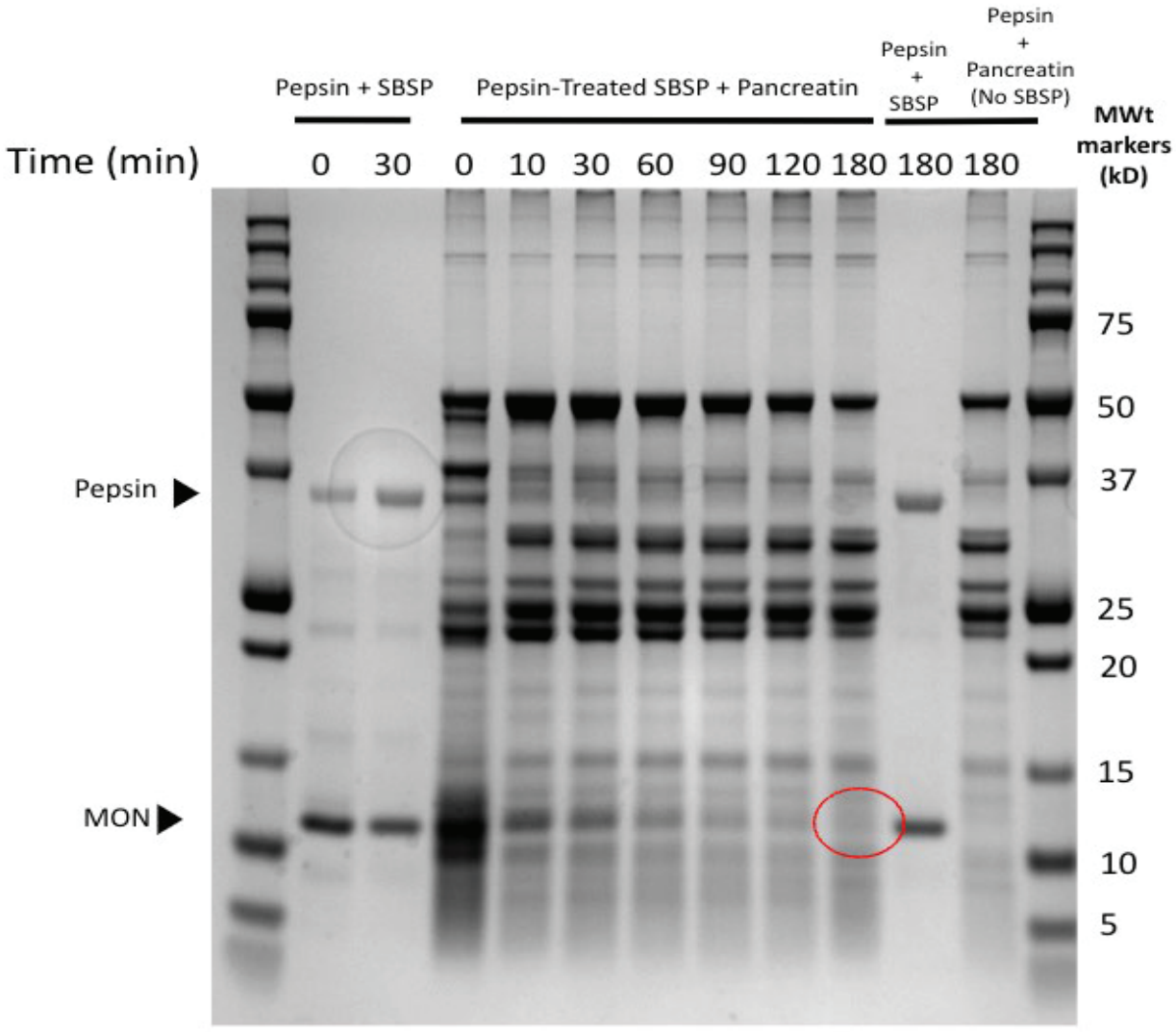
To investigate the relative susceptibility and resistance of MON in SBSP to enzymatic digestion, an experiment was designed with adaptations from the method of Thomas *et al*., 2004 to observe the extent of proteolysis of SBSP containing MON compared to bovine serum albumin (BSA) over time in the presence of pepsin (Freeman *et al.,* 2024). Pepsin activity was checked using the enzyme activity assay protocol from the supplemental material adapted in Brodkorb *et al*. (2019). Purified SBSP (containing 74% MON, 90% total protein) was digested with 10U pepsin/µg of SBSP in simulated gastric fluid (SGF) at pH 2. Samples were taken after timepoints of 0, 0.5, 1, 2, 5, 10, 20, and 30 minutes at 37°C incubation. Proteolysis was stopped with 200mM sodium bicarbonate (pH 11) after each of these timepoints up to 30 minutes. BSA was used as a positive control known to be digested by pepsin. The results shown in the sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 8) qualitatively show that MON in SBSP purified from the production process described appears to begin to be digested by pepsin after 30 minutes of incubation.

Figure 8. SDS-PAGE of SBSP containing MON protein digests in the presence of pepsin after a range of incubation time periods, with molecular weight (MWt) markers



As a result, a further 2-phase digestibility study of MON in SBSP was conducted for longer incubation periods using the adapted INFOGEST 2.0 static *in vitro* model simulating gastrointestinal food digestion (Brodkorb *et al*., 2019). In this study, SBSP (containing 74% MON, 90% total protein) was digested with 10U pepsin/µg of SBSP in SGF at pH 2 for 30 minutes at 37°C (Freeman *et al.,* 2024). The samples were then pH adjusted to 7 and added in a ratio of 1:1 to pancreatin in simulated intestinal fluid (SIF). In this new environment, digestibility of MON in SBSP was tested up to 180 minutes at 37°C incubation. A negative control containing pepsin and pancreatin, but no SBSP, was incubated following the same 2-phase process and conditions. In addition, pepsin only with SBSP was carried through to 180 minutes to demonstrate proteolytic activity in parallel with samples undergoing the 2-phase condition. The resultant digests were run on an SDS-PAGE gel against molecular weight standards. Qualitatively, the majority of MON in SBSP was digested after 180 minutes with pancreatin (Figure 9).

Figure 9. SDS-PAGE of SBSP containing MON protein digests with a single phase of 30 min digestion with pepsin followed by digestion with pancreatin over a range of incubation time periods, with molecular weight (MWt) markers and pepsin + pancreatin as negative controls



6.4.1.2. In Silico Digestibility of MON in SBSP

The digestibility of MON was evaluated *in silico* using PeptideCutter,[5] a publicly available bioinformatics tool which predicts potential cleavage sites in a peptide sequence by proteases (Freeman *et al.,* 2024). PeptideCutter allows for the selection of one or multiple digestive chemicals or enzymes and, upon analysis of a query peptide sequence, provides an output stating the predicted number of cleavage sites for each selected digestive chemical or enzyme along with the identity of the resulting cleaved peptides. Briefly, predicted cleavage of the amino acid sequence of MON in SBSP was assessed using pepsin (pH <1.3) and trypsin digestions to

[5] https://web.expasy.org/peptide_cutter/

simulate the predominant enzymes present in the human stomach and small intestine, respectively. MON was predicted to have 15 pepsin and 15 trypsin cleavage sites, resulting in an average fragment size of 3.43 amino acids. It has been proposed that if a protein digest is composed of peptides <9 amino acid residues in length, the allergenic potential can be considered to be low (Naegeli *et al*., 2017).

The results from both the *in vitro* digestibility studies presented and the *in silico* digestibility tool applied indicate that the majority of SBSP containing MON would be broken down by human digestive enzymes upon ingestion, thus reducing likelihood to pose a toxigenic or allergenic risk to consumers. These indicator data can be combined with further bioinformatic analysis of homology with known allergens in a weight of evidence approach to support safety.

6.4.2. In Silico Toxigenicity of SBSP

The UniProtKB database of proteins[6] was reviewed to tag those proteins known to be animal, plant, fungal, bacterial, or viral toxins with the keyword 'toxin' (KW-0800), resulting in a curated database of >7,600 proteins against which new proteins can be compared. Sequence homology searches were conducted using BLAST comparing the FASTA sequences for MON and the 5 production host cell proteins which together make up 95% of proteins present in SBSP (see Appendix G) against all protein sequences listed in UniProt. The five host cell proteins are all present at >1% w/w and include superoxide dismutase, SCP domain-containing protein, chitin deacetylase, fusion protein (identical to Rpl40Bp), and ACB domain-containing protein (see Appendix A for the complete proteomic characterization of SBSP). The BLAST searches were conducted using the default algorithm and scoring parameters (BLOSUM62 scoring matrix with default gap costs and composition adjustments) and an expected threshold of 0.01. The BLAST search results of proteins identified as having sequence homology for each host cell protein evaluated were then further searched to identify hits against proteins confirmed as 'toxins' in the reviewed database maintained by UniProt using the search terms '(keyword: KW-0800) AND (reviewed:true)'.

None of the proteins identified by the BLAST search as having sequence homology with MON or the host cell proteins in question were identified by UniProt as confirmed toxins. The findings of this review suggest no evidence of toxigenic potential for MON and the 5 residual proteins from the production strain present above 1% w/w in SBSP.

6.4.3. Pre-Clinical Studies of SBSP

6.4.3.1. Overview of Pre-Clinical Studies of SBSP

To further support an understanding of the safety of SBSP for the intended use as a general-purpose sweetener, Oobli, Inc. sponsored a series of pre-clinical studies on SBSP, including *in vitro* genotoxicity (bacterial reverse mutation and mammalian micronucleus assays), and 14-day and 90-day oral (dietary) toxicity studies in rats. These studies, along with results from the *in*

[6] www.uniprot.org

vitro and *in silico* protein digestion and *in silico* allergenicity reported elsewhere in this dossier, are also reported in the published literature (Freeman et al., 2024).

6.4.3.2. Test Materials

All pre-clinical studies were conducted with SBSP containing MON. The material used in the 90-day study was produced from an initial fermentation feedstock source that with a lower protein and MON content compared to SBSP that is currently manufactured (Table 15). Quality control measures have since been optimized to ensure a higher MON content and consistency across batches, as demonstrated by the representative production batch data presented in Table 5.

Table 15. Composition of SBSP used in the toxicity studies vs. production batch data

Parameter	Specification	Average Production Batch Data [a]	Batch Used in Genotoxicity Studies	Batch Used in 90-Day Toxicity Study
Moisture, % w/w	<10	5.07	Not available	7.05
Total protein, as is, % w/w	>70	89.99	85	37.1
MON, % w/w	>50	63	55	21
Fat by fatty acid profile, g/100g	<1	0.05	Not available	0.02
Ash, % w/w	<5	1.53	Not available	0.52
Carbohydrate, % w/w	<15	3.35	Not available	55.35

[a] Values represent average of data from three production batches presented in Table 5.

Relative to currently made production batches of SBSP, the SBSP material used in the 90-day toxicity study had a lower concentration of total protein (including MON), and a higher concentration of carbohydrate. Analysis of this SBSP test material confirms the absence of free sugars (as reducing sugars) and the presence of high molecular-weight carbohydrate, including fiber and starch (measured with AOAC 2011.25 and AOAC 979.10, respectively). These results indicate that the carbohydrates co-purified with the SBSP product used in the 90-day toxicity study are likely either a soluble fiber (e.g., mannan) produced during the fermentation, or the indigestible portion of the fermentation feedstock (i.e., residual corn fiber and corn starch).

Spectral analysis of the SBSP batch used in the 90-day toxicity study identified proteins with assigned UniProt accession numbers for proteins originating from *K. phaffii* (see data in Appendix A, Table 2). Comparison of the spectral analysis results from the 90-day toxicity study batch and the average production batch shows that the material used in the 90-day toxicity study provided 97% of proteins in the production batches. These results demonstrate that the material used in the 90-day toxicity study is consistent with the host cell protein content of the product meeting the established specifications.

6.4.3.3. Genotoxicity

SBSP was evaluated for the potential to cause genotoxicity in a bacterial reverse mutation assay and an *in vitro* mammalian micronucleus assay in human lymphocytes (Freeman *et al.,* 2024).

The genotoxicity studies were conducted by Product Safety Laboratories, New Jersey (USA), under Good Laboratory Practices (GLP).

A bacterial reverse mutation assay was conducted with SBSP (containing by weight 55% MON, 85% total protein) in accordance with U.S. FDA Toxicological Principles for the Safety Assessment of Food Ingredients, Redbook 2000, IV.C. 1. a. "Bacterial Reverse Mutation Test" (FDA, 2007) and the ICH S2 (R1) Guidance on Genotoxicity Testing and Data Interpretation for Pharmaceuticals Intended for Human Use (FDA, 2012) under GLP. The assay was conducted using *Salmonella typhimurium* (*S. typhimurium*) strains TA98, TA100, TA1535, and TA1537, and *Escherichia coli* (*E. coli*) strain WP2 uvrA, both in the presence and absence of metabolic activation by S9 mix up to 5000 µg/plate. There were no concentration-related or substantial test substance related increases in the number of revertant colonies in both the absence and presence of S9 using either the plate incorporation or the pre-incubation method. Based on the results of this study, SBSP was negative for mutagenicity with and without metabolic activation in all tested strains (Freeman *et al.,* 2024).

An *in vitro* mammalian micronucleus assay in human lymphocytes was conducted with SBSP (containing by weight 55% MON, 85% total protein) in accordance with the OECD Guidelines for Testing of Chemicals, Section 4, No. 487, "*In Vitro* Mammalian Cell Micronucleus Test", adopted 29 July, 2016, under GLP. For this study, human peripheral blood lymphocytes were obtained from healthy and non-smoking donors. In all experiments, numbers of micronucleated cells were within or below the historical control limits of the negative control and did not show a biologically relevant increase compared to the concurrent negative control. No statistically significant enhancement ($p<0.05$) of cells with micronuclei was noted in the dose groups of the test item evaluated under all conditions with and without metabolic activation. Under the conditions of this study, SBSP is considered non-mutagenic with respect to clastogenicity and/or aneugenicity in the *in vitro* mammalian cell micronucleus test. Based on the results of the available testing, SBSP containing MON is not genotoxic in bacterial or mammalian cells (Freeman *et al*., 2024).

6.4.3.4. Sub-chronic Toxicity

SBSP (containing by weight 21% MON, 37.1% total protein) was evaluated in an oral 14-day range-finding/palatability toxicity study in rats (non-GLP); the study design was based on OECD Test Guideline 407 (2008). Rats (5/sex/dose) were administered the test material at 0, 0.6, 1.2, or 2.4% in the diet, nominally equivalent to 0, 500, 1000, or 2000 mg/kg bw/day. Animals were observed at least twice daily for viability and cage-side observations for conducted daily. Body weights were recorded on Day 0, prior to the day of dosing, and on days 3, 7, 10, and 14. Food consumption was measured to coincide with body weight measurements. Food efficiency and test material intake was calculated. All animals were subject to gross necropsy at the termination of the study. There were no adverse, treatment-related effects reported in the study at any dose level for any parameter evaluated. Based on the results of this study, the upper dose level of 2000 mg/kg bw/day was selected for the 90-day toxicity study (Freeman *et al*., 2024).

In the 90-day toxicity study, conducted under GLP and in compliance with OECD Test Guideline 408 (2018), rats (10/sex/group) were offered diets prepared containing 0, 500, 1000, or 2000 mg/kg bw/day of SBSP (containing by weight 21% MON, 37.1% total protein). The results

of the 90-day toxicity study provide an overall nominal NOAEL of 2000 mg/kg bw/day with test material intakes of 0, 492, 972, and 1954 mg/kg bw/day in males and 0, 489, 959, and 1967 mg/kg bw/day in females for SBSP. The dosages with regards to the content of MON (21% MON) were 0, 103, 203, and 408 mg/kg bw/day in males and 0, 102, 200, and 411 mg/kg bw/day in females.

Results of the concentration verification, homogeneity, and stability analyses were acceptable (±15%). No mortality or clinical effects were reported in the study. No treatment-related effects were reported on mean bodyweight, mean bodyweight gain, food consumption, and food efficiency. No treatment-related effects were reported following ophthalmological evaluation or in hematology, clinical chemistry, thyroid hormone, and urinalysis analyses. Any statistically significant differences from control were considered not treatment related because there was no dose response, the effect was within the range of historical control values, the magnitude of the change was not considered adverse, and/or the effect was only reported in one sex. No effects were reported on organ weights (absolute or relative to bodyweight) and no treatment-related macropathological or histopathological abnormalities were reported (Freeman *et al*., 2024).

The results of the 14-day and 90-day toxicity studies provide an overall nominal NOAEL of 2000 mg/kg bw/day (1954 mg/kg bw/day in males and 1967 mg/kg bw/day in females) with no adverse effects reported at the highest dose level tested (Freeman *et al*., 2024). The NOAEL for SBSP was 1954 mg/kg bw/day. The NOAEL adjusted for the content of MON (21% MON) was 408 mg/kg bw/day. A 100-fold uncertainty factor was applied to the NOAEL to account for interspecies (rat to human) and intraspecies (human to human) differences. The resulting acceptable daily intake (ADI) for MON is 4.1 mg/kg bw/day and 19.5 mg/kg bw/day for SBSP.

6.4.4. Pre-Clinical Studies of Other Single-Chain Monellin Ingredients

Rega *et al*., 2017 investigated the potential mutagenicity and genotoxicity of an engineered single-chain monellin produced by *E. coli*. The testing included genotoxicity and mutagenicity tests using *S. typhimurium* (strain TA97a, TA98, TA100, and TA1535) and *E. coli* (strain WP2 pkM101) at concentrations up to 500 mg/plate. Guideline-compliant study protocols provide maximum concentrations of 5 mg/plate. Based on findings up to 5 mg/plate, the engineered proteins were non-genotoxic and non-mutagenic.

The safety of a single-chain monellin (purity not reported) produced in *K. phaffii* was assessed in acute, sub-chronic, and chronic toxicity tests and genotoxicity tests (Novik *et al*., 2023). The single-chain monellin test substance was Chain B and Chain A with a Gly-Phe linkage. The reported oral LD_{50} was >5000 mg/kg bw in rats and mice (species not specified). In the subacute (21-day) oral toxicity study, single-chain monellin was administered via gavage to guinea pigs (9/sex/dose) at 0, 1.45, or 14.5 mg/kg bw/day; no effects were reported at the highest dose tested. In the chronic toxicity study in rats, single-chain monellin was administered via gavage to outbred rats (10/sex/dose; species not specified) at 0 (vehicle control), 1.43, or 14.3 mg/kg bw/day for a period of 150 days. A positive control group was administered sucrose at 4824 mg/kg bw/day. The authors reported a statistically significant difference in relative body weight gain in male rats at 1.43 mg/kg bw/day (-9%) and 14.3 mg/kg bw/day (-9%) when compared to the control group. However, the dose groups administered sucrose (4824 mg/kg bw/day) also

demonstrated a decrease in mean body weight gain (-12%) compared to the control group. In a bacterial reverse mutation assay (*S. typhimurium* strains TA 98, TA 97, and TA 100), single-chain monellin (up to 50,000 ug/mL) was reported as negative. Single-chain monellin was reported as negative in *in vivo* micronucleus and *in vivo* chromosome aberrations tests in mice (species not specified) up to 5,000 mg/kg bw/day.

6.5. Allergenicity

Consistent with the recommendations in the Codex Alimentarius assessment scheme for potential allergenicity of proteins derived from recombinant-DNA microorganisms (CAC, 2003), consideration has been given to the history of safe use for the host microorganism and protein derived from it. A search of the scientific literature found no reported cases of allergy and no apparent adverse hypersensitivity reactions to monellin, serendipity berry, or *Dioscoreophyllum cumminsii*. Search details are provided in Appendix B. Four of the identified articles addressed the production and characterization of monoclonal antibodies to monellin for the purpose of identifying antigenic epitopes and developing immunoassays for quantitation of sweet tasting proteins in various plant extracts/preparations. Two articles noted the immunological cross-reactivity of antibodies to thaumatin with monellin, with multiple identical tripeptides found to be in common (Iyengar *et al*., 1979). The authors proposed that these tripeptides may be involved in the interaction of these sweet-tasing proteins with the sweet-taste receptor in humans.

Barre and colleagues propose that although there are limited sequence and tertiary structural homologies between monellin and the cystatin protein orzacrystatin, which is known to cause hypersensitivity in humans, this structural relationship could indicate possible allergenicity potential for monellin (Barre *et al*., 2015). However, the postulated allergenic potential of monellin remains to be demonstrated in humans. These investigators make the suggestion that the closest phylogenetic relationships for monellin and cystatins are observed with allergens of animal origin, namely Fel d 3 from cat and Ani s 4 from anisakis or cystatin from egg white. This information has been applied when considering potential for cross-reactive adverse allergic reactions in the following sections.

6.5.1. Related Allergens and Taxonomy

WHO/IUIS Allergen Nomenclature sub-committee is responsible for maintaining and developing this unique, unambiguous and systematic nomenclature for allergenic proteins in a publicly accessible online database. A minimal criterion of demonstrated IgE binding to the suggested allergen using sera from patients allergic to the specific source is required for a protein to be accepted and formally assigned nomenclature by the WHO/IUIS sub-committee. A search for the terms '*Dioscoreophyllum cumminsii'* and 'monellin' was performed in the WHO/IUIS allergens database.[7] The WHO/IUIS allergens database had no entries for any allergens from plant species *Dioscoreophyllum cumminsii.* It had also no entries for the protein monellin. A search for the terms *'Dioscoreophyllum cumminsii'* and 'monellin' was also performed in the

[7] WHO/IUIS Allergen Database

AllergenOnline,[8] Allergome,[9] and COMPARE[10] allergen databases and the AllergenOnline celiac protein database.[11] No entries for any allergens from the species *Dioscoreophyllum cumminsii* and no entries for the protein monellin were found in any of these four databases.

As monellin is documented to be structurally related to the cystatin family of thiol proteinase inhibitors (Murzin *et al*., 1993) without sharing significant amino acid sequence similarity, searches were also undertaken in the WHO/IUIS allergen database for the terms phytocystatin, cystatin, and stefin B in the 'biochemical name' search for 'All allergens', and in AllergenOnline allergen and celiac protein databases, and the Allergome and COMPARE allergen database**s** for the terms phytocystatin, cystatin, and stefin B in the text search boxes. The WHO/IUIS, AllergenOnline, Allergome, COMPARE, and AllergenOnline celiac protein databases returned no entries for stefin B. The Allergen Online celiac protein database had no entries for phytocystatin, cystatin and stefin B. However, several of the different allergen databases did identify entries for types of cystatin and phytocystatin and cross-reacting known allergens. These results are presented in Table 16.

Table 16. Results from searches in WHO/IUIS, COMPARE, AllergenOnline, and Allergome allergen databases

Species	IUIS Name	Common Name	Size	Source	Allergen Database
Actinidia deliciosa (Green kiwi fruit)	Act d 4 Act d 4.0101	Phytocystatin	11 kDa	Ingestion	1, 2, 3, 4
Actinidia chinensis (Gold kiwi fruit)	Act c 4	Phytocystatin, cystatin	11 kDa	Ingestion	2, 3
Canis familiaris (C. lupus familiaris) (Domestic dog)	Can f 8 Can f 8.0101	Cystatin	14 kDa	Airway	3
Felis domesticus (F. catus) (Domestic cat)	Fel d 3 Fel d 3.0101	Cystatin-A	11 kDa	Airway	1, 2, 3
Ambrosia artemisiifolia Ambrosia elatior (Mugwort / ragweed-related species, short ragweed)	Amb a CPI	Cystatin Proteinase Inhibitor	-	Airway	3
Anisakis simplex (Parasitic worm)	Ani s 4 Ani s 4.0101	Cysteine Proteinase Inhibitor	9 kDa	Ingestion	3
Glycine hispida Glycine max (Soy, soyabean)	Gly m CPI	Cysteine Proteinase Inhibitor	-	Ingestion	3

[8] AllergenOnline
[9] Allergome
[10] COMPARE
[11] http://www.allergenonline.org/celiacproteinbrowse.shtml

Species	IUIS Name	Common Name	Size	Source	Allergen Database
Solanum tuberosum (Potato)	Sola t 3 Sola t 3.0101 Sola t 3.0102	Cysteine Protease Inhibitor	21kDa	Ingestion	3

Data sources: 1 WHO/IUIS allergen database http://allergen.org/; 2 COMPARE allergen database http://db.comparedatabase.org/; 3 Allergome allergen database https://www.allergome.org/; 4 AllergenOnline allergen database http://www.allergenonline.com/

6.5.2. Bioinformatics Allergenicity Analysis

Bioinformatics analysis can be used to compare novel proteins with the peptide sequences and three-dimensional structure of known allergens to obtain predictions of allergenicity potential and of cross-reactivity potential with known IgE-mediated allergens and with gluten-like allergens. There are multiple bioinformatic tools currently available to support this analysis.

Alignments with high identity scores from a FASTA sequence comparison search may indicate a potential for allergenic cross-reactions. FAO/WHO (2001) has recommended that IgE cross-reactivity between a particular protein and allergen be considered when there is greater than 35% identity over a sliding "window" of 80 amino acids. The 35% identity threshold was based on data indicating protein cross-reactivity occurring between Bet v 1 and vegetable proteins at approximately 40% protein identity (Scheurer *et al*., 1999). The 80 amino acid window was selected to represent the size of a “typical” protein domain. Codex Alimentarius Commission developed guidelines for the evaluation of the potential allergenicity of novel proteins, in particular those derived from recombinant DNA plants (FAO/WHO, 2009). They recommended a bioinformatics search using a FASTA or a BLASTP algorithm and suggested that matches of at least 35% identity over segments of at least 80 amino acids may indicate the possibility of cross-reactivity. This Codex Alimentarius guideline has become a widely applied standard for possible risks of cross-reactivity. Similarly, evaluation of the potential for novel proteins to elicit celiac disease or activate MHC class II restricted T cells in human subjects with celiac disease can be performed using bioinformatics comparisons with known celiac inducing peptides. It has been proposed that identity matches of less than 45% over at least one-half of the FASTA aligned CD protein and those with an E score smaller than 1×10^{-16} using this database are unlikely to present a risk of inducing celiac disease (AllergenOnline, 2023).

The allergen database and tools chose for bioinformatics analysis for allergenicity are AllergenOnline allergen database and celiac database and FASTA36 and sliding 80mer window comparisons, the COMPARE allergen database and the associated COMPASS search software, and AllerCatPro 2.0 allergen database modelling tool.

6.5.2.1. AllergenOnline

The AllergenOnline allergen database is curated by the Food Allergen Research and Resource Program at the University of Nebraska. AllergenOnline provides access to a peer reviewed allergen list and sequence searchable database intended for the identification of proteins that may present a potential risk of allergenic cross-reactivity. The sequence search routines allow

comparison between a protein sequence with the sequences in the current AllergenOnline database, which is updated on an annual basis. The most predictive search is the overall FASTA alignment where identity matches greater than 50% indicate possible cross-reactivity. A further precautionary search can be done using a sliding window of 80 amino acid segments of each protein to find identities greater than 35% according to CODEX Alimentarius guidelines (CAC, 2009). Based on historical data, the AllergenOnline research group state that allergen cross-reactivity is not likely for proteins with less than 50% identity over the entire protein sequence and is fairly common above 70% identity.

6.5.2.2. COMprehensive Protein Allergen REsource (COMPARE)

The COMPARE (7th iteration) database (van Ree *et al.,* 2021), is a publicly accessible allergen sequence data resource created by an international collaborative scientific group coordinated by the Health and Environmental Sciences Institute. Allergens are first identified via "rule-based" text sorting algorithms which scan through public sequence and literature repositories. An independent peer-review panel of allergy experts from the public sector then reviews the candidate sequences from the automated selections and decides on the final content of the database on an annual basis. A sequence is selected only if it has literature supporting evidence of IgE binding. Approved allergens are included in the COMPARE database along with their supporting literature containing evidence of allergenicity. COMPARE is equipped with a comparative sequence search software COMPASS, incorporating the open source FASTA software package (FASTA v36). The criteria are oriented to identifying amino acid sequence alignments that might be observed between two or more amino acid sequences for any sequence compared with the COMPARE database. Search options in COMPASS offer options to perform three independent types of sequence comparisons: Full Length Sequence search; 80-mer sliding window FASTA search, and 8-mer FASTA search.

6.5.2.3. AllerCatPro 2.0

The AllerCatPro 2.0 database has been developed to predict potential allergenicity of proteins using similarity of both their amino acid sequences and three-dimensional structural similarity compared with a dataset of known allergens (Maurer-Stroh *et al*., 2019). The known allergen databases are a compilation from WHO/International Union of Immunological Societies, COMPARE, AllergenOnline, UniProtKB and Allergome databases. AllerCatPro 2.0 checks the similarity of the query protein with 714 representatives in the 3D model/structure database of known allergens as well as the most comprehensive dataset of reliable proteins associated with allergenicity (4979 protein allergens). In addition to comparing the similarity of the query protein with the dataset of known allergens, AllerCatPro 2.0 also predicts the similarity of the query sequence to datasets of 165 autoimmune allergens and 162 low allergenic proteins separately. If a significant sequence similarity is found, AllerCatPro 2.0 then identifies hits of similar proteins associated with autoimmune diseases and/or similar proteins of low allergenic potential and presents the sequence identity to the closest hit. In case of a hit, links are provided to view further details on the most similar allergens with results for cross-reactivity, protein information (UniProt/NCBI), functionality (Pfam, InterPro, SUPFAM), as well as clinical relevance of IgE prevalence (Allergome) and allergen information, including the most similar 3D surface epitope. AllerCatPro 2.0 also identifies all similar allergens that have significant sequence similarity to the query protein and refers to the number with the link in potential cross-reactivity of the output

table, as well as all possible similar autoimmune allergens displayed in the link and all possible similar low allergenic potential allergens in the link of the output table.

The AllerCatPro 2.0 query tool workflow is explained on their website. The presence of gluten-like repeats of glutamine (gluten-like Q-repeats) is first evaluated as this is independent of any other similarity scores and a gluten-like prediction does not lead to a result for 'strong evidence' of allergenicity potential in this model. Secondly, the AllerCatPro 2.0 algorithm then checks systematically the similarly to a 3D model/structure database. Results showing similarity are expressed. If the model cannot find any structure hits, then a linear-window approach is applied to predict the queried sequence as protein allergen with strong evidence if the rule of 35% identity over 80 residues is found. The E-value indicates the probability due to chance, and thus, the closer the E-value is to zero, the more significant (but not random) the similarity is towards the query protein sequence. If still no hit is found, then the hexamer hit approach is applied seeking homology with at least three short hexamers with known allergens. If then AllerCatPro 2.0 can still not find any hits, a prediction of 'no evidence' for allergenicity is assigned.

6.5.3. FASTA Sequences

The National Center for Biotechnology Information (NCBI) Protein database contains a collection of protein sequences from several sources, including SwissProt and PIR components of UniProt; Protein Research Foundation (PRF), Protein Data Bank (PDB), and translations of coding regions on sequences in Entrez Nucleotide International Sequence Database Collaboration (DDBJ / EMBL / GenBank).

A search for monellin from *Dioscoreophyllum cumminsii (Stapf) Diels* (i.e., serendipity berry) in the NCBI protein database returned 92 entries for monellin specifically from *Dioscoreophyllum cumminsii (Stapf) Diels* origin. These were refined in the NCBI Identical Protein Groups database down to just 5 entries: 3 unnamed protein products, monellin chain A, and monellin chain B. The unnamed protein product 2O9U_X differs from MON in one additional amino acid and in one inversion between two amino acids.

The details and FASTA sequences for these 5 proteins and for MON are presented in Appendix C. MON is identical to monellin NCBI 1IV9_A, which is made up of native monellin A chain and native monellin B chain linked through an additional glycine residue.

6.5.4. Bioinformatics Analysis of FASTA Sequences for Monellin

A FASTA sequence comparison routine was used to compare each of the 5 protein sequences listed in the NCBI database for monellin and the FASTA sequence for MON found in SBSP (Appendix C) to entries in the AllergenOnline database, the COMPARE Database, and to apply the AllerCatPro 2.0 tool for estimation of allergenicity. The purpose of this FASTA sequence comparison is to evaluate whether the query protein sequence is identical to, or homologous with, known or putative allergens and with gluten in a database.

Results from the FASTA bioinformatics comparison sequence searches are provided in Appendix D. The output includes a list of aligned sequences from best to least similar, statistical

scoring of each alignment (Expectation, or E-score value), a percent identity of the overlapping alignment and the best alignment of the query and aligned protein. A summary of the findings is provided below and in Table 17:

- None of the 5 proteins identified in the NCBI protein database for monellin from *Dioscoreophyllum cumminsii* or MON showed matches with a >35% identity over 80mer amino acid sequence in the AllergenOnline FASTA searches or showed >50% alignment with known allergens in the AllergenOnline database.
- None of the 5 proteins identified in the NCBI protein database for monellin from *Dioscoreophyllum cumminsii* or MON showed matches with a >35% identity over 80mer amino acid sequence in the COMPARE database FASTA searches or showed >50% alignment with known allergens in the COMPARE database.
- None of the 5 proteins identified in the NCBI protein database for monellin from *Dioscoreophyllum cumminsii* or MON were found to have evidence of allergenicity following analysis using the AllerCatPro 2.0 tool using an E value threshold 0.001 and no evidence of similarity to known gluten allergens.

Table 17. Summary of allergenicity concerns for native Monellin A, native Monellin B and MON

Database	Search type	Criteria for Match identification	Native Monellin Chain A	Native Monellin Chain B	MON
Allergen Online Database (Version 21)	Full-length alignment search	Sequence alignment of ≥35% and *E* value cutoff = 1	None	None	None
	80 amino acid alignment search (sliding window of 80-amino acid sequences)	Sequence alignment of ≥35% and *E* value cutoff = 1	None	None	None
	8-mer approach	Exact match over the sequence of 8 amino acids	None	None	None
COMPARE online allergen database (Updated 26 Jan. 2023)	Full-length alignment search	Sequence alignment of ≥35% and *E* value cutoff = 1	None	None	None
	80 amino acid alignment search (sliding window of 80-amino acid sequences)	Sequence alignment of ≥35% and *E* value cutoff = 1	None	None	None
	8-mer approach	Exact match over the sequence of 8 amino acids	None	None	None

Database	Search type	Criteria for Match identification	Native Monellin Chain A	Native Monellin Chain B	MON
AllerCatPro 2.0 tool (Nguyen *et al.,* 2022)	Predicted most similar allergen (non-autoimmune allergens)	Sequence alignment of ≥35% identical to allergens over 80 residue window E value threshold 0.001	No significant hits No evidence of allergenicity	No significant hits No evidence of allergenicity	No significant hits No evidence of allergenicity
	Similarity to Gluten-like proteins (# of Q-repeats)	Hit if the score for a 9-mer is within one std. dev. of the average of the FARRP "Celiac disease peptides"	None	None	None

Barre *et al.* (2015) suggest that the closest phylogenetic relationships for monellin and cystatins are observed with allergens of animal origin, namely Fel d 3 from cat, Ani s 4 from anisakis and cystatin from egg white. The sequence homology comparison investigations found some similarity between these known allergens and native monellin chain B, and MON proteins containing monellin chain B sequences. However, the level of sequence identity was <35% and does not indicate a significant likelihood for cross-reactivities with these well-documented allergens.

6.5.5. Bioinformatics Analysis of FASTA Sequences for Production Strain Host Cell Proteins in SBSP

For completeness, and in line with the recommendations in the Codex Alimentarius assessment scheme for potential allergenicity of proteins derived from recombinant-DNA microorganisms (CAC, 2003), the potential allergenicity of the host cell proteins originating from the *K. phaffii* derived production strain which are found in SBSP was assessed (Appendix E). Potential allergenicity of the host cell proteins was assessed using the same approach used to assess potential allergenicity of the single chain monellin MON expressed by the production strain in SBSP. The FASTA amino acid sequences for the five host cell proteins present at a >1% total protein in SBSP were assessed using the AllergenOnline allergen database and celiac protein database and the AllerCatPro 2.0 model to estimate potential allergenicity and risk of eliciting celiac disease.

Of these five proteins, one host cell protein, Cu/Zn-superoxide dismutase (SOD) was proposed by the AllerCatPro 2.0 modelling tool to have strong evidence of allergenicity potential. The AllergenOnline allergen database search produced hits for SOD against olive pollen allergens with E values smaller than 1e-030 and as such indicate some likelihood of cross-reactivity in individuals allergic to olive pollen. SOD are ubiquitous metalloproteins found in a wide range of organisms to which consumers are frequently exposed. The evidence for allergenicity potential for Cu/Zn-SOD appears to be of limited relevance as it relates to the close homology and 3D

epitope structure with the documented inhalation allergen Neo fi Cu/Zn-SOD from *Neosartorya fischeri* (an environmental mold) and related *Aspergillus* spp., and to records documenting IgE binding prevalence for 66 subjects to Ole e 5 *(Olea* europaea; olive tree pollen), identified as cross-reacting with Cu/Zn-SOD.

A further host cell protein, SCP domain-containing protein was found to have weak evidence of allergenicity potential in both allergen database models for sequence homology and allergenicity potential.SCP domain-containing protein is part of CRISP family and the pathogen-related protein-1 (PR-1-like) protein superfamily, which has a ubiquitous presence across a wide range of organisms to which consumers are commonly and frequently exposed. The AllergenOnline allergen database results indicate a low potential for cross-reactivity for SCP domain-containing protein with pathogenesis related protein PR-1 from pollen sources. The AllerCatPro 2.0 results also predict some potential for cross-reactions upon consumption for individuals sensitized to Cuc m 3 from *Cucumis melo* (cantaloupe or muskmelon) and sensitized via the inhalation route to airborne Art v 2 from *Artemis vulgaris*, but the evidence is very limited. Two reports show evidence of IgE prevalence for Cuc m 3 from *Cucumis melo* (cantaloupe or muskmelon) and Art v 2 from *Artemis vulgaris* (mugwort / ragweed): Reported positive IgE skin prick tests to Cuc m 3 from *Cucumis melo* (cantaloupe or muskmelon) in 2 / 17 subjects and Cuc m 3 IgE binding to sera from 12 of 17 patients allergic to melon (Asensio *et al*., 2004). A separate study performed skin prick tests in 19 patients allergic to mugwort and 10 control patients, showing an Art v 2 sensitization prevalence of 58% (Arilla *et al.,* 2007). Therefore there may be potential for cross-reactions upon consumption for individuals sensitized to Cuc m 3 from *Cucumis melo* (cantaloupe or muskmelon) and sensitized via the inhalation route to airborne Art v 2 from *Artemis vulgaris*. The remaining 3 host cell proteins were found to have no significant hits (>35% sequence homology, E-value threshold 0.001) against known protein allergens, or auto-allergens.

There are no indications of any cross-reactivity with known celiac inducing peptides for any of the 5 host cell proteins analyzed using the AllergenOnline celiac database tools or the AllerCatPro 2.0 tools. The results of these evaluations indicate that the host cell proteins from *K. phaffii* are unlikely to present unacceptable risks of allergy or allergic cross-reactivity, compared to risks presented by foods containing ingredients from other yeasts and molds approved for food use.

6.5.6. Summary of Allergenicity Assessment

Searches in WHO/IUIS, AllergenOnline, Allergome, or COMPARE allergen databases found no entries for any allergens from the species *Dioscoreophyllum cumminsii* and no entries for the native protein monellin. An *in silico* allergenicity predictivity tools in AllergenOnline and in AllerCatPro 2.0 found no evidence for the allergenicity potential for MON and for the monellin entries from *Dioscoreophyllum cumminsii* listed in the NCBI protein database, including monellin chain A and monellin chain B.

There were no entries found for proteins from *K. phaffii* in the WHO/IUIS, AllergenOnline, Allergome, or COMPARE allergen databases. Of the five host cell proteins known to be present in SBSP at >1% total protein, the *in silico* AllergenOnline and AllerCatPro 2.0 allergenicity

predictivity tools found some indicative evidence of allergenicity for two proteins and no evidence of gluten-related adverse reactions. This indicative evidence appears to be of limited relevance as the IgE prevalence data was very low and related to homology for Cu, Zn SOD and SCP domain-containing protein with inhalation allergens and proteins found ubiquitously in a wide range of organisms to which consumers are commonly and frequently exposed. These proteins are also present in low abundance in SBSP.

The investigations do not indicate that either MON nor SBSP containing MON derived from a modified *K. phaffii* production strain are expected to have potential for allergenicity, cross-reactive adverse allergic reactions, or adverse gluten-related adverse reactions.

6.6. Safety Assessment

The subject of the GRAS assessment is serendipity berry sweet protein (SBSP) containing single-chain monellin (MON) expressed in *Komagataella phaffii* (*K. phaffii*) and produced via precision fermentation. SBSP containing MON is intended for use as a general-purpose sweetener in foods, excluding infant formula and meat and poultry products, at levels determined by GMP. SBSP is produced following cGMP. The estimated daily intakes were calculated assuming that MON will replace all high-intensity sweetener intake, and that MON is approximately 2700-fold sweeter than sucrose on a weight basis. The estimated daily intakes for "high" consumers among populations of children and adults with diabetes or without diabetes are up to 0.58 mg/kg bw/day for SBSP, and up to 0.37 mg/kg bw/day for MON. The safety of SBSP was evaluated by considering safety of the native plant-derived monellin, safety of the host organism (*K. phaffii*) used to produce MON, safety of the production strain expressing MON, and safety of the SBSP product.

Serendipity berries have been consumed by indigenous communities in West Africa for years; monellin is a sweet protein isolated from the fruit of the berry. While there is historical consumption of the berry, there are limited data in the literature on the safety of serendipity berry or native monellin and its potential for toxicity or allergenicity when used as an ingredient. There is no evidence of consumption of serendipity berry in the United States.

The parental organism, *K. phaffii,* is a nonpathogenic, non-toxigenic, and well-characterized yeast with a history of safe use in the food industry. The parental strain is derived from well characterized host organism, *K. phaffii* NRRL Y-11430. The safety of the production strain was established through confirmation of the absence of plasmids containing antibiotic resistance genes and maintenance of genomic stability throughout the production process. Application of the Pariza and Johnson decision tree (Pariza & Johnson, 2001) concludes that *K. phaffii* is accepted as a safe and suitable production strain for a food ingredient.

The safety of SBSP containing MON was examined comprehensively through assessments of *in silico* allergenicity, *in vitro* and *in silico* protein digestion*, in vitro* genotoxicity (reverse mutation and mammalian micronucleus assays), and 14-day and 90-day oral (dietary) toxicity studies in rats. The results of the assessments are as follows:

- Searches in WHO/IUIS, AllergenOnline, Allergome, or COMPARE allergen databases found no entries for any allergens from the species *Dioscoreophyllum cumminsii* and no entries for the native protein monellin, and no entries for proteins from *K. phaffii*.
- There was no indication of allergenicity of native monellin and MON in SBSP in the *silico* analyses. Likewise, the investigations indicate that neither native monellin chain A or native monellin chain B are expected to have potential for allergenicity or cross-reactive adverse allergic reactions.
- Some indicative evidence of allergenicity for two host cell proteins was identified in the bioinformatics allergenicity analysis. These findings are of limited relevance, however, as the IgE prevalence data were very low and related to homology for Cu, Zn SOD and SCP domain-containing protein with inhalation allergens and proteins found ubiquitously in a wide range of organisms to which consumers are commonly and frequently exposed. These proteins are also present in low abundance in SBSP.
- An assessment of *in silico* toxigenicity found that neither MON or the 5 residual host cell proteins from the production strain present at >1% w/w in SBSP were confirmed toxins, thus suggesting no evidence of toxigenic potential for MON and the residual proteins.
- Results from both *in vitro* and *in silico* protein digestibility assessments indicate that MON in SBSP is not resistant to digestion, would be digested to a significant degree by human pancreatic and intestinal enzymes upon ingestion, and therefore likelihood to pose a toxigenic or allergenic risk to consumers is reduced (Freeman *et al*., 2024).
- SBSP was shown to be non-genotoxic in the *in vitro* assays and no adverse effects were reported in the 14-day or 90-day toxicity studies up to the highest dose tested. The 90-day toxicity study supports a no observed adverse effect level (NOAEL) for SBSP of 1954 mg/kg bw/day, which corresponds to a NOAEL for MON (21% purity) of 408 mg/kg bw/day (Freeman *et al*., 2024).
- Reported findings from published *in vivo* studies on another single-chain monellin produced from engineered *K. phaffii* indicate a lack of toxicity, which provides corroborative evidence for the absence of toxicity from consumption of SBSP (Novik *et al*., 2023; Rega *et al*., 2017). Findings from published *in vitro* studies indicate no genotoxicity of single-chain monellin produced from engineered *K. phaffii* or engineered *Escherichia coli (E. coli)*, which provide corroborative evidence for the absence of genotoxicity from consumption of SBSP.
- A 100-fold uncertainty factor was applied to the NOAEL to account for interspecies (rat to human) and intraspecies (human and human) differences.
- The acceptable daily intake (ADI) of MON in SBSP was calculated from the adjusted NOAEL of 408 mg/kg bw/day in males from the 90-day toxicity study and the 100-fold uncertainty factor, resulting in and ADI of 4.1 mg/kg bw/day.
- Based on an intense sweetener intake assessment methodology (Renwick, 2008) and using a relative sweetness of 2700X between MON and sucrose and 1700X between SBSP and sucrose (SBSP is approximately 63% MON by weight), the estimated daily intake (EDI) of SBSP and MON were calculated for average and "high" consumers:
    - The EDI of SBSP for "high" consumers in populations of non-diabetic adults, diabetic adults, non-diabetic children, and diabetic children are up to 0.40 mg/kg bw/day, 0.53 mg/kg bw/day, 0.58 mg/kg bw/day, and 0.53 mg/kg bw/day, respectively.

- The EDI of MON for "high" consumers in populations of non-diabetic adults, diabetic adults, non-diabetic children, and diabetic children are up to 0.25 mg/kg bw/day, 0.33 mg/kg bw/day, 0.37 mg/kg bw/day, and 0.34 mg/kg bw/day, respectively.
- The EDIs of SBSP and MON by the population of "high" consumer non-diabetic children therefore provide a conservative approach for evaluating the safety of the intended use. The estimated daily intakes for "high" consumers among populations of children and adults with diabetes or without diabetes are up to 0.58 mg/kg bw/day for SBSP, and up to 0.37 mg/kg bw/day for MON.
- The EDI for MON in SBSP of 0.37 mg/kg bw/day for the population of "high" consumer non-diabetic children is below the ADI for MON in SBSP of 4.1 mg/kg bw/day. Since the EDI is below the ADI, the proposed use can be concluded to be safe.

6.7. Safety Conclusion

The intended use of SBSP containing MON is use as a general-purpose sweetener in foods, excluding infant formula and meat and poultry products. The estimated daily intakes for "high" consumers among populations of children and adults with diabetes or without diabetes are up to 0.37 mg/kg bw/day MON. The ADI for MON is 4.1 mg/kg/bw/day. The EDI of MON from the intended use of SBSP is well below the ADI of MON, therefore the intended use can be concluded to be safe. This conclusion of safety is also supported by evidence for the protein indicating it is unlikely to pose a toxigenic risk or allergenic risk to consumers. Therefore, it can be concluded that the intended use of SBSP containing MON expressed in *K. phaffii* in foods and beverages is safe within the meaning of the FD&C Act, i.e., meets the standard of reasonable certainty of no harm.

6.8. GRAS Panel Evaluation

Oobli concluded that SBSP is GRAS for its intended use as a general-purpose sweetener in foods, as described in Section 1.4, on the basis of scientific procedures. Oobli also convened a GRAS Panel to conduct a critical and comprehensive evaluation of the available pertinent data and safety information to determine, under the conditions of intended use as a general-purpose sweetener, if serendipity berry sweet protein (SBSP) containing single polypeptide chain monellin (MON) expressed in Komagataella phaffii (K. phaffii) and produced via precision fermentation, would be considered "generally recognized as safe" ("GRAS") based on scientific procedures. The GRAS Panel was composed of experts who are qualified by scientific training and experience to evaluate the safety of food ingredients. The GRAS Panel consisted of the following qualified scientific experts: Michael W. Pariza, Ph.D. (Professor Emeritus, Department of Food Science, University of Wisconsin-Madison), Richard E. Goodman, Ph.D. (Food Allergy Research and Resource Program, University of Nebraska-Lincoln), and Thomas Vollmuth, Ph.D. (Vollmuth and Associates, LLC). The GRAS Panel independently and critically evaluated all data and information presented herein, and concluded that SBSP is GRAS for its intended use as a general-purpose sweetener in foods as described in Section 1.4, based on scientific procedures.

6.9. Discussion of Information Inconsistent with a GRAS Determination

No information has been identified that would be inconsistent with a finding that the proposed use of SBSP containing MON expressed in *K. phaffii*, meeting appropriate specifications specified herein and used according to good manufacturing practices (GMP), is GRAS.

6.10. Basis for Conclusion that there is Consensus Regarding Safety

The intended use of SBSP containing MON expressed in *K. phaffii* as a general-purpose sweetener in foods, excluding infant formula and meat and poultry products, has been determined to be safe through scientific procedures as set forth in 21 CFR§170.30(b), thus satisfying the technical element of the GRAS determination. Because this safety evaluation was based on generally available and widely accepted data and information, it also satisfies the common knowledge element of a GRAS determination.

Part 7. List of Supporting Data and Information

References

Agboola DA, Fawibe OO, Ogunyale OG, Ajiboye AA. Botanical and protein sweeteners. J Adv Lab Res Bio. 2014;5(4):169-187.

AllerCatPro 2.0: a web server for predicting protein allergenicity potential. Nguyen MN, Krutz NL, Limviphuvadh V, Lopata AL, Gerberick GF, Maurer-Stroh S. Nucleic Acids Research. 2022. doi: 10.1093/nar/gkac446.

AllergenOnline Allergen Database Food Allergy Research and Resource Program (FARRP) University of Nebraska 2023. http://www.allergenonline.com/

AllergenOnline Celiac Database, Food Allergy Research and Resource Program (FARRP) University of Nebraska 2023. http://www.allergenonline.org/celiachome.shtml

Allergome https://www.allergome.org/script/search_step1.php?clear=1

American Type Culture Collection (ATCC). *Komagataella phaffii* Kurtzman, https://www.atcc.org/

Arilla MC, Ibarrola I, Puente Y, Daza JC, Martínez A, Asturias JA. Cloning, expression and characterization of mugwort pollen allergen Art v 2, a pathogenesis-related protein from family group 1. Mol Immunol. 2007;44(15):3653-60. doi: 10.1016/j.molimm.2007.04.022.

Asensio T, Crespo JF, Sanchez-Monge R, Lopez-Torrejon G, Somoza ML, Rodriguez J, Salcedo G. Novel plant pathogenesis-related protein family involved in food allergy. J Allergy Clin Immunol. 2004;114(4):896-9. doi: 10.1016/j.jaci.2004.06.014.

Barone GD, Emmerstorfer-Augustin A, Biundo A, Pisano I, Coccetti P, Mapelli V, Camattari A. Industrial production of proteins with *Pichia pastoris-Komagataella phaffii*. Biomolecules. 2023;13(3):441. doi: 10.3390/biom13030441.

Barre A, Caze-Subra S, Gironde C, Bienvenu F, Bienvenu J, Rougé P. Allergénicité des protéines édulcorantes. Rev Fr D'Allerg. 2015;55(5):363-371. https://doi.org/10.1016/j.reval.2014.05.010.

Bernauer L, Radkohl A, Lehmayer LGK, Emmerstorfer-Augustin A. *Komagataella phaffii* as emerging model organism in fundamental research. Front Microbiol. 2021;11:607028. doi: 10.3389/fmicb.2020.607028.

Bilal M, Ji L, Xu S, Zhang Y, Iqbal HMN, Cheng H. Bioprospecting and biotechnological insights into sweet-tasting proteins by microbial hosts-a review. Bioengineered. 2022;13(4):9815-9828. doi: 10.1080/21655979.2022.2061147.

Brodkorb A, Egger L, Alminger M, Alvito P, Assunção R, Ballance S, Bohn T, Bourlieu-Lacanal C, Boutrou R, Carrière F, Clemente A, Corredig M, Dupont D, Dufour C, Edwards C, Golding M, Karakaya S, Kirkhus B, Le Feunteun S, Lesmes U, Macierzanka A, Mackie AR, Martins C, Marze S, McClements DJ, Ménard O, Minekus M, Portmann R, Santos CN, Souchon I, Singh RP, Vegarud GE, Wickham MSJ, Weitschies W, Recio I. INFOGEST static *in vitro* simulation of gastrointestinal food digestion. Nat Protoc. 2019;14(4):991-1014. doi: 10.1038/s41596-018-0119-1.

Callahan A, Leonard H, Powell T. Nutrition: Science and Everyday Application, v. 2.0. 2022. https://openoregon.pressbooks.pub/nutritionscience2e/.

Codex Alimentarius Commission (CAC). Foods derived from modern biotechnology, Second edition. Joint FAO/WHO Food Standards Programme. 2009, Rome. https://www.fao.org/3/a1554e/a1554e.pdf.

Codex Alimentarius Commission (CAC). Guideline for the conduct of food safety assessment of foods Produced using recombinant-DNA microorganisms, CAC/GL-46 2003. 2003, Rome. 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

COMprehensive Protein Allergen Resource (COMPARE) Database http://db.comparedatabase.org.

EFSA Panel on Biological Hazards (BIOHAZ); Ricci A, Allende A, Bolton D, Chemaly M, Davies R, Girones R, Koutsoumanis K, Lindqvist R, Nørrung B, Robertson L, Ru G, Fernández Escámez PS, Sanaa M, Simmons M, Skandamis P, Snary E, Speybroeck N, Ter Kuile B, Threlfall J, Wahlström H, Cocconcelli PS, Peixe L, Maradona MP, Querol A, Suarez JE, Sundh I, Vlak J, Barizzone F, Correia S, Herman L. Update of the list of QPS-recommended biological agents intentionally added to food or feed as notified to EFSA 7: suitability of taxonomic units notified to EFSA until September 2017. EFSA J. 2018;16(1):e05131. doi: 10.2903/j.efsa.2018.5131.

Esposito V, Guglielmi F, Martin SR, Pauwels K, Pastore A, Piccoli R, Temussi PA. Aggregation mechanisms of cystatins: a comparative study of monellin and oryzacystatin. Biochemistry. 2010;49(13):2805-10. doi: 10.1021/bi902039s.

Esposito V, Temussi PA. Cystatins: a versatile family. Biomol Concepts. 2011;2(1-2):95-102. doi: 10.1515/bmc.2011.001.

FAO/WHO. Evaluation of Allergenicity of Genetically Modified Foods. Report of a Joint FAO/WHO Expert Consultation on Allergenicity of Foods Derived From Biotechnology. Food and Agriculture Organization of the United Nations, January 22-25, 2001, Rome, Italy. 29 p.

FAO/WHO. Foods derived from modern biotechnology Second edition Codex Alimentarius Rome, 2009 ISBN 978-92-5-105914-2 https://www.fao.org/3/a1554e/a1554e.pdf.

FAO/WHO. Revision of Environmental Health Criteria 240 for the evaluation of enzyme preparations used in the manufacture of foods Section 9.1.4.2 of Chapter 9, Principles Related to Specific Groups of Substances, of Environmental Health Criteria 240 published in Evaluation of certain food additives: eighty-ninth report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva: World Health Organization and Food and Agriculture Organization of the United Nations; 2020 (WHO Technical Report Series, No. 1027).

Fernstrom JD, Munger SD, Sclafani A, de Araujo IE, Roberts A, Molinary S. Mechanisms for sweetness. J Nutr. 2012;142(6):1134S-41S. doi: 10.3945/jn.111.149567.

Freeman EL, Ward R, Murphy MM, Wang T, Ryder J. Comprehensive safety assessment of serendipity berry sweet protein produced from *Komagataella phaffii*. Regul Toxicol Pharmacol. 2024;147:105562. doi: 10.1016/j.yrtph.2024.105562.

Fry C. Natural low-calorie sweeteners. In: Woodhead Publishing Series in Food Science, Technology and Nutrition, Natural Food Additives, Ingredients and Flavourings, Editor(s): David Baines, Richard Seal. Woodhead Publishing, 2012, pp 41-75, ISBN 9781845698119, https://doi.org/10.1533/9780857095725.1.41.

Gitler C. Chapter 2: Protein Digestion and Absorption in Nonruminants. In: Mammalian Protein Metabolism Volume I, 1964, Pages 35-69. Protein Digestion and Absorption in Nonruminants. Editor: Munro HN. Academic Press.

Govaerts R, Nic Lughadha E, Black N, Turner R, Paton A. The World Checklist of Vascular Plants, a continuously updated resource for exploring global plant diversity. Sci Data. 2021;8(1):215. Doi: 10.1038/s41597-021-00997-6. Accessed 26 June 2023.

Iyengar RB, Smits P, van der Ouderaa F, van der Wel H, van Brouwershaven J, Ravestein P, Richters G, van Wassenaar PD. The complete amino-acid sequence of the sweet protein thaumatin I. Eur J Biochem. 1979;96(1):193-204. doi: 10.1111/j.1432-1033.1979.tb13029.x.

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(1):1700297. doi: 10.1002/mnfr.201700297.

Joye I. Protein digestibility of cereal products. Foods. 2019;8(6):199. doi: 10.3390/foods8060199.

Kant R. Sweet proteins--potential replacement for artificial low calorie sweeteners. Nutr J. 2005;4:5. doi: 10.1186/1475-2891-4-5.

Kim SH, Kang CH, Kim R, Cho JM, Lee YB, Lee TK. Redesigning a sweet protein: increased stability and renaturability. Protein Eng. 1989;2(8):571-5. doi: 10.1093/protein/2.8.571.

Kinghorn AD, Compardre CM. Less Common High-Potency Sweeteners. In: Alternative Sweeteners, 4th Edition, Ed: Nabors LO. 2012, p. 229.

Kojima I, Nakagawa Y. The role of the sweet taste receptor in enteroendocrine cells and pancreatic β-cells. Diabetes Metab J. 2011;35(5):451-7. doi: 10.4093/dmj.2011.35.5.451.

Laffitte A, Neiers F, Briand L. Functional roles of the sweet taste receptor in oral and extraoral tissues. Curr Opin Clin Nutr Metab Care. 2014;17(4):379-85. doi: 10.1097/MCO.0000000000000058.

Macfarlane GT, Cummings JH, Allison C. Protein degradation by human intestinal bacteria. J Gen Microbiol. 1986;132(6):1647-56. doi: 10.1099/00221287-132-6-1647.

Maurer-Stroh S, Krutz NL, Kern PS, Gunalan V, Nguyen MN, Limviphuvadh V, Eisenhaber F, Gerberick GF. AllerCatPro-prediction of protein allergenicity potential from the protein sequence. Bioinformatics. 2019;35(17):3020-3027. doi: 10.1093/bioinformatics/btz029.

Metcalfe DD, Astwood JD, Townsend R, Sampson HA, Taylor SL, Fuchs RL. Assessment of the allergenic potential of foods derived from genetically engineered crop plants. Crit Rev Food Sci Nutr. 1996;36 Suppl:S165-86. doi: 10.1080/10408399609527763.

Murzin AG. Sweet-tasting protein monellin is related to the cystatin family of thiol proteinase inhibitors. J Mol Biol. 1993;230(2):689-94. doi: 10.1006/jmbi.1993.1186.

Naegeli H, Birch AN, Casacuberta J, De Schrijver A, Gralak MA, Guerche P, Jones H, Manachini B, Messéan A, Nielsen EE, Nogué F, Robaglia C, Rostoks N, Sweet J, Tebbe C, Visioli F, Wal JM, Eigenmann P, Epstein M, Hoffmann-Sommergruber K, Koning F, Lovik M, Mills C, Moreno FJ, van Loveren H, Selb R, Fernandez Dumont A. EFSA Panel on Genetically Modified Organisms (GMO). Guidance on allergenicity assessment of genetically modified plants. EFSA J. 2017;15(6):e04862. doi: 10.2903/j.efsa.2017.4862.

Nguyen MN, Krutz NL, Limviphuvadh V, Lopata AL, Gerberick GF, Maurer-Stroh S. AllerCatPro 2.0: a web server for predicting protein allergenicity potential. Nucleic Acids Res. 2022 Jul 5;50(W1):W36-W43. doi: 10.1093/nar/gkac446.

Novik TS, Koveshnikova EI, Kotlobay AA, Sycheva LP, Kurochkina KG, Averina OA, Belopolskaya MV, Sergiev PV, Dontsova OA, Lazarev VN, Maev IV, Kostyaeva MG, Eremeev AV, Chukina SI, Lagarkova MA. Sweet-tasting natural proteins brazzein and monellin: Safe sugar substitutes for the food industry. Foods. 2023;12(22):4065. doi: 10.3390/foods12224065.

OECD, 2008. Test No. 407: Repeated dose 28-day oral toxicity study in rodents. www.oecd-ilibrary.org/content/publicatio n/9789264070684-en. doi: 10.1787/9789264070684-en.

OECD, 2018. Test No. 408: Repeated dose 90-day oral toxicity study in rodents. www.oecd-ilibrary.org/content/publicatio n/9789264070707-en. doi: 10.1787/9789264070707-en.

OECD, 2023. Test No. 487 *In vitro* mammalian cell micronucleus test. www.oecd-ilibrary.org/content/publication/9789264 264861-en. doi 10.1787/9789264264861-en.

Pariza MW, Johnson EA. Evaluating the safety of microbial enzyme preparations used in food processing: update for a new century. Regul Toxicol Pharmacol. 2001;33(2):173-86. doi: 10.1006/rtph.2001.1466.

Picone D, Temussi PA. Dissimilar sweet proteins from plants: oddities or normal components? Plant Sci. 2012;195:135-42. doi: 10.1016/j.plantsci.2012.07.001.

Portune KJ, Benítez-Páez A, Del Pulgar EM, Cerrudo V, Sanz Y. Gut microbiota, diet, and obesity-related disorders-the good, the bad, and the future challenges. Mol Nutr Food Res. 2017;61(1). doi: 10.1002/mnfr.201600252.

Rega MF, Siciliano A, Gesuele R, Lofrano G, Carpentieri A, Picone D, Guida M. Ecotoxicological survey of MNEI and Y65R-MNEI proteins as new potential high-intensity sweeteners. Environ Sci Pollut Res Int. 2017;24(10):9734-9740. doi: 10.1007/s11356-017-8626-0.

Renukuntla J, Vadlapudi AD, Patel A, Boddu SH, Mitra AK. Approaches for enhancing oral bioavailability of peptides and proteins. Int J Pharm. 2013;447(1-2):75-93. doi: 10.1016/j.ijpharm.2013.02.030.

Renwick AG. The use of a sweetener substitution method to predict dietary exposures for the intense sweetener rebaudioside A. Food Chem Toxicol. 2008;46 Suppl 7:S61-9. doi: 10.1016/j.fct.2008.05.009.

Reyes TF, Agrawal P, Chan T, Green R, Matulka RA. The safety of soy leghemoglobin protein preparation derived from *Pichia pastoris* expressing a soy leghemoglobin gene from *glycine max*: *In vitro* and *in vivo* studies. J Toxicol. 2023;2023:7398724.

Reyes TF, Chen Y, Fraser RZ, Chan T, Li X. Assessment of the potential allergenicity and toxicity of Pichia proteins in a novel leghemoglobin preparation. Regul Toxicol Pharmacol. 2021;119:104817. doi: 10.1016/j.yrtph.2020.104817.

Saraiva A, Carrascosa C, Ramos F, Raheem D, Pedreiro S, Vega A, Raposo A. Brazzein and monellin: Chemical analysis, food industry applications, safety and quality control, nutritional profile and health impacts. Foods. 2023;12(10):1943. doi: 10.3390/foods12101943.

Scheurer S, Son DY, Boehm M, Karamloo F, Franke S, Hoffmann A, Haustein D, Vieths S. Cross-reactivity and epitope analysis of Pru a 1, the major cherry allergen. Mol Immunol. 1999;36(3):155-67. doi: 10.1016/s0161-5890(99)00033-4.

Smith NJ, Grant JN, Moon JI, So SS, Finch AM. Critically evaluating sweet taste receptor expression and signaling through a molecular pharmacology lens. FEBS J. 2021;288(8):2660-2672. doi: 10.1111/febs.15768.

Steinhardt J, Zaiser EM. Hydrogen ion equilibria in native and denatured proteins. Adv Protein Chem. 1955;10:151-226. doi: 10.1016/s0065-3233(08)60105-9.

Sturmberger L, Chappell T, Geier M, Krainer F, Day KJ, Vide U, Trstenjak S, Schiefer A, Richardson T, Soriaga L, Darnhofer B, Birner-Gruenberger R, Glick BS, Tolstorukov I, Cregg J, Madden K, Glieder A. Refined *Pichia pastoris* reference genome sequence. J Biotechnol. 2016;235:121-31. doi: 10.1016/j.jbiotec.2016.04.023.

Temussi PA. Why are sweet proteins sweet? Interaction of brazzein, monellin and thaumatin with the T1R2-T1R3 receptor. FEBS Lett. 2002;526(1-3):1-4. doi: 10.1016/s0014-5793(02)03155-1.

Terashima H, Ichikawa M. A comparative ethnobotany of the Mbuti and Efe hunter-gatherers in the Ituri forest, Democratic Republic of Congo. African Study Monographs. 2003;24;1-168. doi: 10.14989/68220.

Thomas K, Aalbers M, Bannon GA, Bartels M, Dearman RJ, Esdaile DJ, Fu TJ, Glatt CM, Hadfield N, Hatzos C, Hefle SL, Heylings JR, Goodman RE, Henry B, Herouet C, Holsapple M, Ladics GS, Landry TD, MacIntosh SC, Rice EA, Privalle LS, Steiner HY, Teshima R, Van Ree R, Woolhiser M, Zawodny J. A multi-laboratory evaluation of a common *in vitro* pepsin digestion assay protocol used in assessing the safety of novel proteins. Regul Toxicol Pharmacol. 2004;39(2):87-98. doi: 10.1016/j.yrtph.2003.11.003.

U.S. Food and Drug Administration (FDA). FDA GRAS Notice (GRN) No. 737. Soy leghemoglobin preparation from a strain of *Pichia pastoris*. 2018. https://www.fda.gov/Food/IngredientsPackagingLabeling/GRAS/NoticeInventory/default.htm. Date of closure: Jul 23, 2018.

U.S. Food and Drug Administration (FDA). FDA GRAS Notice (GRN) No. GRN 1056. 2022. β-lactoglobulin produced by *Komagataella phaffii* strain "yRMK-66". https://www.cfsanappsexternal.fda.gov/scripts/fdcc/index.cfm?set=GRASNotices&id=1056. Date of closure: Feb 15, 2023.

U.S. Food and Drug Administration (FDA). FDA GRAS Notice (GRN) No. GRN 1025. 2021. Pepsin A enzyme preparation produced by *Pichia pastoris* to overexpress the gene encoding pepsin A. https://www.cfsanappsexternal.fda.gov/scripts/fdcc/index.cfm?set=GRASNotices&id=1025. Date of closure: Jan 6, 2023.

U.S. Food and Drug Administration (FDA). FDA GRAS Notice (GRN) No. GRN 204. 2006. Phospholipase C enzyme preparation from *Pichia pastoris* expressing a heterologous phospholipase C gene. https://www.cfsanappsexternal.fda.gov/scripts/fdcc/?set=GRASNotices&id=204&sort=GRN_No&order=DESC&startrow=1&type=basic&search=204. Date of closure: Dec 5, 2006.

U.S. Food and Drug Administration (FDA). FDA GRAS Notice (GRN) No. GRN 1104. 2022. Egg-white protein produced by *Komagataella phaffii* ATCC GSD-1235. https://www.cfsanappsexternal.fda.gov/scripts/fdcc/?set=GRASNotices&id=1104&sort=GRN_No&order=DESC&startrow=1&type=basic&search=1104. Date of closure: Oct 17, 2023.

U.S. Food and Drug Administration (FDA). FDA GRAS Notice (GRN) No. GRN 1001. 2021. Myoglobin preparation from a strain of *Pichia pastoris* expressing the myoglobin gene from Bos taurus. https://www.cfsanappsexternal.fda.gov/scripts/fdcc/?set=GRASNotices&id=1001&sort=GRN_No&order=DESC&startrow=1&type=basic&search=1001. Date of closure: Dec 3, 2021.

U.S. Food and Drug Administration (FDA). FDA GRAS Notice (GRN) No. GRN 967. 2022. Soluble egg-white protein produced by *Komagataella phaffii* strain GSD-1209. https://www.cfsanappsexternal.fda.gov/scripts/fdcc/?set=GRASNotices&id=967&sort=GRN_No&order=DESC&startrow=1&type=basic&search=967. Date of closure: Sept 9, 2021.

U.S. Food and Drug Administration (FDA). International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). 2012. S2(R1) Guidance on Genotoxicity Testing and Data Interpretation for Pharmaceuticals Intended for Human Use (2012). https://www.fda.gov/media/71980/download.

U.S. Food and Drug Administration (FDA). Toxicological Principles for the Safety Assessment of Food Ingredients, Redbook 2000, IV.C. 1. A. 2007. “Bacterial Reverse Mutation Test.

van Ree R, Sapiter Ballerda D, Berin MC, Beuf L, Chang A, Gadermaier G, Guevera PA, Hoffmann-Sommergruber K, Islamovic E, Koski L, Kough J, Ladics GS, McClain S, McKillop KA, Mitchell-Ryan S, Narrod CA, Pereira Mouriès L, Pettit S, Poulsen LK, Silvanovich A, Song P, Teuber SS, Bowman C. The COMPARE database: A public resource for allergen identification, adapted for continuous improvement. Front Allergy. 2021;2:700533. doi: 10.3389/falgy.2021.700533.

World Health Organization and International Union of Immunological Societies (WHO/IUIS) Allergen Nomenclature Sub-committee. No other form or system of allergen nomenclature is recognized by the WHO or the IUIS. http://allergen.org/

Zhou CE, Smith J, Lam M, Zemla A, Dyer MD, Slezak T. MvirDB--a microbial database of protein toxins, virulence factors and antibiotic resistance genes for bio-defence applications. Nucleic Acids Res. 2007;35(Database issue):D391-4. doi: 10.1093/nar/gkl791.

Appendix A. Proteomic Characterization of SBSP

All proteins present in SBSP were identified in three individual production batches and their relative amount was determined using LC-MSMS (Appendix A, Table 1). The host cell proteins detected in SBSP were confirmed to originate from *K. phaffii* and their gene names were identified. Of the proteins present in SBSP, 100% of proteins present were identified in this characterization.

Appendix A, Table 1. Characterization of production strain host cell proteins in 3 production batches of SBSP

Rank	Identified Proteins	Molecular Weight	Relative % protein present				
			Batch 1	Batch 2	Batch 3	Average	Cumulative % total protein
1	MON	11 kDa	87.00	76.79	89.65	84.48	84.48
2	Superoxide dismutase [Cu-Zn] OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0786 PE=3 SV=1	16 kDa	1.05	3.46	2.62	2.38	86.85
3	SCP domain-containing protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0076 PE=4 SV=1	32 kDa	2.09	2.68	2.33	2.37	89.22
4	Chitin deacetylase, together with Cda1p involved in the biosynthesis ascospore wall component OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-1_0103 PE=4 SV=1	35 kDa	1.64	2.57	2.04	2.08	91.31
5	Fusion protein, identical to Rpl40Bp OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0486 PE=3 SV=1	15 kDa	2.54	2.57	0.87	1.99	93.30
6	ACB domain-containing protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-4_0674 PE=4 SV=1	10 kDa	1.20	1.45	1.46	1.37	94.67

Rank	Identified Proteins	Molecular Weight	Relative % protein present				
			Batch 1	Batch 2	Batch 3	Average	Cumulative % total protein
7	Glyceraldehyde-3-phosphate dehydrogenase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0437 PE=3 SV=1	36 kDa	1.05	1.00	0.29	0.78	95.45
8	Protein TOS1 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_c034_0034 PE=3 SV=1	45 kDa	0.90	0.45	0.00	0.45	95.90
9	FK506-binding protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-2_0476 PE=3 SV=1	46 kDa	0.75	0.45	0.15	0.45	96.34
10	Transaldolase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-2_0337 PE=3 SV=1	36 kDa	0.15	1.12	0.00	0.42	96.76
11	Nuclear protein required for transcription of MXR1 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0102 PE=4 SV=1	47 kDa	0.15	0.89	0.00	0.35	97.11
12	Uncharacterized protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-1_0118 PE=4 SV=1	46 kDa	0.30	0.45	0.29	0.35	97.46

Rank	Identified Proteins	Molec-ular Weight	Relative % protein present				
			Batch 1	Batch 2	Batch 3	Aver-age	Cumu-lative % total protein
13	6-phosphogluconate dehydrogenase, decarboxylating OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0277 PE=3 SV=1	54 kDa	0.30	0.45	0.15	0.30	97.76
14	Phosphopyruvate hydratase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0082 PE=3 SV=1	47 kDa	0.30	0.56	0.00	0.29	98.04
15	Actin-binding protein of the cortical actin cytoskeleton OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-2_0320 PE=4 SV=1	61 kDa	0.15	0.56	0.15	0.28	98.33
16	Suppressor protein STM1 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-4_0297 PE=4 SV=1	30 kDa	0.15	0.56	0.00	0.24	98.56
17	5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0160 PE=3 SV=1	86 kDa	0.00	0.67	0.00	0.22	98.78
18	Protein that binds to cruciform DNA structures OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0627 PE=4 SV=1	32 kDa	0.15	0.45	0.00	0.20	98.98

Rank	Identified Proteins	Molecular Weight	Relative % protein present				
			Batch 1	Batch 2	Batch 3	Average	Cumulative % total protein
19	ATPase involved in protein folding and the response to stress OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0230 PE=3 SV=1	71 kDa	0.00	0.56	0.00	0.19	99.17
20	Eisosome protein 1 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0722 PE=3 SV=1	91 kDa	0.00	0.45	0.00	0.15	99.32
21	Protein component of the large (60S) ribosomal subunit OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-1_0219 PE=3 SV=1	27 kDa	0.15	0.22	0.00	0.12	99.44
22	Vacuolar proteinase B (YscB), a serine protease of the subtilisin family OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-1_0226 PE=3 SV=1	59 kDa	0.00	0.33	0.00	0.11	99.55
23	Uncharacterized protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0618 PE=4 SV=1	62 kDa	0.00	0.33	0.00	0.11	99.67
24	inorganic diphosphatase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-3_0028 PE=3 SV=1	32 kDa	0.00	0.33	0.00	0.11	99.78

Rank	Identified Proteins	Molec-ular Weight	Relative % protein present				
			Batch 1	Batch 2	Batch 3	Aver-age	Cumu-lative % total protein
25	Primary component of eisosomes OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-4_0569 PE=4 SV=1	35 kDa	0.00	0.22	0.00	0.07	99.85
26	Deoxyuridine 5'-triphosphate nucleotidohydrolase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0478 PE=3 SV=1	17 kDa	0.00	0.22	0.00	0.07	99.93
27	ATPase involved in protein folding and nuclear localization signal (NLS)-directed nuclear transport OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0552 PE=3 SV=1	70 kDa	0.00	0.22	0.00	0.07	**100.00**

Appendix A, Table 2. Characterization of production strain host cell proteins in the 90-day toxicity study batch vs. average of 3 production batches of SBSP

	90-day Toxicity Study Batch		Average of 3 Production Batches		
Accession	**%**	**Rank**	**%**	**Rank**	**Protein Name**
1IV9_A	69.54	1	84.48	1	Monellin-SC
C4R8X7	2.19	5	2.38	2	Superoxide dismutase [Cu-Zn] OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0786 PE=3 SV=1
C4R3H3	2.32	4	2.37	3	SCP domain-containing protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0076 PE=4 SV=1
C4QW42	1.57	7	2.08	4	Chitin deacetylase, together with Cda1p involved in the biosynthesis ascospore wall component OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-1_0103 PE=4 SV=1
C4R0U2	1.57	8	1.99	5	Fusion protein, identical to Rpl40Bp OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0486 PE=3 SV=1
C4R2Z5	1.71	6	-	-	Protein of the SUN family (Sim1p, Uth1p, Nca3p, Sun4p) that may participate in DNA replication OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-2_0064 PE=3 SV=1
C4QY91	2.39	3	1.37	6	ACB domain-containing protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-4_0674 PE=4 SV=1
C4R0P1	0.41	18	0.78	7	Glyceraldehyde-3-phosphate dehydrogenase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0437 PE=3 SV=1
C4QVL4	1.37	9	-	-	1,3-beta-glucanosyltransferase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-3_0226 PE=3 SV=1
C4R6P9	1.02	10	-	-	Mitochondrial outer membrane and cell wall localized SUN family member OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0046 PE=3 SV=1
C4R0Z8	0.96	11	-	-	Uncharacterized protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0539 PE=4 SV=1
C4QYW7	0.89	12	-	-	Lectin-like protein with similarity to Flo1p, thought to be expressed and involved in flocculation OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-4_0584 PE=4 SV=1
C4R9F6	4.85	2	0.45	8	Protein TOS1 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_c034_0034 PE=3 SV=1

Accession	90-day Toxicity Study Batch		Average of 3 Production Batches		Protein Name
	%	Rank	%	Rank	
C4QYF3	0.75	14	-	-	Endo-beta-1,3-glucanase, major protein of the cell wall, involved in cell wall maintenance OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-4_0426 PE=3 SV=1
C4R2G3	-	-	0.45	9	FK506-binding protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-2_0476 PE=3 SV=1
C4R245	0.82	13	0.42	10	Transaldolase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-2_0337 PE=3 SV=1
C4QZC2	0.48	17	-	-	Phosphatidylglycerol/phosphatidylinositol transfer protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_FragB_0077 PE=4 SV=1
C4R6V3	-	-	0.35	11	Nuclear protein required for transcription of MXR1 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0102 PE=4 SV=1
C4QW71	0.41	19	-	-	glucan endo-1,3-beta-D-glucosidase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-1_0130 PE=3 SV=1
C4R0Q7	0.41	20	-	-	Major exo-1,3-beta-glucanase of the cell wall, involved in cell wall beta-glucan assembly OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0454 PE=3 SV=1
C4QVL5	0.34	21	-	-	1,3-beta-glucanosyltransferase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-3_0227 PE=3 SV=1
C4R2D7	0.34	22	-	-	Flocculation protein FLO11 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=FLO11 PE=1 SV=1
C4R894	0.34	23	-	-	Glycosidase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0559 PE=3 SV=1
C4QW56	-	-	0.35	12	Uncharacterized protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-1_0118 PE=4 SV=1
C4QVL7	0.27	25	-	-	Cell wall protein with similarity to glucanases OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-3_0229 PE=3 SV=1
C4QVR8	0.27	26	-	-	Mucin family member OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-3_0276 PE=4 SV=1
C4QZC5	0.2	27	-	-	Mucin-like protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_FragB_0067 PE=4 SV=1
C4QZH9	0.2	28	-	-	Cell wall protein with similarity to glucanases OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0052 PE=4 SV=1

Accession	90-day Toxicity Study Batch		Average of 3 Production Batches		Protein Name
	%	Rank	%	Rank	
C4R3C4	0.2	29	-	-	Uncharacterized protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0030 PE=4 SV=1
C4R6L9	0.2	30	-	-	Cytochrome c, isoform 1 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0018 PE=3 SV=1
C4R9F4	0.2	31	-	-	1,3-beta-glucanosyltransferase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_c034_0032 PE=3 SV=1
C4R430	-	-	0.30	13	6-phosphogluconate dehydrogenase, decarboxylating OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0277 PE=3 SV=1
C4QYT0	0.14	33	-	-	Vacuolar proteinase B (YscB), a serine protease of the subtilisin family OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-4_0548 PE=3 SV=1
C4QYX3	0.14	34	-	-	Protein component of the small (40S) subunit, essential for control of translational accuracy OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-4_0589 PE=3 SV=1
C4R1Q1	0.14	35	-	-	Uncharacterized protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0887 PE=4 SV=1
C4R300	0.14	36	-	-	Nucleoside diphosphate kinase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-2_0059 PE=3 SV=1
C4R564	0.14	37	-	-	Type I transmembrane sorting receptor for multiple vacuolar hydrolases OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0653 PE=4 SV=1
C4R6J5	0.14	38	-	-	Nuclear transport factor 2 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_1114 PE=4 SV=1
C4R7G7	0.14	39	-	-	FAS1 domain-containing protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0303 PE=4 SV=1
C4R7G9	0.14	40	-	-	O-glycosylated protein required for cell wall stability OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0305 PE=4 SV=1
C4R7H1	0.14	41	-	-	Secreted beta-glucosidase adg3 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0307 PE=3 SV=1
C4R9C9	0.14	42	-	-	Contains GLEYA adhesin domain OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_c034_0002 PE=4 SV=1
C4QV25	0.07	43	-	-	Essential RNA-binding G protein effector of mating response pathway OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-3_0045 PE=4 SV=1

Accession	90-day Toxicity Study Batch		Average of 3 Production Batches		Protein Name
	%	Rank	%	Rank	
C4QWA3	0.07	44	-	-	N-glycosylated protein involved in the maintenance of bud site selection during bipolar budding OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-1_0161 PE=4 SV=1
C4QZB0	0.07	45	-	-	Elongation factor 1-alpha OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_FragB_0052 PE=3 SV=1
C4R2B9	0.07	46	-	-	Uncharacterized protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-2_0271 PE=3 SV=1
C4R2E7	0.07	47	-	-	Thioredoxin reductase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-2_0480 PE=3 SV=1
C4R4Y8	0.07	48	-	-	ATP synthase subunit alpha OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0576 PE=3 SV=1
C4R635	0.07	49	-	-	Glycoprotein involved in cell wall beta-glucan assembly OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0960 PE=3 SV=1
C4R6S4	0.07	50	-	-	Uncharacterized protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0917 PE=4 SV=1
C4R743	0.07	51	-	-	SsuA/THI5-like domain-containing protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0193 PE=4 SV=1
Q9P4D1	0.07	52	-	-	Actin OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=ACT1 PE=1 SV=1
C4R3H8	0.27	24	0.29	14	phosphopyruvate hydratase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0082 PE=3 SV=1
C4R262	-	-	0.28	15	Actin-binding protein of the cortical actin cytoskeleton OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-2_0320 PE=4 SV=1
C4QY12	0.14	32	0.24	16	Suppressor protein STM1 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-4_0297 PE=4 SV=1
C4QZU2	-	-	0.22	17	5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0160 PE=3 SV=1
C4R537	-	-	0.20	18	Protein that binds to cruciform DNA structures OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0627 PE=4 SV=1
C4R3X8	0.55	16	0.19	19	ATPase involved in protein folding and the response to stress OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0230 PE=3 SV=1

Accession	90-day Toxicity Study Batch		Average of 3 Production Batches		Protein Name
	%	Rank	%	Rank	
C4R1J7	-	-	0.15	20	Eisosome protein 1 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0722 PE=3 SV=1
C4QWG6	-	-	0.12	21	Protein component of the large (60S) ribosomal subunit OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-1_0219 PE=3 SV=1
C4QWH2	-	-	0.11	22	Vacuolar proteinase B (YscB), a serine protease of the subtilisin family OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-1_0226 PE=3 SV=1
C4R184	-	-	0.11	23	Uncharacterized protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0618 PE=4 SV=1
C4QV10	-	-	0.11	24	inorganic diphosphatase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-3_0028 PE=3 SV=1
C4QYV1	-	-	0.07	25	Primary component of eisosomes OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-4_0569 PE=4 SV=1
C4R0T4	-	-	0.07	26	Deoxyuridine 5'-triphosphate nucleotidohydrolase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0478 PE=3 SV=1
C4R887	0.68	15	0.07	27	ATPase involved in protein folding and nuclear localization signal (NLS)-directed nuclear transport OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0552 PE=3 SV=1
TOTAL	**100.0**	**52**	**100.0**	**27**	

Appendix B. Literature Searches

0000Search	Database	Search Terms	Hits	Date
#1	PubMed	(monellin) OR (serendipity berry) OR (Dioscoreophyllum cumminsii) Filters: English	218	June 19, 2023
#2	Science Direct	(monellin OR serendipity berry OR Dioscoreophyllum cumminsii) in Title, Abstract or author-specified keywords; No other filters	89	June 20, 2023
#3	PubMed	Search: (monellin) OR (serendipity berry) OR (Dioscoreophyllum cumminsii) Filters: English Indexed since June 1, 2023	2	November 7, 2023
#4	Science Direct	Search: (monellin OR serendipity berry OR Dioscoreophyllum cumminsii) in Title, Abstract or author-specified keywords; No other filters	32	November 7, 2023
#5	PubMed	((Komagataella phaffii) OR (Pichia pastoris)) AND (safe* OR toxi* OR adverse* OR risk) Filters: English Indexed since October 1, 2022	56	November 11, 2023
#6	PubMed	((Komagataella phaffii) OR (Pichia pastoris)) AND (metabolis*[ti] OR absorb*[ti] OR absorp*[ti] OR excret*[ti] OR eliminat*[ti] OR pharmacokinetic*[ti] OR pharmacodynamic*[ti] OR toxicokinetic*[ti] OR toxicodynamic*[ti] OR bioavailab*[ti] OR biotransform*[ti] OR digest*[ti]) Filters: English Indexed since November 1, 2023	9	November 11, 2023
#7	PubMed	((Komagataella phaffii) OR (Pichia pastoris)) AND (metabolis*[ti] OR absorb*[ti] OR absorp*[ti] OR excret*[ti] OR eliminat*[ti] OR pharmacokinetic*[ti] OR pharmacodynamic*[ti] OR toxicokinetic*[ti] OR toxicodynamic*[ti] OR bioavailab*[ti] OR biotransform*[ti] OR digest*[ti]) Filters: English Indexed since November 1, 2023	3	January 29, 2024
#8	PubMed	((Komagataella phaffii) OR (Pichia pastoris)) AND (safe* OR toxi* OR adverse* OR risk) Filters: English Indexed since November 1, 2023	16	January 29, 2024
#9	PubMed	(monellin) OR (serendipity berry) OR (Dioscoreophyllum cumminsii) Filters: English Indexed since November 1, 2023	5	January 29, 2024

Appendix C. FASTA Sequences for Monellin Proteins from *Dioscoreophyllum cumminsii* (serendipity berry) and for MON in SBSP

The NCBI Protein database is a collection of sequences from several sources, including SwissProt and PIR components of UniProt; Protein Research Foundation (PRF); Protein Data Bank (PDB); and translations of coding regions on sequences in Entrez Nucleotide International Sequence Database Collaboration (DDBJ / EMBL / GenBank).

Monellin isolated from *Dioscoreophyllum cumminsii* is made up of two noncovalently associated polypeptide chains: monellin chain A contains 44 amino acid residues and monellin chain B has 50 residues. The search for monellin from *Dioscoreophyllum cumminsii* in the NCBI protein database[12] returned 85 entries. These were refined in the NCBI Identical Protein Groups database[13] to 5 entries: monellin chain A, monellin chain B, and 3 unnamed protein products. A search was also conducted for MON in SBSP.

1) Monellin Chain A

RefSeq Selected Product: P02881.1, 45 amino acids

Source	CDS Region in Nucleotide	Protein	Name
Swiss-Prot	N/A	P02881.1	Monellin chain A
PDB	N/A	4MON_A	
PDB	N/A	4MON_C	
PAT	N/A	AAC85138.1	Sequence 1 from patent US 5739409
PAT	N/A	AAE40194.1	Sequence 1 from patent US 6001410
PAT	N/A	AAQ71910.1	Sequence 1 from patent US 5487983
PAT	N/A	ABA13044.1	Sequence 4 from patent US 6913906

FASTA Sequence

>sp|P02881.1|MONA_DIOCU RecName: Full=Monellin chain A; AltName: Full=Monellin chain I FREIKGYEYQLYVYASDKLFRADISEDYKTRGRKLLRFNGPVPPP

2) Monellin Chain B

RefSeq Selected Product:P02882.2, 50 amino acids

---

[12]NCBI protein database results from search terms “monellin”[All Fields] AND (“Dioscoreophyllum cumminsii” [Organism] OR “Dioscoreophyllum cumminsii” [All Fields]) https://www.ncbi.nlm.nih.gov/protein?term=(monellin)%20AND%20Dioscoreophyllum%20cumminsii

[13] NCBI protein database Identical Protein Groups results from search terms “monellin”[All Fields] AND (“Dioscoreophyllum cumminsii” [Organism] OR “Dioscoreophyllum cumminsii” [All Fields]) https://www.ncbi.nlm.nih.gov/ipg?term=(monellin)%20AND%20Dioscoreophyllum%20cumminsii

Source	CDS Region in Nucleotide	Protein	Name
Swiss-Prot	N/A	P02882.2	Monellin chain B
PDB	N/A	1KRL_B	
PDB	N/A	1KRL_D	
PAT	N/A	AAE40196.1	Sequence 3 from patent US 6001410

FASTA Sequence
>sp|P02882.2|MONB_DIOCU RecName: Full=Monellin chain B; AltName: Full=Monellin chain II GEWEIIDIGPFTQNLGKFAVDEENKIGQYGRLTFNKVIRPCMKKTIYEEN protein product (1KRL_A; 3MON_G)

3) Unnamed Protein Product (1KRL_A; 3MON_G)

RefSeq Selected Product:1KRL_A, 44 amino acids

Source	CDS Region in Nucleotide	Protein	Name
PDB	N/A	1KRL_A	
PDB	N/A	1KRL_C	
PDB	N/A	3MON_A	
PDB	N/A	3MON_C	
PDB	N/A	3MON_E	
PDB	N/A	3MON_G	
PRF	N/A	761801A	monellin A

FASTA Sequence
>pdb|3MON|G Chain G, MONELLIN
REIKGYEYQLYVYASDKLFRADISEDYKTRGRKLLRFNGPVPPP

4) Unnamed Protein Product (3MON_B; 4MON-D)

RefSeq Selected Product:3MON_B, 50 amino acids

Source	CDS Region in Nucleotide	Protein	Name
PDB	N/A	3MON_B	
PDB	N/A	3MON_D	
PDB	N/A	3MON_F	
PDB	N/A	3MON_H	
PDB	N/A	4MON_B	
PDB	N/A	4MON_D	
PRF	N/A	761801B	monellin B
PAT	N/A	AAC85139.1	Sequence 2 from patent US 5739409
PAT	N/A	AAE40195.1	Sequence 2 from patent US 6001410

PAT	N/A	AAQ71911.1	Sequence 2 from patent US 5487983

FASTA Sequence
>pdb|4MON|D Chain D, Monellin
GEWEIIDIGPFTQNLGKFAVDEENKIGQYGRLTFNKVIRPCMKKTIYENE

5) Unnamed Protein Product (2O9U_X)

RefSeq Selected Product:2O9U_X, 97 amino acids

Source	CDS Region in Nucleotide	Protein	Name
PDB	N/A	2O9U_X	
INSDC	JQ282905.1; 1-294 (+)	AFF58925.1	monellin (synthetic construct)
PAT	N/A	AAR46452.1	Sequence 5 from patent US 6610514

FASTA Sequence
>pdb|2O9U|X Chain X, Monellin chain B and Monellin chain A
MGEWEIIDIGPFTQNLGKFAVDEENKIGQYGRLTFNKVIRPCMKKTIYENEGFREIKGYE
YQLYVYASDKLFRADISEDYKTRGRKLLRFNGPVPPP

6) MON in SBSP

MON produced by the production strain is 96 amino acids. A FASTA sequence was generated using the EMBOSS Seqret conversion tool (https://www.ebi.ac.uk/Tools/sfc/emboss_seqret/).

FASTA sequence:
>EMBOSS_001
GEWEIIDIGPFTQNLGKFAVDEENKIGQYGRLTFNKVIRPCMKKTIYEENGFREIKGYEY
QLYVYASDKLFRADISEDYKTRGRKLLRFNGPVPPP

Appendix D. Bioinformatics Analysis - Monellin

Codex Alimentarius Commission (2003) developed guidelines for the evaluation of the potential allergenicity of novel proteins, recommending bioinformatics search using a FASTA or a BLASTP algorithm, and suggesting that matches of at least 35% identity over segments of at least 80 amino acids may indicate the possibility of cross-reactivity.

Similarly, evaluation of the potential for novel proteins to elicit celiac disease or activate MHC class II restricted T cells in human subjects with celiac disease can be performed using bioinformatics comparisons with known celiac inducing peptides. It has been proposed that identity matches of less than 45% over at least one-half of the FASTA aligned CD protein and those with an E score smaller than 1×10^{-16} using this database are unlikely to present a risk of inducing celiac disease (AllergenOnline, 2023).

Bioinformatics tools were applied using the FASTA sequences for MON and the monellin entries from *Dioscoreophyllum cumminsii* listed in the NCBI protein database, including monellin chain A and monellin chain B, as presented in Appendix C.

AllergenOnline FASTA Sequence Search Comparison

Methods

The AllergenOnline allergen database version used for these FASTA sequence comparison searches was No. 21 (released on February 14, 2021) with 2233 peer reviewed amino acid sequences entries categorized into 913 taxonomic-protein groups of unique proven or putative allergens (food, airway, venom/salivary and contact) from 430 species. Sequence comparison searches were performed using the FASTA bioinformatics tool to elucidate the degree of similarity in identity between a test FASTA amino acid sequence with amino acid sequences for putative allergens in the AllergenOnline database. Additionally, a specific search was performed using a sliding 80mer window comparison, comparing every possible 80 amino acid segment of the query protein to determine whether there were matches of greater than 35% identity. Matches with E-values larger than 1e-7 are not likely to identify relevant matches when seeking to identify proteins that may share immunologic or allergic cross-reactivity, while matches with E-values smaller than 1e-30 are much more likely to be cross-reactive in at least some allergic individuals. A search was also performed against the AllergenOnline Celiac Database of proteins version 3, April 2022. Query sequences were screened against the set of 7 full length CD eliciting proteins using FASTA version 35.04.

Results

The AllergenOnline allergen database sequence analysis tools returned results showing no matches greater than 35% identity between putative allergens in the AllergenOnline allergen database and MON or the monellin entries from *Dioscoreophyllum cumminsii* listed in the NCBI protein database, including monellin chain A and monellin chain B. The AllergenOnline celiac

database sequence analysis tools returned results showing no matches greater than 35% identity between known celiac peptides in the AllergenOnline celiac database and MON or the monellin entries from *Dioscoreophyllum cumminsii* listed in the NCBI protein database, including monellin chain A and monellin chain B.

COMPARE Allergen Database Sequence Search Results

Methods

The COMprehensive Protein Allergen REsource (COMPARE) 2023 database (https://comparefasta.comparedatabase.org/, released on 26 January 2023) is the seventh iteration. The comparative sequence software COMPASS (COMPare Analysis of Sequences with Software) was applied to assess the degree of shared sequence similarity between an amino acid sequence of interest ("query sequence") and allergen sequences within the COMPARE 2023 database, incorporating the open source FASTA software package (FASTA v36). Three different types of sequence comparison searches were performed: Full Length Sequence search; 80-mer sliding window FASTA search, and 8-mer FASTA search.

Results

The COMPASS tool returned results showing no matches greater than 35% identity between putative allergens in the COMPASS allergen database and MON or the monellin entries from *Dioscoreophyllum cumminsii* listed in the NCBI protein database, including monellin chain A and monellin chain B.

AllerCatPro 2.0 Allergenicity Tool Search Results

Methods

The AllerCatPro 2.0 tool uses comparative sequence software to compare protein sequences in FASTA format and generates an AllerCatPro 2.0 output table with the result for strong, weak, or no evidence for allergenicity per protein based on corresponding workflow decisions and, in case of a hit, the possibility to view the most similar allergens with detailed results for cross-reactivity, protein information (UniProt/NCBI), functionality (Pfam, InterPro, SUPFAM), as well as clinical relevance of IgE prevalence (Allergome) and allergen information, including the most similar 3D surface epitope. AllerCatPro 2.0 also identifies all similar allergens that have significant sequence similarity to the query protein and refers to the number with the link in potential cross-reactivity of the output table, as well as all possible similar autoimmune allergens displayed in the link and all possible similar low allergens in the link of the output table.

Results

None of the 6 proteins entered into the AllerCatPro 2.0 tool were found to have significant hits with an E-value threshold of 0.001 to known allergens (non-autoimmune allergens), and there was no similarity to autoimmune allergens or proteins of low allergenicity. Therefore. the AllerCatPro 2.0 tool has returned a result of 'no evidence' for the allergenicity potential for MON and for the monellin entries from *Dioscoreophyllum cumminsii* listed in the NCBI protein database, including monellin chain A and monellin chain B.

Appendix E. Bioinformatics Analysis – Production Strain Host Cell Proteins >1% of Total Protein in SBSP

Bioinformatics tools were applied using the FASTA sequences for the 5 production strain host cell proteins present with an average level >1% in SBSP (Appendix A) as per the searches performed in Appendix D.

AllergenOnline FASTA Sequence Search Comparison Results

The bioinformatics results indicate some likelihood of cross-reactivity in individuals allergic to olive pollen for the Superoxide dismutase [Cu-Zn] host cell protein.

There are no indications from these AllergenOnline celiac protein comparisons that would suggest relevant cross-reactivity for SCP domain-containing protein from the production strain and known allergens or gluten-like peptides. There are some unequivocal indications from the AllergenOnline putative allergen protein comparisons of potential for allergen cross-reactivity with pathogenesis related protein PR-1 from pollen sources.

There are no indications from these AllergenOnline allergen and celiac protein comparisons that would suggest relevant cross-reactivity for chitin deacetylase from the production strain and known allergens or gluten-like peptides.

There is a potential similarity between a protein from the dust mite *Lepidoglyphus destructor*, however this is at an E value confidence of >1e-1 and as such is not expected to be relevant. There are no indications from these AllergenOnline allergen and celiac protein comparisons that would suggest relevant cross-reactivity for chitin deacetylase from the production strain and known allergens or gluten-like peptides.

There are no indications from these AllergenOnline allergen and celiac protein comparisons that would suggest relevant cross-reactivity for ACB domain containing protein from the production strain and known allergens or gluten-like peptides.

AllerCatPro 2.0 Allergenicity Tool Search Results

Of the 5 host cell proteins present at a >1% total protein in SBSP, one was proposed by the AllerCatPro 2.0 modelling tool to have strong evidence of allergenicity potential (Superoxide dismutase [Cu-Zn]), and one was proposed to have weak evidence of allergenicity potential (SCP domain-containing protein). The remaining 3 proteins were found to have no significant hits (E-value threshold 0.001) against known protein allergens, auto-allergens or gluten-like Q-repeats.

Summary of Allergenicity Potential

There are no indications of any cross-reactivity with known celiac inducing peptides for any of the 5 host cell proteins analyzed using the AllergenOnline Celiac Database tools or the

AllerCatPro 2.0 tools. There were no indications of relevant allergen cross-reactivity for Chitin deacetylase, Fusion protein, or ACB domain-containing protein.

The specific sequence for C4R8X7 superoxide dismutase in *K. phaffii* was found to have 31 hits of predicted cross-reactivity >35% homology by AllerCatPro 2.0. These hits were proteins from four species: Neo fi Cu/Zn-SOD from *Neosartorya fischeri* (an environmental mold), Amb t 13 from *Ambrosia trifida* (giant ragweed pollen), Sola I SOD from *Solanum lycopersicum* (tomato), and Ole e 5 from *Olea europaea* (olive tree pollen). The AllergenOnline bioinformatics results also indicate some likelihood of cross-reactivity in individuals allergic to olive pollen for a Superoxide dismutase [Cu-Zn] host cell protein. Superoxide dismutases (SOD) are ubiquitous metalloproteins found across multiple taxonomic groups in both prokaryotes and eukaryotes, and are part of the Cu, Zn SOD-like protein superfamily. The presence of Cu, Zn SOD is ubiquitous across a wide range of organisms to which consumers are commonly and frequently exposed. There is a lack of published evidence for prevalence of such broad cross-reactivity. There are no records documenting IgE prevalence of reactivity for these potentially cross-reacting proteins except for Ole e 5, which has recorded evidence of 66 subjects upon skin prick testing to olive tree pollen proteins, including Ole e 5.[14] Therefore, there may be potential for cross-reactions upon consumption for individuals sensitized via the inhalation route to airborne Ole e 5 from *Olea europaea*.

The specific sequence for C4R3H3 SCP domain-containing protein in *K. phaffii* was found to have 53 hits of predicted cross-reactivity, but only 13 of these were with >35% homology. Of these 13, only two had reports of IgE prevalence: Cuc m 3 from *Cucumis melo* (cantaloupe or muskmelon) and Art v 2 from *Artemis vulgaris* (Mugwort / ragweed). The AllergenOnline bioinformatics results also indicated cross-reactivity with pathogenesis related protein PR-1 from pollen sources. SCP domain-containing protein is part of CRISP family and the pathogen-related protein – 1 (PR-1-like) protein superfamily. These are found in a wide range of organisms. The presence of the Pr-1-like superfamily is ubiquitous across a wide range of organisms to which consumers are commonly and frequently exposed. Again, there is a lack of published evidence for prevalence of such broad cross-reactivity. The predicted most similar allergen is Cuc m 3 from *Cucumis melo* (cantaloupe or muskmelon), with reported positive IgE skin prick tests in 2 of 17 and Cuc m 3 IgE binding to sera from 12 of 17 patients allergic to melon (Asensio *et al.*, 2004). Skin prick tests, performed in 19 patients allergic to mugwort and 10 control patients, showed an Art v 2 sensitization prevalence of 58% (Arilla *et al*., 2007). Therefore, there may be potential for cross-reactions upon consumption for individuals sensitized to Cuc m 3 from *Cucumis melo* (cantaloupe or muskmelon) and sensitized via the inhalation route to airborne Art v 2 from *Artemis vulgaris*.

[14] https://www.allergome.org/script/dettaglio.php?id_molecule=493

DEPARTMENT OF HEALTH AND HUMAN SERVICES
Food and Drug Administration

GENERALLY RECOGNIZED AS SAFE *(GRAS)* NOTICE (Subpart E of Part 170)

Form Approved: OMB No. 0910-0342; Expiration Date: 07/31/2022
(See last page for OMB Statement)

FDA USE ONLY

GRN NUMBER	DATE OF RECEIPT
001183	Feb 28, 2024
ESTIMATED DAILY INTAKE	INTENDED USE FOR INTERNET
NAME FOR INTERNET	
KEYWORDS	

Transmit completed form and attachments electronically via the Electronic Submission Gateway *(see Instructions)*; OR Transmit completed form and attachments in paper format or on physical media to: Office of Food Additive Safety *(HFS-200)*, Center for Food Safety and Applied Nutrition, Food and Drug Administration,5001 Campus Drive, College Park, MD 20740-3835.

SECTION A – INTRODUCTORY INFORMATION ABOUT THE SUBMISSION

1. Type of Submission *(Check one)*

☒ New ☐ Amendment to GRN No. _____ ☐ Supplement to GRN No. _____

2. ☒ All electronic files included in this submission have been checked and found to be virus free. *(Check box to verify)*

3 Most recent presubmission meeting *(if any)* with FDA on the subject substance *(yyyy/mm/dd)*: 2023/10/30

4 For Amendments or Supplements: Is your amendment or supplement submitted in response to a communication from FDA? *(Check one)* ☐ Yes ☐ No If yes, enter the date of communication *(yyyy/mm/dd)*: _____

SECTION B – INFORMATION ABOUT THE NOTIFIER

1a. Notifier	Name of Contact Person: Jason Ryder	Position or Title: Chief Technology Officer		
	Organization *(if applicable)*: Oobli, Inc.			
	Mailing Address *(number and street)*: 202 Cousteau Place, Suite 210			
City: Davis	State or Province: California	Zip Code/Postal Code: 95618	Country: United States of America	
Telephone Number: 510-684-5610	Fax Number:	E-Mail Address: jason.ryder@oobli.com		
1b. Agent or Attorney *(if applicable)*	Name of Contact Person: Mary Murphy	Position or Title: Principal Scientist		
	Organization *(if applicable)*: Exponent, Inc.			
	Mailing Address *(number and street)*: 1150 Connecticut Avenue, NW Suite 1100			
City: Washington	State or Province: District of Columbia	Zip Code/Postal Code: 20036	Country: United States of America	
Telephone Number: 202-772-4953	Fax Number:	E-Mail Address: mmurphy@exponent.com		

SECTION C – GENERAL ADMINISTRATIVE INFORMATION

1. Name of notified substance, using an appropriately descriptive term

serendipity berry sweet protein (SBSP) containing monellin (MON) expressed in Komagataella phaffii

2. Submission Format: *(Check appropriate box(es))*

☒ Electronic Submission Gateway

☐ Electronic files on physical media

☐ Paper

If applicable give number and type of physical media

_____

3. For paper submissions only:

Number of volumes _____

Total number of pages _____

4. Does this submission incorporate any information in CFSAN's files? *(Check one)*

☐ Yes *(Proceed to Item 5)* ☒ No *(Proceed to Item 6)*

5. The submission incorporates information from a previous submission to FDA as indicated below *(Check all that apply)*

☐ a) GRAS Notice No. GRN _____

☐ b) GRAS Affirmation Petition No. GRP _____

☐ c) Food Additive Petition No. FAP _____

☐ d) Food Master File No. FMF _____

☐ e) Other or Additional *(describe or enter information as above)* _____

6. Statutory basis for conclusions of GRAS status *(Check one)*

☒ Scientific procedures *(21 CFR 170.30(a) and (b))* ☐ Experience based on common use in food *(21 CFR 170.30(a) and (c))*

7. Does the submission (including information that you are incorporating) contain information that you view as trade secret or as confidential commercial or financial information? (see 21 CFR 170.225(c)(8) and 170.250(d) and (e))

☐ Yes *(Proceed to Item 8*

☒ No *(Proceed to Section D)*

8. Have you designated information in your submission that you view as trade secret or as confidential commercial or financial information *(Check all that apply)*

☐ Yes, information is designated at the place where it occurs in the submission

☐ No

9. Have you attached a redacted copy of some or all of the submission? *(Check one)*

☐ Yes, a redacted copy of the complete submission

☐ Yes, a redacted copy of part(s) of the submission

☐ No

SECTION D – INTENDED USE

1. Describe the intended conditions of use of the notified substance, including the foods in which the substance will be used, the levels of use in such foods, and the purposes for which the substance will be used, including, when appropriate, a description of a subpopulation expected to consume the notified substance.

The intended use of SBSP containing MON expressed in *K. phaffii* is use as a general-purpose sweetener in foods at levels determined by current good manufacturing practice (cGMP), excluding use in infant formula and meat and poultry products.

2. Does the intended use of the notified substance include any use in product(s) subject to regulation by the Food Safety and Inspection Service (FSIS) of the U.S. Department of Agriculture?

(Check one)

☐ Yes ☒ No

3. If your submission contains trade secrets, do you authorize FDA to provide this information to the Food Safety and Inspection Service of the U.S. Department of Agriculture?

(Check one)

☐ Yes ☐ No , you ask us to exclude trade secrets from the information FDA will send to FSIS.

SECTION E – PARTS 2 -7 OF YOUR GRAS NOTICE

(check list to help ensure your submission is complete – PART 1 is addressed in other sections of this form)

☒ PART 2 of a GRAS notice: Identity, method of manufacture, specifications, and physical or technical effect (170.230).

☒ PART 3 of a GRAS notice: Dietary exposure (170.235).

☒ PART 4 of a GRAS notice: Self-limiting levels of use (170.240).

☒ PART 5 of a GRAS notice: Experience based on common use in foods before 1958 (170.245).

☒ PART 6 of a GRAS notice: Narrative (170.250).

☒ PART 7 of a GRAS notice: List of supporting data and information in your GRAS notice (170.255)

Other Information

Did you include any other information that you want FDA to consider in evaluating your GRAS notice?

☐ Yes ☒ No

Did you include this other information in the list of attachments?

☐ Yes ☐ No

SECTION F – SIGNATURE AND CERTIFICATION STATEMENTS

1. The undersigned is informing FDA that Jason Ryder
(name of notifier)

has concluded that the intended use*(s)* of serendipity berry sweet protein (SBSP) containing monellin (MON) expressed in Komagataella
(name of notified substance)

described on this form, as discussed in the attached notice, is (are) not subject to the premarket approval requirements of the Federal Food, Drug, and Cosmetic Act based on your conclusion that the substance is generally recognized as safe recognized as safe under the conditions of its intended use in accordance with § 170.30.

2. Oobli, Inc. agrees to make the data and information that are the basis for the conclusion of GRAS status available to FDA if FDA asks to see them; agrees to allow FDA to review and copy these data and information during customary business hours at the following location if FDA asks to do so; agrees to send these data and information to FDA if FDA asks to do so.
(name of notifier)

Exponent, Inc., 1150 Connecticut Avenue, NW, Suite 1100, Washington, DC 20036
(address of notifier or other location)

The notifying party certifies that this GRAS notice is a complete, representative, and balanced submission that includes unfavorable, as well as favorable information, pertinent to the evaluation of the safety and GRAS status of the use of the substance.The notifying party certifies that the information provided herein is accurate and complete to the best or his/her knowledge. Any knowing and willful misinterpretation is subject to criminal penalty pursuant to 18 U.S.C. 1001.

3. Signature of Responsible Official, Agent, or Attorney	Printed Name and Title	Date (mm/dd/yyyy)
DocuSigned by: [redacted] D0AD848849D24FB...	Jason Ryder Chief Technology officer	2/28/2024

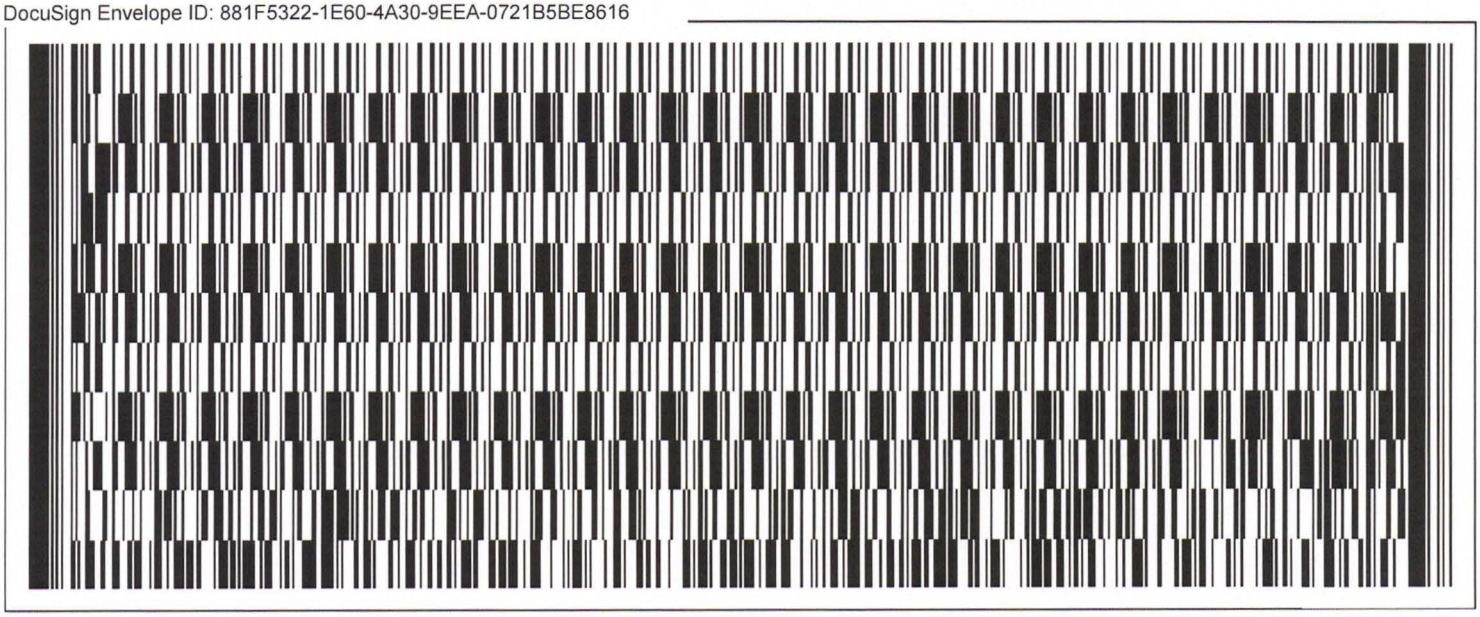
SECTION G – LIST OF ATTACHMENTS

List your attached files or documents containing your submission, forms, amendments or supplements, and other pertinent information. Clearly identify the attachment with appropriate descriptive file names (or titles for paper documents), preferably as suggested in the guidance associated with this form. Number your attachments consecutively. When submitting paper documents, enter the inclusive page numbers of each portion of the document below.

Attachment Number	Attachment Name	Folder Location (select from menu) (Page Number(s) for paper Copy Only)
1	Insert Clear GRASNotice_SBSP	
	Insert Clear	
	Insert Clear	
	Insert Clear	
	Insert Clear	
	Insert Clear	
	Insert Clear	
	Insert Clear	
	Insert Clear	

Add Continuation Page

OMB Statement: Public reporting burden for this collection of information is estimated to average 170 hours per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to: Department of Health and Human Services, Food and Drug Administration, Office of Chief Information Officer, PRAStaff@fda.hhs.gov. (Please do NOT return the form to this address). An agency may not conduct or sponsor, and a person is not required to respond to, a collection of information unless it displays a currently valid OMB control number.



Chemistry Questions

1. Please provide the CAS number for the native monellin.

Response: The Chemical Abstracts Service (CAS) number for monellin, *Dioscoreophyllum cumminsii,* is CAS No. 101794-80-3. This was verified in CAS SciFinder, a webtool managed by CAS. Another CAS number is reported for monellin, 9062-83-3; however, this CAS number was deleted from CAS in 1977 (Email from CAS to Exponent, Sept 13, 2024).

2. Please clarify if the single-chain monellin is secreted into the fermentation media.

Response: Single-chain monellin is secreted into the fermentation media.

3. Please provide the color of the monellin preparation final product.

Response: The color of the monellin preparation final product is off-white.

4. Please state whether any of the raw materials used in the fermentation are or are derived from major food allergens. If any of the raw materials used are, or are derived from major food allergens, please discuss why these materials do not pose a safety concern.

Response: The raw materials used in the fermentation are not major food allergens nor are they derived from major food allergens.

5. Please provide a statement that all raw materials and processing aids used in the manufacturing process are approved for their respective uses via a regulation in Part 21 of the U.S. Code of Federal Regulations, are the subject of an effective food contact notification, or are GRAS for their intended use.

Response: All raw materials and processing aids used in the manufacturing process are approved for their respective uses via a regulation in Part 21 of the U.S. Code of Federal Regulations or are GRAS for their intended use.

Microbiology Questions

1. The specification listed for the aerobic plate count (APC) is 10,000 CFU/g and the batch analysis results were 80, 420, and 550 CFU/g for the three analyzed batches.

    a. First, we request that you consider reducing the specification to align more closely with the results from the batch analyses which are 2-3 orders of magnitude lower than the current specification.

Response: The specification of no more than 10,000 cfu/g for APC is consistent with the specification for this parameter as identified in numerous GRAS notices for ingredients produced via fermentation of an engineered strain of *K. phaffii,* e.g., Brazzein produced by *Komagataella phaffii* expressing a gene encoding for brazzein from *Pentadiplandra brazzeana* (GRN 1142), non-animal egg-white protein (GRN 1104), β-lactoglobulin produced by *Komagataella phaffii* strain yRMK-66 (GRN 1056), myoglobin produced via fermentation of an engineered strain of *K. phaffii* (GRN 1001), and LegH preparation produced via fermentation of an engineered strain of *K. phaffii* (GRN 737). We recognize that SBSP batch data demonstrate that the APC count is well below the specification, though analytical data on the other ingredients mentioned above also indicate that APC levels are well below the specification.

    b. Second, we request a brief discussion of the positive APC values measured in your batch analyses. For an ingredient made under controlled fermentation, we do not necessarily expect the APC test to measure positives consistently. Please discuss the potential sources of the organisms identified by this test and how this is not expected to impact safety of the final ingredient.

Response: The production of SBSP was detailed in the GRAS notice. As indicated in the text and accompanying figure (Figure 4) in GRN 1183, production of the SBSP ingredient begins in a closed fermentation system. The downstream purification process and packaging stage are not performed under aseptic conditions. Following the fermentation, the supernatant (which contains the SBSP substance) undergoes filtration. While this step is not part of a closed system, the microfiltration step would be expected to exclude all pro- and eukaryotic microorganisms from the process stream. Subsequently, the ultrafiltration step is a sanitary but not aseptic system, therefore some exposure to microorganisms is possible. The SBSP substance is spray dried and packaged in the final step. Environmental exposure to microorganisms detected with the APC could occur post spray drying, including filling and packaging operations.

The SBSP ingredient is produced under current Good Manufacturing Practices (cGMP) and Hazard Analysis Critical Control Points (HACCP) in a facility meeting FSSC 22000 requirements for food safety management systems. Variation across batches in the microbial load of finished products is not uncommon. The positive APC values are all well within the established specification for the ingredient and these levels are not expected to impact safety of the ingredient. We note that the ingredient is currently tested for a range of microorganisms, including possible pathogens and indicators as part of product specifications, and these data consistently demonstrate that the ingredient meets specifications established to support a safe and suitable food substance.

2. The notifier indicates that the consumption of non-single-chain-monellin proteins (i.e., copurified *K. phaffii* proteins) is similar to that of previous GRNs using this production organism. Please clarify if the ratios of the different proteins are expected to be similar to the referenced GRNs or if the current manufacturing process is expected to concentrate certain copurifying *K. phaffii* proteins.

Response: Appendix A of GRN 1183 summarizes the average concentration of each of the 27 proteins identified in batch data of the SBSP product. Excluding monellin, which accounts for on average 84% of the total proteins in the product, the average concentration of all other proteins does not exceed 2.4% for any single protein, and only five proteins have an average concentration exceeding 1%. The remaining 21 proteins were present at an average of concentration of <1%. These data indicate that the SBSP production process does not substantially concentrate any of the copurified *K. phaffii* proteins.

Table 1 below presents information provided in Appendix A of GRN 1183 along with the reported average content of host cell proteins in another GRAS ingredient produced from an engineered strain of *K. phaffii*, namely Brazzein produced by *K. phaffii* expressing a gene encoding for brazzein from *Pentadiplandra brazzeana* (GRN 1142). FDA completed review of GRN 1142 and sent the notifier a letter of no questions since GRN 1183 was filed, thus the substance that is the subject of GRN 1142 was not discussed in GRN 1183. The data in Table 1 below show that another GRAS ingredient produced from an engineered strain of *K. phaffii* contains some of the same copurifying *K. phaffii* proteins (the substance that is the subject of GRN 1142 also contains numerous copurifying proteins not listed in Table 1 below). The relative ratios of specific copurifying *K. phaffii* proteins are not identical in the ingredients (as indicated by differences in ranked contributions), though concentrations of all copurifying proteins are small relative to the active protein.

As reviewed in Table 13 of GRN 1183, other GRAS ingredients produced from an engineered strain of *K. phaffii* contain copurifying proteins. The concentrations of specific

copurifying proteins were not quantified in GRNs 1104 and 967. In GRNs 737 and 1056, concentrations of copurifying proteins were quantified; however, the unique identifiers for the co-purifying proteins are not Uniprot Protein ID identifiers, so direct comparison cannot be readily completed.

Collectively, the available information demonstrates that other GRAS ingredients produced from an engineered strain of *K. phaffii* contain copurifying proteins. There is some overlap among the proteins that copurify in the production of SBSP and proteins that copurify in the production of other ingredients, though concentrations of all copurified proteins are relatively low and the data indicate that relative ratios of the copurify proteins differ based on ranked contributions.

Microbiology Question, Table 1: Concentration of copurifying proteins in GRN 1183 and another GRAS ingredient produced from an engineered strain of *K. phaffii*

UNIPROT ID	GRN 1183: SBSP	GRN 1142: OFSP[a]	Protein Name
-	84.48 [Monellin-SC]	46.78 [Brazzein]	Active Protein
C4R8X7	2.38	0.61	Superoxide dismutaseOS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0786 PE=3 SV=1
C4R3H3	2.37	0.68	SCP domain-containing protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0076 PE=4 SV=1
C4QW42	2.08	0.83	Chitin deacetylase, together with Cda1p involved in the biosynthesis ascospore wall component OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-1_0103 PE=4 SV=1
C4R0U2	1.99	6.23	Fusion protein, identical to Rpl40Bp OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0486 PE=3 SV=1
C4QY91	1.37	0.19	ACB domain-containing protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-4_0674 PE=4 SV=1
C4R0P1	0.78	0.16	Glyceraldehyde-3-phosphate dehydrogenase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0437 PE=3 SV=1
C4R9F6	0.45	0.47	Protein TOS1 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_c034_0034 PE=3 SV=1
C4R2G3	0.45	0.56	FK506-binding protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-2_0476 PE=3 SV=1
C4R245	0.42	3.15	Transaldolase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-2_0337 PE=3 SV=1

UNIPROT ID	GRN 1183: SBSP	GRN 1142: OFSP[a]	Protein Name
C4R6V3	0.35	0.39	Nuclear protein required for transcription of MXR1 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0102 PE=4 SV=1
C4QW56	0.35	0.39	Uncharacterized protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-1_0118 PE=4 SV=1
C4R430	0.30	0.25	6-Phosphogluconate dehydrogenase, decarboxylating OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0277 PE=3 SV=1
C4R3H8	0.29	0.16	Phosphopyruvate hydratase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0082 PE=3 SV=1
C4R262	0.28	0.14	Actin-binding protein of the cortical actin cytoskeleton OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-2_0320 PE=4 SV=1
C4QY12	0.24	0.47	Suppressor protein STM1 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-4_0297 PE=4 SV=1
C4QZU2	0.22	0.39	5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0160 PE=3 SV=1
C4R537	0.20	0.5	Protein that binds to cruciform DNA structures OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0627 PE=4 SV=1
C4R3X8	0.19	0.08	ATPase involved in protein folding and the response to stress OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr3_0230 PE=3 SV=1
C4R1J7	0.15	0.03	Eisosome protein 1 OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0722 PE=3 SV=1
C4QWG6	0.12	0.72	Protein component of the large (60S) ribosomal subunit OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-1_0219 PE=3 SV=1
C4QV10	0.11	0.03	Inorganic diphosphatase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-3_0028 PE=3 SV=1
C4QWH2	0.11	0.14	Vacuolar proteinase B (YscB), a serine protease of the subtilisin family OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-1_0226 PE=3 SV=1
C4R184	0.11	0.19	Uncharacterized protein OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0618 PE=4 SV=1
C4R887	0.07	0.27	ATPase involved in protein folding and nuclear localization signal (NLS)-directed nuclear transport OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr4_0552 PE=3 SV=1
C4QYV1	0.07	0.22	Primary component of eisosomes OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr1-4_0569 PE=4 SV=1

UNIPROT ID	GRN 1183: SBSP	GRN 1142: OFSP[a]	Protein Name
C4R0T4	0.07	0.06	Deoxyuridine 5'-triphosphate nucleotidohydrolase OS=Komagataella phaffii (strain GS115 / ATCC 20864) OX=644223 GN=PAS_chr2-1_0478 PE=3 SV=1

[a] The complete list of proteins copurifying in the production of OFSP is provided in the Appendix of GRN 1142.

Toxicology Question

1. Is the sweetening effect of single-chain monellin transient, similar to polysaccharide-based sweeteners, or are effects on taste receptors sustained for a significant duration? Please note that such activity has been reported in sensory studies of other protein-based sweeteners, such as miraculin.

Response: The sweetening effect of serendipity berry sweet protein containing single-chain monellin (SBSP) was investigated in a time-intensity sensory evaluation. Trained panelists (N=10) were provided SBSP (single-chain monellin purity 19%) and 5% sucrose solutions at equivalent sweetness. The normalized sucrose solution reached a peak sweetness intensity at approximately 5 seconds while the SBSP solution reached a peak sweetness intensity at approximately 6 seconds (Figure 1). The sweetness intensity of both solutions rapidly declined within 20 seconds and progressively decreased to zero thereafter. These findings indicate that the sweetening effect of single-chain monellin is transient and not sustained.

Toxicology Question, Figure 1. Sweet taste time intensity of SPSP vs sucrose