Safety Assessment Considerations for Proteins from Novel Sources

Supratim Choudhuri, PhD

Division of Food Ingredients

Center for Food Safety and Applied Nutrition

US Food and Drug Administration

Conflict of Interest Statement

I declare no conflict of interest

Objectives

- General biological aspects of proteins in food.
- Known and potential adverse effects of proteins in food.
- Principles of safety assessment of new proteins in food.

Proteins

- Structural components (e.g., actin, keratin).
- Intra- and extracellular transporters (e.g., ferritin and transferrin).
- Membrane transporters (uptake and efflux).
- Antibodies.
- Enzymes.
- Messengers (e.g., growth hormone).
- Offense-defense effectors (e.g., plant and animal toxins).

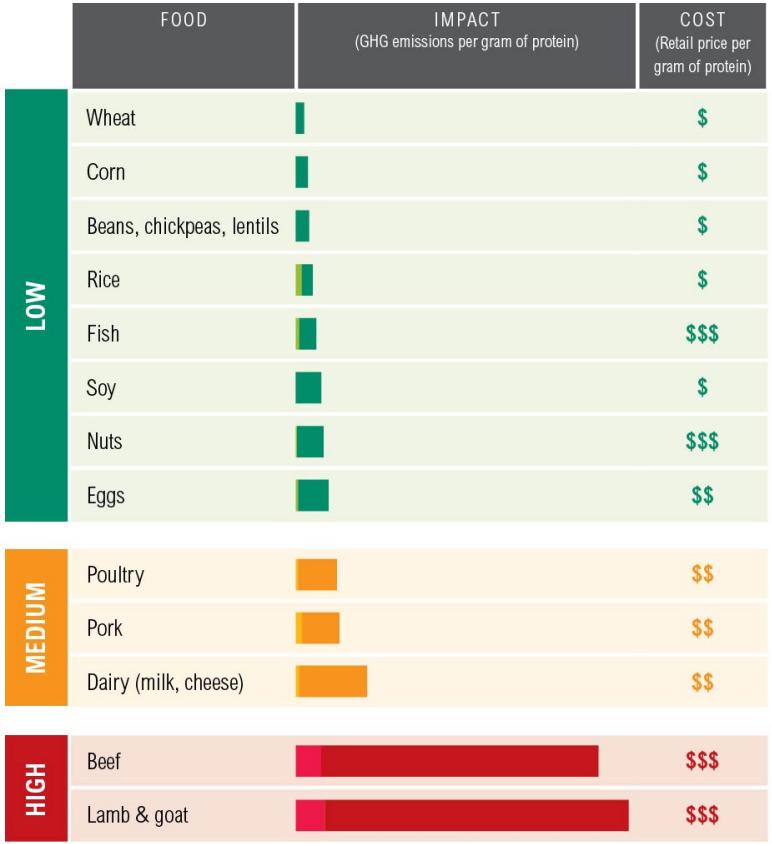
Source of Proteins in Food: Shifting Preference

Dietary proteins are macronutrients.

Increasing consumer preference for plant-based proteins:

(1) increasing health consciousness.

(2) increasing environmental awareness.



WORLD RESOURCES INSTITUTE
PROTEIN SCORECARD
wri.org/shiftingdiets

FDA's Experience: Submissions for Proteins from Novel Sources

High Proposed Daily Intake

- Plant sources (Increasing trend).
- Fungal sources (*Fusarium* sp.).
- Milk protein.

Low Daily Intake

- Bacterial or endogenous proteins genetically engineered into plants (e.g., EPSPS, PAT).
- Enzymes used in cheesemaking (e.g., chymosin).
- Enzymes used as processing aid.

Adverse Effects of Proteins

- Excess protein intake
- Misfolded proteins
- Allergenicity and toxicity

The BIG EIGHT

- Big 8: Milk, Eggs, Peanuts, Tree nuts, Soy, Wheat, Finfish, and Shellfish.
  Source labeling is the primary means of protecting consumers that are allergic to these foods.
- Allergenicity by other foods: Low incidence (e.g., certain fruits, vegetables, and seeds; example–kiwi, celery, sesame, mustard, lentil).
- Incidence of food allergy: 6-8% in children and 1-2% in adults. Often develops during infancy/early childhood.
- About 80% of all food allergies in children → peanuts, milk, or eggs.
  - Childhood food allergies (including soy allergy) are usually outgrown.
  - Allergies to peanuts, tree nuts, and fish are rarely resolved in adulthood.

Characteristics of Allergenic/Toxic Proteins

Some proposed characteristics (not universal):

- High levels in the source (hence high dietary intake).
- Heat stability.
- -S-S- bonds.
- Stability in SGF (particularly in lower pepsin-to-protein ratio).
- Glycosylation.
- Allergenic/toxic structural motifs in some proteins.

Allergenic and Toxic Plant Proteins in Food

A small number of protein families/superfamilies:

- Cupins (e.g., vicilins, legumins).
- Prolamins (e.g., 2S albumins, LTPs, tree nuts, protease and α-amylase inhibitors, cereal prolamins).
- Plant defense system and pathogenesis-related proteins (e.g., Bet V1, Chitinase, Thaumatin-like proteins).
- Seeds contain various naturally occurring compounds toxic to humans and animals, such as the trypsin inhibitors, phytic acid, lectins.

If a protein is from an allergenic (or toxic) source, the protein is assumed to be allergenic (or toxic) unless demonstrated otherwise.

Facts To Remember
Safety Assessment of Total Proteins or Individual Proteins

- Two issues: toxicity and allergenicity (primary)
- Most food allergens are proteins, but most food proteins are not allergenic or toxic to the majority of the population.
- Any source of protein could be allergenic to some individuals in the population.

The protein source and its history of use in food are important considerations in the safety assessment of both total proteins and individual proteins.

Safety Assessment of Individual Proteins

Weight-of-Evidence (WOE) approach for both toxicity and allergenicity (FAO/WHO; CODEX):

- Source of the protein (non-allergenic/non-toxigenic/non-pathogenic; protein family).
- Dietary intake level.
- Biological function (mode of action).
- History of safe use.
- Sequence comparison with known allergens and toxins.
- *In vitro* digestibility in simulated gastric fluid (SGF).
- Oral toxicity studies.
- Additional studies as needed.
- Case-by-case approach.

Safety Assessment of New Protein: Some Examples

Pea protein (Total Protein)

Protein from *Fusarium* sp. (Total Protein)

β-lactoglobulin (Individual Protein)

Soy leghemoglobin (Individual Protein)

ALL GRAS SUBMISSIONS ARE PUBLICLY AVAILABLE

Safety Assessment of Pea Protein: Total Protein

- Consumption of pea protein is a new use, but pea has been historically widely consumed.
- Intake: Up to ~50 g.
- Safety data:
  - ADME studies in humans.
  - Subchronic toxicity studies in rats.
  - Mutagenicity and genotoxicity studies.
  - Human studies (tolerance, satiation, benefit).
  - A discussion on the low allergenicity of pea and labeling to inform consumers.

Safety Assessment of *Fusarium venenatum* Protein: Total Protein

Protein from *Fusarium venenatum* (mycoprotein).

Over 40-50 years of history of use by now.

Approved by U.K. in 1985 for use in food; FDA GRAS in 2001.

Intake: Up to 46 g for the general population.

Safety data:

- Compliance with several specifications.
- Several long-term published animal toxicological studies in rats, dogs, baboons.
- Several human tolerance studies (adverse GI symptoms ~1 per 140,000).
- Notifier concluded from a published investigational study that ~5% of the reported adverse effects were true allergies whereas the rest were food intolerances.
- The product is labeled.

Safety Assessment of β-lactoglobulin: Individual Protein

- Expressed in *Trichoderma reesei*.
- High levels of proposed intake (up to 48 g).

Safety assessment approach:

- Safety of *T. reesei* as the host of expressed protein (non-toxigenic, non-allergenic, non-pathogenic, widely used organism for expressing proteins/enzymes used in food).
- History of exposure to β-lactoglobulin through milk.
- Source labeling for people allergic to milk.

Soy Leghemoglobin: Individual Protein

- It is a new and novel protein (not known to have been consumed before).
- It is obtained from a novel tissue source (root nodules).
- Expressed in *Pichia pastoris*.
- Soybean is a known allergenic food (one of the big eight).

Safety Assessment of Soy Leghemoglobin (Toxicology)

FACTS to remember:

- Soybean contains several naturally occurring compounds toxic to humans and animals, such as the trypsin inhibitors, phytic acid, lectins, soybean toxin (SBTX).
- Trypsin inhibitor is a concern in uncooked soy (cooking destroys it).
- The toxins and antinutrients are all found in seeds.
- Soy leghemoglobin is in root nodules, not in seeds.
- Many other foods eaten regularly also have these toxins/anti-nutrients (e.g., corn, beans, nuts).

Safety Assessment of Soy Leghemoglobin (Toxicology…contd.)

- 28-day feeding study in rats: no evidence of systemic adverse effects.
- Ames assay and ChromAb assay: Non-mutagenic and non-clastogenic.
- Toxin Database search and BLASTp did not return any positive hits.
- *P. pastoris* is a non-toxigenic, non-allergenic, non-pathogenic host for protein expression, widely used for expressing proteins/enzymes used in food.

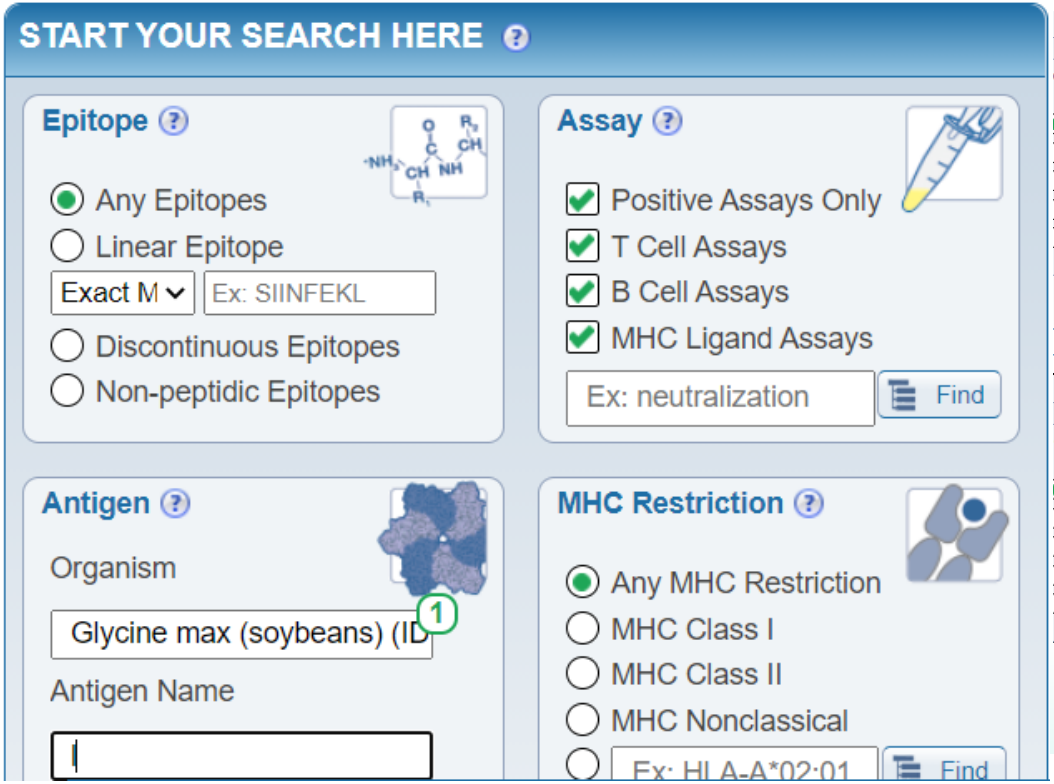
Safety Assessment of Soy Leghemoglobin (Allergenicity)

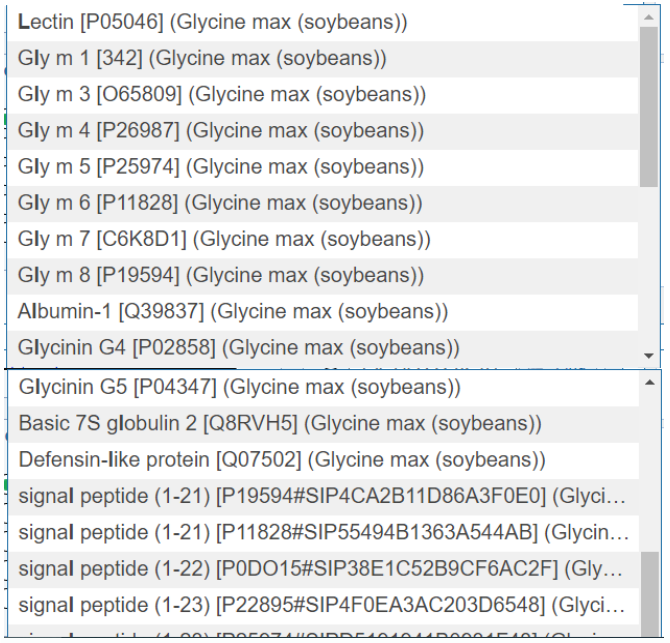
- Source of the protein: The primary soybean allergens are in seeds.
- Soy allergy often starts in infancy, but most children outgrow soy allergy.

Species	Allergen	Biochemical name	MW(SDS-PAGE)	Route of Allergen Exposure	Date Created	Modified Date
Glycine max (Soybean)						
	Gly m 1	Hydrophobic protein from soybean	7 kDa	Food	24-06-2003	2019-08-27
	Gly m 2	Defensin	8 kDa	Airway	25-06-2003	2019-08-27
	Gly m 3	Profilin	14 kDa	Food	25-06-2003	2019-08-27
	Gly m 4	Pathogenesis-related protein, PR-10, Bet v 1 family member	17 kDa	Food	03-11-2006	2019-08-27
	Gly m 5	Beta-conglycinin (vicilin, 7S globulin)	see subunits	Food	19-08-2008	2019-08-27
	Gly m 6	Glycinin (legumin, 11S globulin)	see subunits	Food	19-08-2008	2019-08-27
	Gly m 7	Seed biotinylated protein	76.2 kDa	Food	24-09-2012	2019-08-27
	Gly m 8	2S albumin	28 kDa	Food	19-11-2013	2019-08-27

WHO-IUIS Database

Safety Assessment of Soy Leghemoglobin (Allergenicity…contd.)





IMMUNE EPITOPE DATABASE AND ANALYSIS RESOURCE

Safety Assessment of Soy Leghemoglobin (Allergenicity…contd.)

No	Allergen - Scientific Name	Type	Species - Scientific Name	Species - Common Name	Keywords	Class
1	Gly m 1	foods	*Glycine max*	soybean	P34, 34 kDa maturing seed vacuolar thiol protease; Hydrophobic seed protein; EC	IUIS
2	Gly m 1.0101	foods	*Glycine max*	soybean	Gly m IA	IUIS
3	Gly m 1.0102	foods	*Glycine max*	soybean	Gly m IB	IUIS
4	Gly m 2	foods	*Glycine max*	soybean	Hull allergen; Cytokinin	IUIS
5	Gly m 3	foods	*Glycine max*	soybean	Profilin; GmPRO1	IUIS
6	Gly m 3.0101	foods	*Glycine max*	soybean	N	IUIS
7	Gly m 3.0102	foods	*Glycine max*	soybean	N	IUIS
8	Gly m 4	foods	*Glycine max*	soybean	SAM22, PR-10 protein	IUIS
9	Gly m 5	foods	*Glycine max*	soybean	Beta-conglycinin (vicilin, 7S globulin)	IUIS
10	Gly m 5.0101	foods	*Glycine max*	soybean	Beta-conglycinin (vicilin, 7S globulin); Alpha subunit	IUIS
11	Gly m 5.0201	foods	*Glycine max*	soybean	Beta-conglycinin (vicilin, 7S globulin); Alpha subunit	IUIS
12	Gly m 5.0301	foods	*Glycine max*	soybean	Beta-conglycinin (vicilin, 7S globulin); Beta subunit	IUIS
13	Gly m 5.0302	foods	*Glycine max*	soybean	Beta-conglycinin (vicilin, 7S globulin); Beta subunit	IUIS
14	Gly m 6	foods	*Glycine max*	soybean	Glycinin (legumin, 11S globulin)	IUIS
15	Gly m 6.0101	foods	*Glycine max*	soybean	Glycinin (legumin, 11S globulin), G1 subunit	IUIS
16	Gly m 6.0201	foods	*Glycine max*	soybean	Glycinin (legumin, 11S globulin), G2 subunit	IUIS
17	Gly m 6.0301	foods	*Glycine max*	soybean	Glycinin (legumin, 11S globulin), G3 subunit	IUIS
18	Gly m 6.0401	foods	*Glycine max*	soybean	Glycinin (legumin, 11S globulin), G4 subunit	IUIS
19	Gly m 6.0501	foods	*Glycine max*	soybean	Glycinin (legumin, 11S globulin), G5 subunit	IUIS
20	Gly m Bd28K	foods	*Glycine max*	soybean	MP27/MP32 homologue	non-IUIS
21	Gly m conglycinin	foods	*Glycine max*	soybean	conglycinin, 7S seed storage protein	non-IUIS
22	Gly m glycinin G1	foods	*Glycine max*	soybean	glycinin, 11S seed storage protein	non-IUIS
23	Gly m glycinin G2	foods	*Glycine max*	soybean	glycinin, 11S seed storage protein	non-IUIS
24	Gly m lectin	foods	*Glycine max*	soybean	lectin, agglutinin	non-IUIS
25	Gly m TI	foods	*Glycine max*	soybean	Kunitz trypsin inhibitor	non-IUIS

SearchTerm: *Glycine max*

SDAP

Safety Assessment of Soy Leghemoglobin (Allergenicity…contd.)

COMPARE
Comprehensive Protein Allergen Resource

ABOUT DATABASE PROCESS DEVELOPMENT TIMELINE EVENTS FAQ CONTACT US

Powered by FoodRisk.org

COMPARE Database

COMPARE 2020 DB Release Date: 01/29/2020

SEARCH: Glycine max

Run COMPASS (COMPare Analysis of Sequences with Software)

Article Lookup | 2020 FastA (PDF) | 2020 FastA (Text) | Download Data | Documentati

Search Term: Glycine max

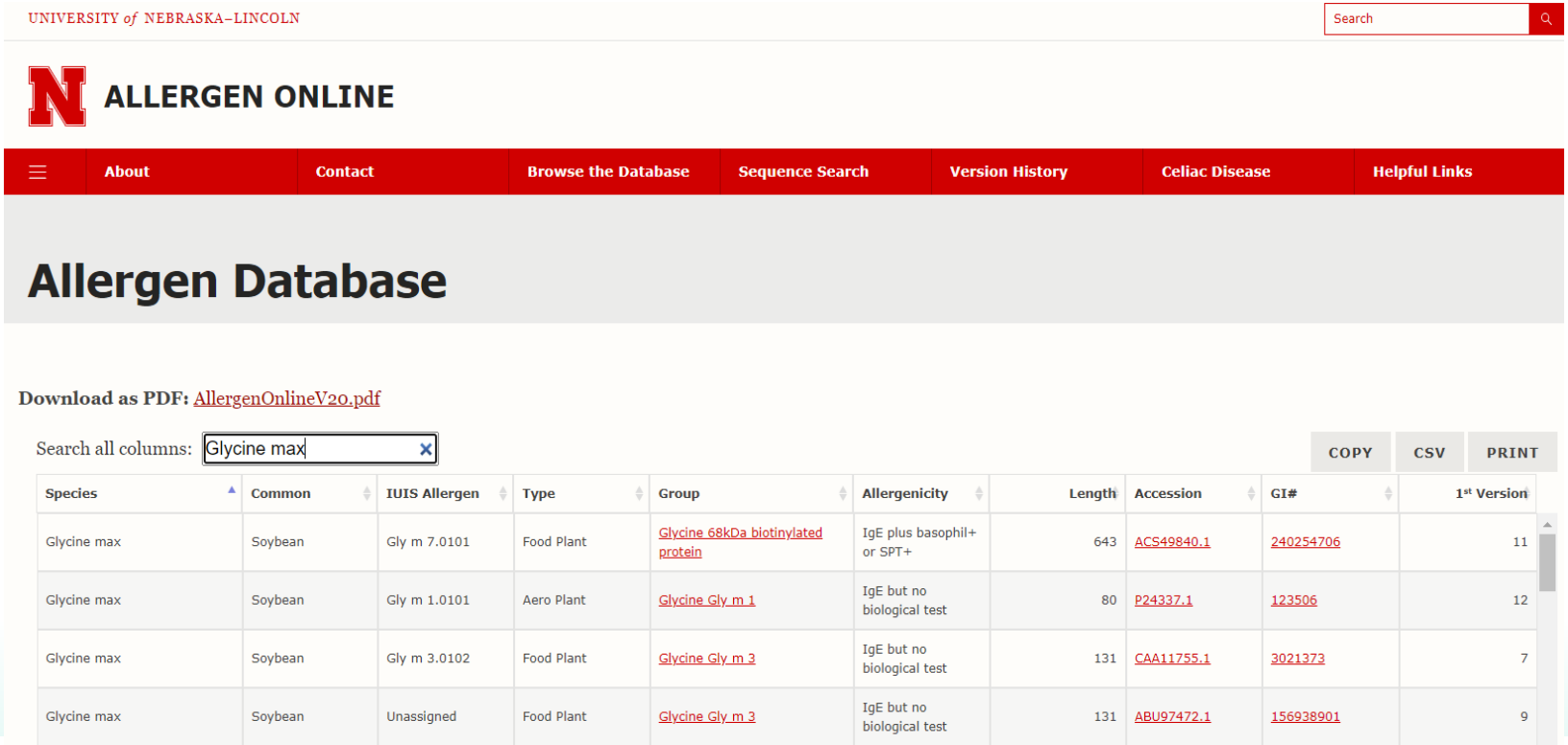
Showing 43 total record(s) foun

Species	Common Name	Description	Accession	Length	Year Adopted	
Glycine max	soybean	Allergen Gly m 7.0101; seed biotinylated protein 68 kDa isoform	ACS49840.1	643	2020H	VIEW
Glycine max	soybean	Allergen Gly m 5; beta-conglycinin (vicilin, 7S globulin) (precursor)	BAA23361.2	416	2020	VIEW
Glycine max	Soybean	glycinin A3B4 subunit [Glycine max]	BAB15802.1	517	2015	VIEW
Glycine max	Soybean	2S albumin precursor [Glycine max]	NP_001238443.1	158	2015	VIEW
Glycine max	Soybean	profilin [Glycine max]	CAA11756.1	131	2015	VIEW
Glycine max	Soybean	beta-conglycinin alpha prime subunit [Glycine max]	BAB64303.1	621	2015	VIEW

SearchTerm: *Glycine max*

43 Entries

Safety Assessment of Soy Leghemoglobin (Allergenicity…contd.)



UNIVERSITY of NEBRASKA–LINCOLN

Search

ALLERGEN ONLINE

About | Contact | Browse the Database | Sequence Search | Version History | Celiac Disease | Helpful Links

Allergen Database

Download as PDF: AllergenOnlineV20.pdf

Search all columns: Glycine max

COPY CSV PRINT

Species	Common	IUIS Allergen	Type	Group	Allergenicity	Length	Accession	GI#	1st Version
Glycine max	Soybean	Gly m 7.0101	Food Plant	Glycine 68kDa biotinylated protein	IgE plus basophil+ or SPT+	643	ACS49840.1	240254706	11
Glycine max	Soybean	Gly m 1.0101	Aero Plant	Glycine Gly m 1	IgE but no biological test	80	P24337.1	123506	12
Glycine max	Soybean	Gly m 3.0102	Food Plant	Glycine Gly m 3	IgE but no biological test	131	CAA11755.1	3021373	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m 3	IgE but no biological test	131	ABU97472.1	156938901	9

SearchTerm: *Glycine max*

43 Entries

Safety Assessment of Soy Leghemoglobin (Allergenicity…contd.)

Search in multiple databases did not reveal any report of allergenicity of soy leghemoglobin.

Safety Assessment of Soy Leghemoglobin (Allergenicity…contd.)

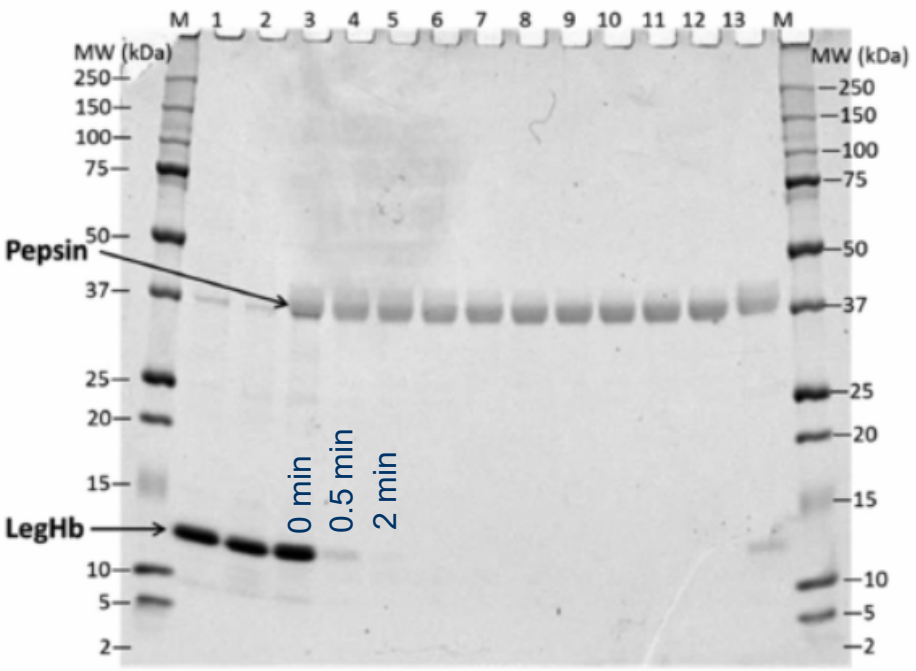
- Dietary intake (90th percentile): 6.7 mg/kg bw/d (= <1% of IOM daily protein intake recommendation).
- Biological function (mode of action): Oxygen-binding like hemoglobin and myoglobin. Similar 3D structure predicted.
- History of safe use: Not reportedly consumed in the US.
- **Digestibility:** 10:1 and 1:1 (U of pepsin:μg protein).
- **Bioinformatic analysis:** Sequence comparison with known toxins and allergens – (I) sequence alignment-based, (II) SVM module-based.

Digestibility

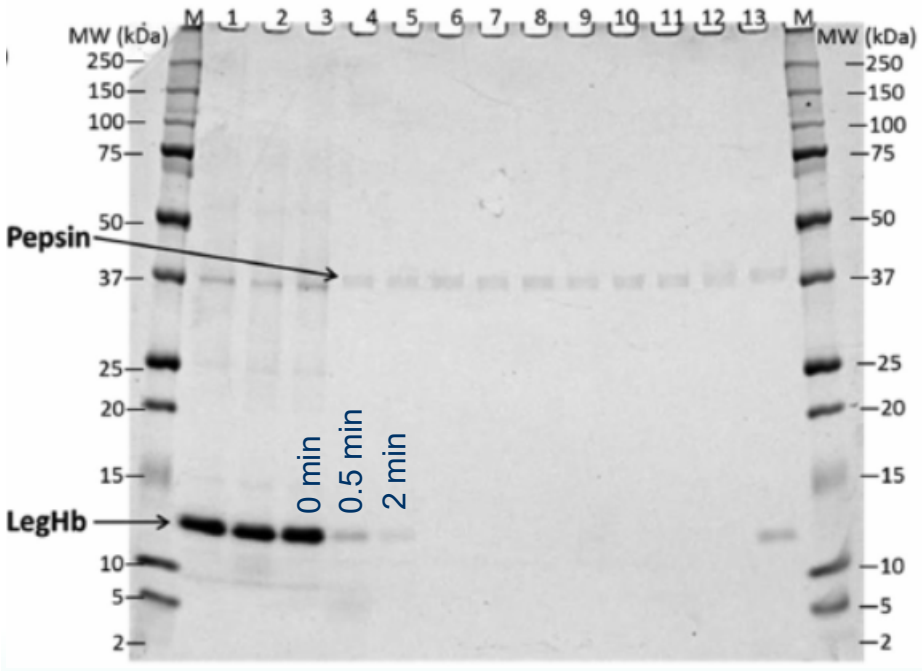
The Utility of Digestibility Studies

- One element of the WOE approach (FAO/WHO; CODEX).
- Allergenic proteins may show greater resistance to SGF *in vitro*, specially under lower pepsin to protein, such as 1U pepsin:1μg protein or 1U pepsin:10 μg protein (different views exist in the literature).
- The notifier performed the assay using the usual 10:1 ratio (10U pepsin: 1μg protein), as well as 1:1 ratio.

In Vitro Digestibility Assay of Soy Leghemoglobin



10U pepsin



1U pepsin

In Silico Digestibility Prediction of Soy Leghemoglobin

mgaftekqea lvsssfeafk anipqysvvf ytsilekapa akdlfsflsn gvdpsnpklt
ghaeklfglv rdsagqlkan gtvvadaalg sihaqkaitd pqfvvvkeal lktikeavgd
kwsdelssaw evaydelaaa ikkaf

Predicted Cutting site (C-terminal from the aa)	Fragment span (aa#)	Fragment length (# of aa)
19 29	20-29	10
35 43	36-43	8
48 65	49-65	17
89 110	90-110	21
111 125	112-125	14
126 136	127-136	10

Name of enzyme	No. of cleavages	Positions of cleavage sites
Pepsin (pH1.3)	33	3 4 10 11 15 16 18 19 29 30 34 35 43 45 46 47 48 59 65 66 68 69 76 77 88 89 103 110 111 125 126 136 137

Leghemoglobins (overwhelming # of hits), various enzymes, σ-factor, Se-binding protein. No hits with known toxins or allergens.

Many of the above fragments can be digested by trypsin (*in silico* prediction).

***In silico* digestibility prediction complements but is NOT a substitute for digestibility experiments**

Bioinformatics (*In Silico* Analysis)

The Utility of Bioinformatics

- Bioinformatic (*in silico*) analysis is an element of the WOE approach (FAO/WHO; CODEX).
- The paradigm: Search for 8 aa identity for potential cross-reactive epitopes; >35% identity over a sliding window of 80 or more aa; (FAO/WHO and CODEX).
- >50% aa identity to known allergens (Aalberse, 2000).
- The predictions are based on sequence alignment-based approach. Other allergenicity prediction algorithms have been developed (e.g., SVM and ACC prediction methods).

Bioinformatic Analysis of Soy Leghemoglobin

- Sequence alignment-based analysis predicted soy leghemoglobin to be a non-allergen.
- SVM-based analysis predicted that soy leghemoglobin could be a potential allergen.
- However, SVM-based analysis also predicts Mb and certain chains of Hb as potential allergens.
- High false positive prediction rate.
- Common function: oxygen-binding.

Why Use Non Sequence-Alignment Based Methods?

- For a novel protein, the potential allergenic and toxicological epitopes may be unique.

Utility of Bioinformatics (Case-by-Case Approach: A Different Example)

Bioinformatic analysis can provide direction to further studies.

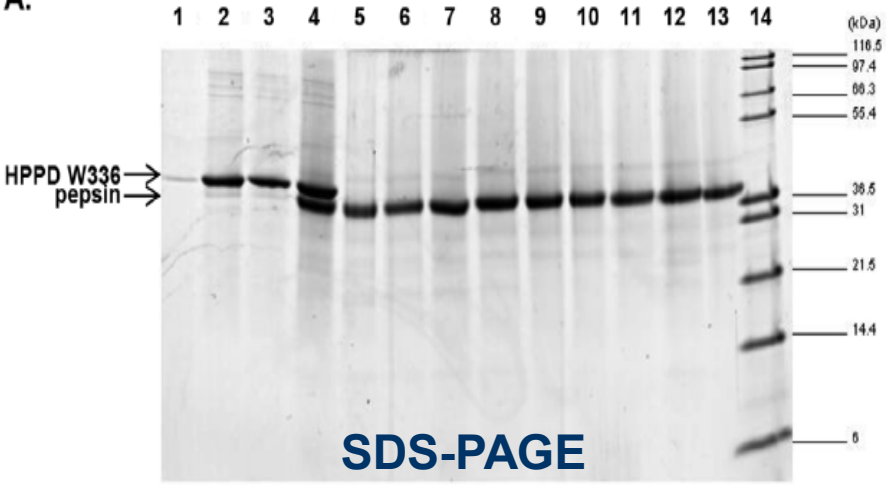
Example: 4-hydroxyphenylpyruvate dioxygenase (HPPD) from *Pseudomonas fluorescens*

Is HPPD a Hemolytic Protein?

- HPPD W336 was engineered into MST-FGØ72-2 GE soybean line to confer resistance to the herbicide isoxaflutole.
- HPPD W336 differs from the native HPPD protein from *Pseudomonas fluorescens* by a glycine to tryptophan (G→W) substitution at position 336.
- Sequence analysis predicted that HPPD proteins could be a potential hemolysin.
- HPPD is digested by pepsin *in vitro*.

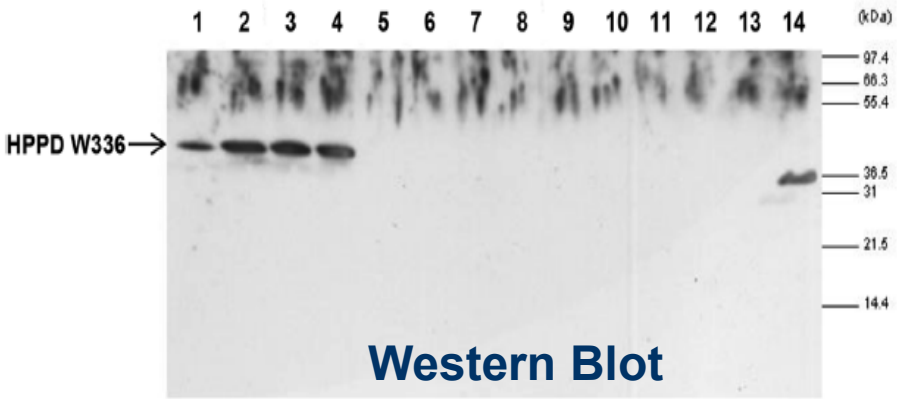
In Vitro Digestibility Assay of HPPD (10:1)

A.



Lanes 1-3: HPPD Controls
Lane 4: 0 min
Lane 5: 0.5 min
Lane 6: 2 min
Lane 7: 5 min ……

B.



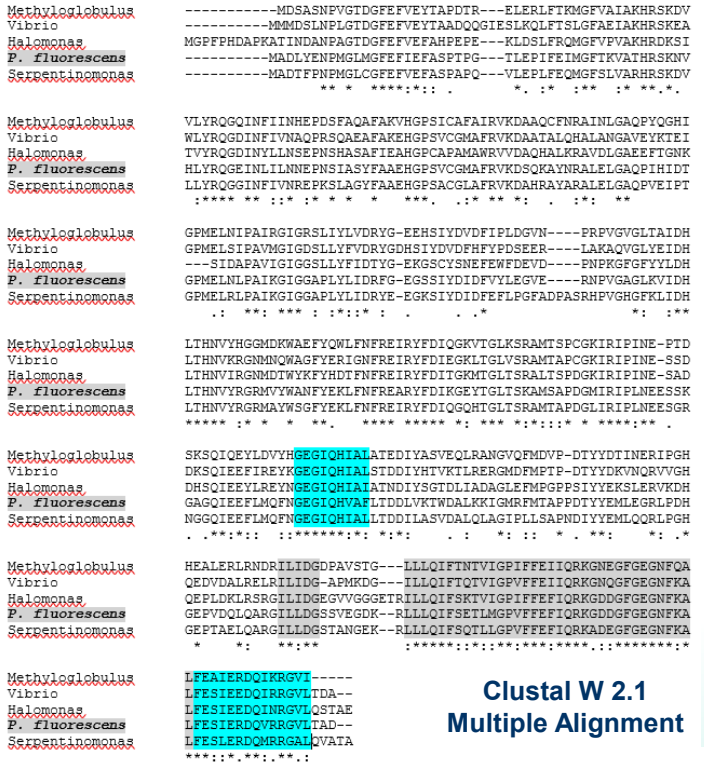
Lanes 1-3: HPPD Controls
Lane 4: 0 min
Lane 5: 0.5 min
Lane 6: 2 min
Lane 7: 5 min ……

Analysis of Amino Acid Identity/Similarity Among Different HPPD Proteins

	M. morosus		*S. raichei*		*Halomonas*		*V. mimicus*	
	% ID	% SIM	% ID	% SIM	% ID	% SIM	% ID	% SIM
Pseudomonas fluorescens	54.5	72.3	72.6	83.5	53.2	68.0	54.7	69.2

M. morosus: *Methyloglobulus morosus*
S. raichei: *Serpentinomonas raichei*
V. mimicus: *Vibrio mimicus VM573*

In Silico Digestibility Prediction of HPPD and Fragment Analysis



Clustal W 2.1 Multiple Alignment

Name of enzyme	No. of cleavages	Positions of cleavage sites
***in silico* analysis of digestibility**		3 4 10 11 13 14 15 16 18 19 26 31 35 36 48 49 56 57 58 59 69 70 82 83 95 97 98 111 112 114 124 126 127 131 141 142 144 145 156 157 162 163 177 178 181
Pepsin (pH1.3)	86	182 184 185 190 199 200 229 230 231 233 234 244 245 248 249 255 256 263 274 275 279 288 289 295 296 309 311 312 315 316 321 322 323 324 332 333 337 338 340 342 355

***In silico* digestibility prediction complements but is NOT a substitute for digestibility experiments**

The *In Silico* Analysis Justifies Additional Studies on Potential Hemolytic Activity of HPPD

- HPPD W336 was incubated with whole blood at 1, 10 and 100 µg/mL concentrations. It did not cause hemolysis (positive control saponin caused maximal hemolysis).
- Mice treated with 2000 mg HPPD/kg bw (acute study) did not show any signs of hemolysis or any clinical signs.
- 28-day feeding study was also conducted; no hemolysis.
- So, both *in vitro* and *in vivo* studies demonstrated the absence of hemolytic activity of HPPD.

The Publication: HPPD is Not a Hemolysin

Regulatory Toxicology and Pharmacology 97 (2018) 170–185

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

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

Characterization and safety evaluation of HPPD W336, a modified 4-hydroxyphenylpyruvate dioxygenase protein, and the impact of its expression on plant metabolism in herbicide-tolerant MST-FGØ72-2 soybean

Rozemarijn Dreesen[a,*], Annabelle Capt[b], Regina Oberdoerfer[c], Isabelle Coats[d], Kenneth, Edward Pallett[a]

[a] Bayer CropScience N.V. – Innovation Center, Tech Lane Ghent Science Park 38, B-9052, Gent, Belgium
[b] Bayer S.A.S., Bayer CropScience, 355 rue Dostoïevski, 06903, Sophia Antipolis, France
[c] Bayer A.G., CropScience Division, Alfred-Nobel-Straße 50, 40789, Monheim, Germany
[d] Bayer CropScience L.P., 2 T.W. Alexander Drive, Research Triangle Park, NC, 27709, USA

Innovation and Food Safety: The Consumer, The Food Industry, and The FDA

Some take-home points:

- Allergenicity of food proteins is not widespread but could be a concern for some sensitive consumers.
- Avoidance is the key.
- Source labeling in packaged food is a useful tool to warn susceptible consumers.
- Responsible innovation practices by food industry is important (e.g., a responsible practice will not introduce a known allergenic protein in corn, a food that cannot be labeled to warn the consumers).
- Continuous engagement of all stakeholders with FDA in the process.

Summary

- Allergenicity of food proteins is not widespread but could be a concern for some sensitive consumers.
- The WOE approach has served us well for the last 20+ years.
- Depending on the protein, further targeted studies may be needed (e.g., HPPD). Hence, a case-by-case approach to safety assessment is important.
- Further improvements in non sequence alignment-based prediction methods and allergenic epitope mapping of various food proteins should increase prediction accuracy and confidence in the outcome of the analysis.
- Allergen source labeling in packaged food is an important regulatory tool to protect consumers.
- Responsible innovation practices by the food industry can ensure safe food.

Acknowledgements

SOT
OFAS Management
OFAS Colleagues