Biotechnology Consultation - Note to File
Biotechnology Notification File No. 000155

Date

July 28, 2017

Subject

HB4 Soybean with altered tolerance to environmental stresses

Keywords

Soybean; *Glycine max* (L.) Merr.; conventional variety Williams 82; *HaHB4* gene from *Helianthus annuus* (common sunflower); HAHB4 transcription factor; HD-Zip family of transcription factors; altered tolerance to environmental stresses; *bar* gene from *Streptomyces hygroscopicus*; phosphinothricin N- acetyltransferase; PAT; selectable marker for plant transformation; New Protein Consultation No. (NPC) 000016; OECD unique identifier IND-00410-5; Verdeca, LLC

Purpose

This document summarizes the Food and Drug Administration's (FDA's, our) evaluation of biotechnology notification file (BNF) No. 000155. Verdeca, LLC (Verdeca) submitted a safety and nutritional assessment of genetically engineered soybean, transformation event HB4 soybean, which we received on May 17, 2016. We received additional information from Verdeca on December 22, 2016, and May 8 and 10, 2017. We evaluated the information in Verdeca's submissions to ensure that regulatory and safety issues regarding human or animal food derived from HB4 soybean have been resolved prior to commercial distribution.

In our evaluation, we considered all information provided by Verdeca as well as publicly available information and information in the agency's files. Here we discuss the outcome of the consultation, but do not intend to restate the information provided in the final consultation in its entirety.

Intended Effects

The intended effect of the modification in HB4 soybean is to increase tolerance to a broad array of environmental stresses that reduce crop yield. To confer this trait, Verdeca introduced the *HaHB4* gene from common sunflower (*Helianthus annuus*), which encodes the HAHB4 transcription factor (HAHB4).[1] According to Verdeca, expression of the *HaHB4* gene in soybean augments the plant's response to stresses that are

---

[1] HAHB4 is the subject of New Protein Consultation No. (NPC 000016). FDA responded to NPC 000016 on August 7, 2015.

encountered during soybean growth, thereby enabling a greater seed yield than conventional commercial soybean varieties exposed to comparable environmental pressures.

Verdeca also introduced the *bar* gene from *Streptomyces hygroscopicus*. This gene encodes L-phosphinothricin acetyltransferase (PAT), which confers tolerance to the herbicide glufosinate and which was used by Verdeca as a selectable marker during development of HB4 soybean. According to Verdeca, the *bar* gene was only necessary for the plant transformation process and is not intended to provide field herbicide resistance.

Regulatory Considerations

The purposes of this evaluation are (1) to assess whether the developer has introduced into human or animal food a substance requiring premarket approval as a food additive and (2) to determine whether use of the new plant variety in human or animal food raises other regulatory issues under the Federal Food, Drug, and Cosmetic Act (FD&C Act).

Genetic Modification and Characterization

Parental variety

Verdeca used the Williams 82 soybean variety to generate HB4 soybean. Williams 82 is a commercial soybean variety that was first released in 1981.

Transformation Plasmid and Method

Verdeca transformed Williams 82 with the *pIND2-HB4* binary vector using *Agrobacterium tumefaciens* strain EHA101.

The T-DNA region of the binary vector contains two expression cassettes:

- the *HaHB4* gene from *H. annuus* with regulatory elements, including the large promoter fragment of the *HaHB4* gene from *H. annuus* and the 3′-untranslated region (3′-UTR) of the nopaline synthase gene from *A. tumefaciens*;
- the *bar* gene from *S. hygroscopicus* with regulatory elements including the 35S promoter from Cauliflower Mosaic Virus, the 5′ leader sequence from Tobacco Etch Virus, and the sequence from the 3′ terminator of a gene encoding a soybean vegetative storage protein.

Vector elements outside of the T-DNA region include the *aminoglycoside 3′-(O) adenylyltransferase* (*aadA*) gene from the pCN1 plasmid from *Shigella flexneri* serotype 2a. The *aadA* gene confers resistance to spectinomycin/streptomycin antibiotics and was used by Verdeca for plasmid replication and maintenance in bacterial cultures.

Transformation, event selection, and plant regeneration were conducted as follows:

- cotyledonary nodes (explants) were isolated from mature seeds and infected with the Agrobacterium strain carrying the binary vector;
- explants were co-cultured with the Agrobacterium strain for five to seven days;

- explants were placed on selective medium containing cefotaxime, timentin, and vancomycin (to inhibit Agrobacterium growth) and glufosinate (to inhibit the growth of non-transformed plant cells);
- shoot induction of transformants was induced over several weeks by subculturing several times on shoot induction selective medium;
- leafy stems were excised and transferred first to shoot elongation selective medium and later to semi-solid rooting medium;
- rooted plants were transferred to soil mix for growth and further assessment.

Characteristics, Stability, and Inheritance of the Introduced DNA

Verdeca used a combination of techniques, including Southern blot analysis, DNA sequencing (both conventional Sanger capillary sequencing and whole genome sequencing), and bioinformatics to characterize the insertion event in HB4 soybean. Verdeca estimates that it collected sufficient whole genome sequence data to achieve at least 30-fold read coverage for each element in the binary vector. Verdeca compared the sequence of the T-DNA and flanking sequences assembled from whole genome sequence data with (1) sequences determined by conventional Sanger sequencing and (2) sequences in the binary vector. Based on the results of its analyses, Verdeca concludes that HB4 soybean contains a single, intact copy of the T-DNA integrated within a single locus in the soybean genome. Verdeca further concludes that HB4 soybean does not contain any binary vector backbone elements, including the *aadA* gene.

Verdeca monitored the stability and integrity of the T-DNA across six self-pollinated generations (T1, T3, T5, and T6) and assessed T-DNA inheritance in a segregating population. Verdeca concluded that the T-DNA is stable over six generations and inherited as a single locus according to Mendelian inheritance principles.

Verdeca used junction sequence analysis to map the integrated T-DNA to an intergenic region of chromosome 9. Based on comparisons with the Williams 82 genomic sequence, Verdeca determined that the HB4 integration event resulted in a loss of a 142 base pairs from the genome; however, given the intergenic location of the event, Verdeca concluded that no known genes or annotated sequences were disrupted.

Verdeca used bioinformatic analysis to identify potential new open reading frames (ORFs) and to assess the potential of these sequences to produce peptides that could be toxic or allergenic. Sequences within the insert and across the flanking regions were searched for initiation and stop codon pairs capable of producing a peptide of eight amino acids or greater in length. Verdeca identified seventy-four putative peptide sequences, and compared these sequences to those in the FARRP allergen protein database[2] and in the ATDB animal toxin database.[3] Verdeca concludes from the results of these

---

[2] Food Allergy Research and Resource Program AllergenOnline.org database, Version 16.

[3] Animal Toxin Database, College of Life Sciences, Hunan Normal University (He *et al.*, 2008)

comparisons that none of the putative peptides has the potential to be allergenic[4] or toxic.[5]

Protein Characterization

Identity and Function of Introduced Proteins

HB4 soybean expresses the sunflower HAHB4 protein and the bacterial PAT protein.

HAHB4 belongs to the HD-Zip family of transcription factors. It is characterized by two functional domains: a homeobox domain (HD) that binds DNA and a leucine zipper (LZ) motif that is involved in dimerization through protein-protein interaction. The HD and LZ domains are found in many transcriptional activators in plant and animal kingdom, but their association in a single polypeptide is unique to plants, including sunflower from which the HAHB4 has been obtained. Various abiotic and biotic stresses, such as drought, salinity, darkness, insect feeding, and chemical exposure, induce the expression of HAHB4, which, in turn, regulates the expression of genes that respond to these stresses.

Verdeca reports that the HAHB4 protein sequence in HB4 soybean was determined to have some differences compared to the sunflower HAHB4 sequence (GenBank Accession number AAA63768.2), including a four amino acid deletion and three substitutions. The deletion keeps the DNA in frame and the amino acid substitutions are either conservative or putative variants of the original sequence. Based on its examination of the amino acid sequence changes, Verdeca concludes that the effects of these changes are not expected to produce significant differences in the general properties of the protein.

PAT is an enzyme that confers tolerance to glufosinate ammonium herbicides by catalyzing the conversion of phosphinothricin to N-acetylphosphinothricin. Verdeca explains that it used glufosinate selection to inhibit growth of non-transformed plant cells in order to select for the plant cells transformed with the T-DNA.

Protein Expression Levels

The levels of HAHB4 and PAT in HB4 soybean were determined from seed and leaf samples collected from six field sites in Argentina (2012-2013) and five field sites in the United States (2013). Leaf tissue samples were collected from three or more soybean plants from each of the four plots at each site.

Due to the expected low-level expression of HAHB4, Verdeca developed and validated a specific and sensitive LC-MS/MS method for the detection of HAHB4. Verdeca combined the denaturing protein gel electrophoresis of HAHB4 with in-gel trypsin digestion of the bands, and used the entire sample for the LC-MS/MS detection of tryptic

---

[4] Similarity to known allergens was defined as (1) sequence identities of greater than 35% across an 80 amino acid window or (2) alignments of eight contiguous amino acids over the whole length of the putative peptide.

[5] Relevant homology to known toxins was set at E score $< 1\text{x}10^{-5}$.

fragments. Appropriate controls and standards were used. Verdeca reports that although the levels of HAHB4 were below the lower limit of quantification (LLOQ)[6] in samples from plants grown under field trial conditions, HAHB4 was detectable in some samples. The highest amount of HAHB4 detected in field grown soybean leaves was observed at 5 ng/g dry weight (DW). Verdeca further notes that growth chamber experiments designed to elicit HAHB4 production through exposure to direct and continuous abiotic stress demonstrated similar expression levels, ranging from below the limit of detection to up to 5 ng/g DW in root and 4 ng/g DW in leaf.

An enzyme-linked immunosorbent assay (ELISA) method was used to quantify PAT levels. In seeds, mean PAT levels ranged from 23 – 69 μg/g fresh weight (FW). In forage, mean PAT levels ranged from 5 – 13 μg/g FW. Verdeca noted that these values are comparable to or less than previously reported values of PAT in genetically engineered crops. Verdeca states that the levels of PAT in HB4 soybean were not sufficient for the plants to exhibit glufosinate ammonium tolerance under field conditions.

Toxicity and Allergenicity Assessment of HAHB4

The potential for toxicity of HAHB4 was evaluated by performing bioinformatic analysis to determine its sequence homology with known toxins from the Animal Toxin Database. No significant homology to known protein toxins was found. Additionally, HAHB4 belongs to a large class of transcription factors that are present in many plant species including edible plants. As noted above, the levels of HAHB4 in the plant leaf and seed samples from plants grown under field trial conditions were below the limit of quantification.

The potential for allergenicity of HAHB4 was evaluated by assessing its thermal stability and its digestibility in simulated gastric fluid (SGF), and by comparing its sequence to known allergens. Verdeca reports that HAHB4 was heat stable when held at 90°C for up to 60 minutes but was rapidly digested (< 0.5 min) by pepsin in SGF. Bioinformatic analyses revealed neither sequence identity ≥ 35% in a sliding window of 80 amino acids nor any 8 contiguous amino acid identities between HAHB4 and allergens in any of the databases searched.[7] Verdeca states that, although HAHB4 is heat stable, the weight of the evidence supports the conclusion that HAHB4 lacks allergenic potential.

Verdeca concludes that HAHB4 is unlikely to be toxic or allergenic to humans or animals on the basis that HAHB4 is present at very low concentrations in soybean forage and grain, it is rapidly degraded in SGF, and it lacks sequence homology to known toxins and allergens.

---

[6] The LLOQ for the LC-MS/MS method was determined to be 25 fmol HAHB4/70 μg of both seed and leaf protein, equivalent to 1.37 μg HAHB4/g DW seed and 1.02 μg HAHB4 /g DW leaf tissue.

[7] Verdeca states that it conducted allergen sequence and structural homology searches using the FARRP database (updated January 2016), Allermatch (updated August 2012), and the Structural Database of Allergenic Proteins (updated February 2013).

Toxicity and Allergenicity Assessment of PAT

PAT proteins[8] have been used extensively to confer herbicide tolerance to plants cultivated under field conditions as well as in the laboratory as selectable markers for event selection during genetic engineering. In the United States, residues of the PAT enzyme are exempt from the requirement of a tolerance when used as plant-incorporated protectant inert ingredients in all food commodities (40 CFR 174.522). Additionally, FDA completed several consultations involving PAT proteins expressed in various crops including soybean. Verdeca provides a list of consultations completed by FDA in which the *pat* or the *bar* gene has been introduced into various crops to confer tolerance to the herbicide glufosinate. Verdeca states that these conclusions, along with scientific literature and food and feed approvals by other governments of crops containing either the *pat* or *bar* genes, demonstrate that the presence of PAT in plants does not render them unsafe for consumption in human or animal food.

Human and Animal Food Use

Soybean (*Glycine max*) is grown around the world for a variety of human and animal food and industrial uses. Soybean seeds are processed into oil and meal. Soybean oil is rich in polyunsaturated fatty acids and is commonly used as a salad and cooking oil and in the production of margarine and other human food ingredients. A small fraction of soybean meal is further processed into soy flours and soy proteins for a variety of food uses. Traditional foods prepared from soybeans include tofu, miso, soymilk, tempeh, and soy sauce.

The preponderance of soybean meal is used in animal food, primarily in poultry, swine, and beef and dairy cattle diets. Soybean meal is processed in moist heat to inactivate trypsin inhibitors and lectins, which are antinutrients occurring in raw soybeans.

Composition

Scope of Analyses

Verdeca analyzed the composition of soybean seed and forage from HB4 soybean, from the parental variety Williams 82 (the control), and from commercial reference varieties adapted to the geographic locations of individual field sites in the United States and Argentina. Verdeca states that it selected the reference varieties to represent the range of natural variation among commercial soybean varieties.

Study Design - Compositional Analyses

Verdeca grew HB4 soybean and the control at five sites in the United States during the 2013 growing season and at six sites in Argentina during the 2012-2013 growing season.[9]

---

[8] To date, two PAT proteins have been used in genetically engineered plants: one is encoded by the *bar* gene derived from *S. hygroscopicus*; the other is encoded by the *pat* gene derived from *S. viridochromogenes*. Both PAT proteins have previously been evaluated through FDA's biotechnology consultation process.

[9] According to Verdeca, it selected the subset of eleven field sites for compositional analysis from among a total of twenty-five field trials. The eleven sites were selected prior to planting and the selections were made based on previous trials that demonstrated reliable performance across a varying yield range.

Verdeca states that field sites were selected based on geographical distribution, location within major soybean production areas, and coverage of the diverse environmental conditions for soybeans in these two countries. A minimum of four reference varieties was grown at each of the field sites in the United States while five reference varieties were grown at each of the field sites in Argentina. The soybean varieties were planted using a randomized complete block design with four replicates at each site. Samples were collected from each replicate at each site and prepared for analysis according to published analytical methods.

Verdeca analyzed the composition of seed and forage samples.

- In seed, Verdeca reported the results for the levels of ash, carbohydrates by calculation, moisture, crude protein, crude fat, fiber (acid detergent fiber (ADF), neutral detergent fiber (NDF)), crude fiber, minerals (calcium and phosphorus), vitamins (E and K1[10]), isoflavones (daidzein, genistein, and glycitein), anti-nutrients (phytic acid, lectins, raffinose, stachyose, and trypsin inhibitor), six fatty acids and eighteen amino acids.

- In forage, Verdeca reported the levels of ash, carbohydrates by calculation, moisture, crude protein, crude fat, ADF, NDF, calcium, and phosphorus.

For each site, Verdeca used the results from the individual replicates to calculate mean levels, the standard error of the mean, and the mean difference between HB4 soybean and the control. Statistically significant differences (established at $\alpha = 0.05$) were analyzed using a two-way analysis of variance approach and the Least Significant Difference post-test. Differences between HB4 soybean and the control were compared across all sites (combined-site) and for each individual site.

When statistically significant results between HB4 soybean and control were seen, Verdeca compared mean levels for HB4 soybean to (1) the reference values determined from the commercial varieties grown in its field trials and (2) publicly available data.[11] Verdeca considered the levels of components in HB4 soybean to be within the normal variability of commercial soybean varieties when mean values were within the reference values or the publicly available values.

Results of Analyses

<u>Soybean seed</u>

Verdeca statistically analyzed and reported the results for seed from HB4 soybean, the control, and the commercial reference varieties. In the combined site analyses, statistically significant differences between HB4 soybean and the control were observed

---

[10] Vitamin K1 was measured in samples from U.S. field trials only.

[11] Data provided in the Organisation for Economic Co-operation and Development (OECD) (2012) “Revised Consensus Document on Compositional Considerations for New Varieties of SOYBEAN [*Glycine max* (L.) Merr]: Key Food and Feed Nutrients, Anti-nutrients, Toxicants, and Allergens.” OECD cited data from International Life Science Institute (2010) *Crop Composition Database Version 4.0* or Souci et al. (2008) *Food Composition and Nutrition Tables, 7th revised and completed edition*.

for cysteine, vitamin K1, phytic acid, stachyose, and the isoflavones daidzein, genistein, and glycitein; however, the levels for these components for HB4 soybean were within the reference range, the publicly available data, or both.

In the individual site analyses, there were some statistically significant differences between HB4 soybean and the control. In most cases, HB4 soybean and control were within the variability of the commercial varieties, publicly available values, or both. Some components for HB4 soybean at one or more sites fell outside of the reference varieties, publicly available data or both; these were ADF, vitamin K1, lysine, phytic acid, stachyose, and raffinose. Verdeca concluded, however, that the differences were not consistent overall and HB4 soybean values for these components did not vary substantially from the control.

Soybean forage

Verdeca statistically analyzed and reported the results for forage from HB4 soybean, the control, and the commercial reference varieties. In the combined site analysis, no statistically significant differences between HB4 soybean forage and the control were observed. In the individual site analyses, Verdeca reported statistically significant differences between HB4 soybean forage and the control for carbohydrates and fat at one site and, ADF and NDF at two sites; however, Verdeca noted that the HB4 soybean mean levels at these sites were within the ranges observed for the commercial reference varieties.

Summary of Compositional Analyses

Verdeca concludes that the levels of nutrients, antinutrients, and isoflavones in seed and forage from the HB4 soybean are similar to those from the control and are within the range of the commercial reference varieties.

Endogenous Allergens

Verdeca considered the potential for the genetic modification to alter endogenous soybean allergen levels. Specifically, they considered:

- the avoidance of allergic consumers of soybean products;
- limitations on the relevance of endogenous allergen levels in food safety assessment, as supported by publications on varietal and environmental effects on allergen levels and a lack of threshold data for either sensitization or elicitation phase of an allergic reaction;
- the function of the HAHB4 transcription factor and its target sequences; and
- the results of activity assays of the Kunitz trypsin inhibitor, an anti-nutrient and soybean allergen.

Verdeca concludes that the likelihood of upregulating an endogenous allergen in HB4 soybean to significant levels is low and that, in the unlikely event that levels were up-regulated, the magnitude would be no greater than changes due to environmental conditions, falling within the natural variation range.

Conclusion

FDA evaluated Verdeca's submission to determine whether HB4 soybean raises any safety or regulatory issues with respect to its uses in human or animal food. Based on the information provided by the company and other information available to the agency, FDA did not identify any safety or regulatory issues under the Federal Food, Drug, and Cosmetic Act that would require further evaluation at this time.

Verdeca has concluded that its stress tolerant variety, HB4 soybean, and the human and animal foods derived from it are as safe as and are not materially different in composition or any other relevant parameter from other soybean varieties now grown, marketed, and consumed. At this time, based on Verdeca's data and information, the agency considers Verdeca's consultation on HB4 soybean to be complete.

Robert I. Merker -S
Digitally signed by Robert I. Merker -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=1300070407, cn=Robert I. Merker -S
Date: 2017.08.11 07:39:08 -04'00'
for

Carrie McMahon, Ph.D.
Consumer Safety Officer