

## A Gene-Shuffled Glyphosate Acetyltransferase Protein from *Bacillus licheniformis* (GAT4601) Shows No Evidence of Allergenicity or Toxicity

Bryan Delaney,<sup>\*,1</sup> John Zhang,<sup>\*</sup> Gabrielle Carlson,<sup>\*</sup> Jean Schmidt,<sup>\*</sup> Barb Stagg,<sup>\*</sup> Brad Comstock,<sup>\*</sup> Amy Babb,<sup>†</sup> Carol Finlay,<sup>‡</sup> Robert F. Cressman,<sup>§</sup> Greg Ladics,<sup>§</sup> Amarin Cogburn,<sup>§</sup> Dan Siehl,<sup>¶</sup> Luda Bardina,<sup>||</sup> Hugh Sampson,<sup>||</sup> and Youngshin Han<sup>|||</sup>

<sup>\*</sup>Pioneer Hi-Bred International, Inc., Johnston, Iowa 50131; <sup>†</sup>Charles River Laboratories, Redfield, Arkansas 72132; <sup>‡</sup>DuPont Haskell Laboratory, Newark, Delaware 19714; <sup>§</sup>DuPont Agriculture and Nutrition, Wilmington, Delaware 19880; <sup>¶</sup>Pioneer Hi-Bred International, Inc., Verdia Campus, Redwood City, California 94063; <sup>||</sup>Department of Pediatrics and Immunobiology, Mount Sinai School of Medicine, New York, New York 10029; and <sup>|||</sup>Sungkyunkwan University School of Medicine, Seoul, Korea

Received October 10, 2007; accepted December 11, 2007

The *glyphosate acetyltransferase* (*gat*) gene from *Bacillus licheniformis* was subjected to multiple rounds of gene shuffling to optimize kinetics of corresponding GAT proteins to acetylate the herbicide active ingredient glyphosate. Genetically modified soybeans expressing the *gat4601* gene (356043 soybeans) are tolerant to the application of glyphosate. The current manuscript reports the outcome of the allergenicity and toxicity assessment for the GAT4601 protein. Bioinformatic comparison of the amino acid sequence of GAT4601 did not identify similarities to known allergenic or toxic proteins. *In vitro* studies conducted with heterologously produced GAT4601 protein demonstrated that it was rapidly degraded in simulated gastric fluid containing pepsin (< 30 s) and in simulated intestinal fluid containing pancreatin (< 2 min) and completely inactivated at temperatures above 56°C. The GAT4601 protein expressed *in planta* is not glycosylated and similar protein profiles were observed in flour extracts from 356043 soybeans and nontransgenic near isoline comparator soybeans (Jack) using serum from soy allergic persons. No evidence of adverse effects was observed in mice following acute oral exposure to 2000 mg/kg of GAT4601 protein or in a repeated dose dietary exposure study at doses of 800–1000 mg/kg/day. This comprehensive assessment demonstrates that the GAT4601 protein does not present a risk for adverse effects in humans when used in the context of agricultural biotechnology.

**Key Words:** glyphosate acetyltransferase; bioinformatics; *in vitro* digestibility; allergenicity; toxicology; transgenic crop; gene shuffling.

Three alleles of a *glyphosate acetyltransferase* (*gat*) gene from the ubiquitous *Bacillus licheniformis* soil bacteria were isolated and subjected to numerous rounds of gene shuffling to optimize the ability to acetylate the herbicidal active ingredient glyphosate in the presence of acetyl-CoA (Castle *et al.*, 2004;

Siehl *et al.*, 2005). Following 7–11 rounds of shuffling, the efficiency of the GAT protein to acetylate glyphosate ( $k_{cat}/K_M$ ) was increased approximately 5000 times compared with the native GAT enzyme. Expression of gene-shuffled variants of the GAT protein in field crops confers tolerance to glyphosate (Castle *et al.*, 2004).

The gene from the seventh-round variant of GAT (*gat4601*) results in expression of a protein (GAT4601) with a high degree of similarity to the amino acid sequence of the native GAT protein. Twenty-one amino acid substitutions in GAT4601 were identified relative to the 146 amino acids of native GAT proteins. Only 4 of the 21 substitutions are in the active site of the enzyme, which is defined as lying within 5 Å of either substrate (acetyl-CoA or glyphosate; Siehl *et al.*, 2007). None of the substitutions include amino acids identified through kinetic analysis of site-directed mutants as being involved with substrate binding or catalysis. The four substitutions in the active site (Y31F, V114A, I132T, and I135V) are to slightly smaller side-chains, which appear to accommodate glyphosate, a larger ligand, compared with the best known substrate for the native enzyme (D-2-amino-3-phosphonopropionate).

The *gat4601* gene was transferred into the germplasm of soybeans to produce Optimum GAT soybeans that are tolerant to field applications of glyphosate. Although proteins as a general class of dietary substances are not associated with adverse effects, two safety concerns that have been raised are allergenicity and acute toxicity (Metcalf *et al.*, 1996; Sjoblad *et al.*, 1992). Accordingly, a number of guidance documents have been published by scientific authorities describing processes to assess the potential allergenicity and toxicity of transgenic proteins used in the context of agricultural biotechnology (CODEX, 2003; FAO/WHO, 2001).

No single biochemical property can predict the allergenic potential of individual proteins and there are no validated animal models that accurately predict their allergenic potential.

<sup>1</sup> To whom correspondence should be addressed at Pioneer Hi-Bred International, Inc., 7250 NW 62nd Ave., P.O. Box 552, Johnston, IA 50131-0552. Fax: (515) 334-4478. E-mail: bryan.delaney@pioneer.com.

Consequently, a weight-of-evidence approach that takes into account a variety of relevant factors and experimental observations is used to derive an overall assessment of the allergenic potential of transgenic proteins such as those used on genetically modified (GM) crops (CODEX, 2003; Taylor, 2006). The allergenicity assessments for transgenic proteins are typically based on properties of known allergenic food proteins and include information about the history of exposure and safety of the gene source, amino acid sequence identity compared with known allergenic proteins, physicochemical properties such as glycosylation, stability to pepsin digestion *in vitro* or other digestive enzymes such as pepsin and pancreatin (Astwood and Fuchs, 1996; Ladics *et al.*, 2006; Thomas *et al.*, 2004). In some cases, the effect of processing (e.g., heating) on heterologously produced transgenic proteins may also be assessed. These studies are typically conducted using transgenic proteins obtained from heterologous (i.e., bacterial) expression systems following demonstration that they are equivalent to the protein expressed *in planta*.

Evaluation of the potential toxicity of transgenic proteins is similar to the allergenicity assessment of proteins because both rely on a multiple component strategy rather than emphasizing any one particular type of study. However, it differs in that it is based on a tiered approach rather than weight-of-evidence (Delaney *et al.*, in press). The components of the first tier include an assessment of the history of use of the organism from which the gene was obtained especially when there is a documentable history of human exposure or consumption, bioinformatic comparison of the amino acid sequence of the transgenic protein for similarity to other proteins that are considered toxic, information about the mechanism of action of the protein, *in vitro* stability to digestive enzymes, and exposure assessment. Although the assessment of transgenic proteins using concepts in the first tier is comprehensive, a second tier was developed to be applied on a case-by-case basis depending on the nature of the particular protein. Specific studies within that tier include an assessment of the acute toxicity of the particular transgenic protein and hypothesis-based testing. Repeated dose toxicity studies with transgenic proteins may also be requested in some cases by regulatory authorities (European Commission, 2003).

This manuscript describes the comprehensive safety assessment process that was conducted to evaluate the potential allergenicity and toxicity of the gene-shuffled GAT4601 protein. The results demonstrate that the introduction of the gene-shuffled variant of the GAT4601 protein into 356043 soybeans (i.e., Optimum GAT soybeans) represents no risk for allergenicity or toxicity when used in the context of agricultural biotechnology.

## MATERIALS AND METHODS

**Allergenicity assessment.** Similarity of the amino acid sequence of GAT4601 to known allergenic proteins was evaluated by FASTA34

bioinformatic comparison with the sequences of known allergenic proteins in the Allergen Database (Version 6.0 [January 2006]) at the Food Allergy Research and Research Program at the University of Nebraska (Lincoln, NE). Resistance of GAT4601 protein to digestion in simulated gastric fluid (SGF) containing pepsin containing pancreatin was determined by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) analysis of pure heterologously produced GAT4601 protein following incubation for periods of time ranging from 0.5 to 60 min. Resistance to digestion in simulated intestinal fluid (SIF) was assessed by SDS-PAGE analysis and Western blot analysis of the GAT4601 protein following incubation for periods of time ranging from 0.5 to 60 min. The thermal stability of heterologously produced GAT4601 protein was assessed by determination of the enzymatic activity following exposure of the GAT4601 protein to temperatures ranging from 36°C to 60°C. The glycosylation of GAT4601 protein isolated from Optimum GAT soybeans (Event DP-356043-5; 356043 soybeans) in comparison with a glycoprotein-positive control (horseradish peroxidase) and a glycoprotein-negative control (soybean trypsin inhibitor) was assessed using a GelCode Glycoprotein Staining kit (Pierce Biotechnology, Inc., Rockford, IL). The impact of genetic modification on expression of allergenic proteins in 356043 soybeans was assessed using one-dimensional IgE immunoblot analysis and enzyme-linked immunosorbent assay (ELISA) inhibition analysis of protein extracts from 356043 and control (non-GM) soybeans using sera from patients with documented clinical sensitivity to soybeans or from nonatopic patients.

**Toxicity assessment.** Similarity of the amino acid sequence of GAT4601 to the sequences of proteins in the National Center for Biotechnology Information Protein dataset was assessed by BLASTP bioinformatic analysis followed by manual inspection of sequence matches above a predetermined *E* score cutoff of 1.0. The acute oral toxicity of heterologously produced GAT4601 protein in mice was assessed per OECD 423 Guidelines by analysis of clinical signs and body weights following gavage at the limit dose of 2000 mg/kg (OECD, 2001). The repeated dose oral toxicity of GAT4601 protein was assessed in mice by blending heterologously produced GAT4601 protein into rodent diets (Purina Mills International, PMI LLC Certified Rodent LabDiet 5002, Richmond, IN) at concentrations corresponding to daily doses of approximately 10, 100, and 1000 mg/kg/day. Mice in the control groups consumed PMI 5002 rodent diets. Diets were prepared and administered on a weekly basis for a total of 27 days. The concentration, homogeneity, and stability of the GAT4601 protein were determined by antibody specific ELISA at Pioneer Hi-Bred International, Inc. (Johnston, IA). Assessment of the health of animals consuming the control and experimental diets was carried out, including body weights, feed consumption, clinical observations, ophthalmology, hematology, clinical chemistry, organ weights, and histopathology, per OECD 407 Guidelines (OECD, 1995).

## RESULTS

### *History of Use of B. licheniformis and Proteins from B. licheniformis in Foods*

The *gat4601* gene was derived from *Bacillus licheniformis*; a ubiquitous gram-positive soil bacteria. *Bacillus licheniformis* has a history of safe use as a host organism in the production of food enzymes (e.g., alpha-amylase, cyclodextrin glycosyltransferase, hemicellulase, protease, pullulanase), biocontrol agents, and as a dietary probiotic in the United States, Canada, and Europe (Alexopoulos *et al.*, 2004a,b; European Commission, 2000; Kritas *et al.*, 2006; U.S. FDA, 2001). *Bacillus licheniformis* was determined by the U.S. Environmental Protection Agency (EPA) to present a low risk of adverse effects to human health and the environment (U.S. EPA, 1996). Although there

have been reports of particular detergent enzymes from *Bacillus* species (including *B. licheniformis*) causing occupational asthma in detergent industry workers (Biagini *et al.*, 1996; Hole *et al.*, 2000; Sarlo *et al.*, 1997), there are no documented reports of *B. licheniformis* or proteins derived from this organism causing allergic or toxic reactions in humans or other mammalian species following oral exposure.

#### Bioinformatic Analysis of the GAT4601 Protein Amino Acid Sequence

Although GAT4601 is a synthetic protein produced from shuffling of the native *gat* gene from *B. licheniformis*, it is 86% identical at the amino acid sequence level to the translated protein sequences of each of the three original *gat* alleles from which it was derived. It contains the definitive motif for the Gcn5-related N-acetyltransferase family of N-acetyltransferases; a superfamily of enzymes present in all organisms, including plants, mammals, fungi, algae, and bacteria (Dyda *et al.*, 2000; Marchler-Bauer *et al.*, 2005; Neuwald and Landsman, 1997; Vetting *et al.*, 2005).

No matching sequences of eight or more contiguous amino acids and no significant alignments were observed between the amino acid sequence of GAT4601 when compared with the sequences of known allergenic proteins with a bioinformatic search. Similarly, no significant sequence similarity between the GAT4601 protein and proteins with known toxic or antinutritional qualities was identified using BLASTP bioinformatic analysis.

#### Isolation and Characterization of Heterologously Produced GAT4601 Protein

Recombinant GAT4601 protein for the studies described in subsequent sections was expressed in *Escherichia coli* strain BL21(DE3) and purified as a soluble protein. The microbial GAT4601 protein migrated as a single band with a molecular weight of approximately 17 kDa by SDS-PAGE (Supplemental Fig. S1). The purity of the GAT4601 protein was greater than 95% on a total protein basis. The concentration of GAT4601 protein on a mass basis in the lyophilized powder was determined to be 84% by amino acid composition analysis (data not presented). The GAT4601 protein demonstrated immunoreactivity in Western blot analysis using a GAT-specific antibody, which recognized a single protein band migrating at approximately 17 kDa (Supplemental Fig. S1). Additional analytical studies conducted to characterize the heterologously produced GAT4601 protein demonstrated that it was of high purity and enzymatically active (data not shown).

#### Digestion of GAT4601 in SGF and SIF

The resistance of heterologously produced GAT4601 to digestive enzymes was assessed by incubation in SGF (containing pepsin) or SIF (containing pancreatin) followed by SDS-PAGE analysis to evaluate apparent disappearance of

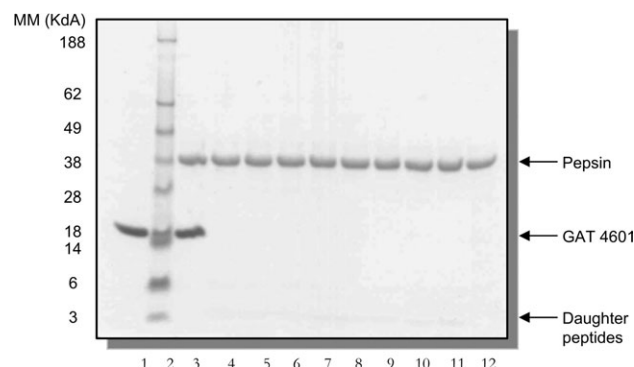


**FIG. 1.** Pepsin digestion of GAT4601 protein. Purified heterologously produced GAT4601 protein was incubated in SGF containing pepsin for 0–60 min at 37°C. Lane 1 contains GAT4601 protein (2.3 µg); lane 2 contains molecular weight markers; lane 3 contains GAT4601 at “time 0”; lanes 4–11 contain GAT4601 following SGF digestion of 30 s, 1, 2, 5, 10, 20, 30, and 60 min, respectively; lane 12 contains SGF control ~60 min.

protein bands. The GAT4601 protein was rapidly hydrolyzed (< 30 s) in SGF (Fig. 1) and in SIF (< 2 min; Fig. 2). Similarly, the GAT4601 protein band was only detectable in Western blot analysis following incubation in SIF for less than 2 min (Supplemental Fig. S2). These results demonstrated that the GAT4601 protein is not resistant to degradation by digestive enzymes.

#### Thermal Stability of GAT4601

The thermal stability of GAT4601 was evaluated by determination of the percent residual enzyme activity following heat treatment compared with the activity of untreated GAT4601 protein. The enzymatic activity of GAT4601 was stable at incubation temperatures up to 49.5°C but decreased at higher temperatures. No enzymatic activity was observed at incubation temperatures above 56.1°C (Fig. 3).

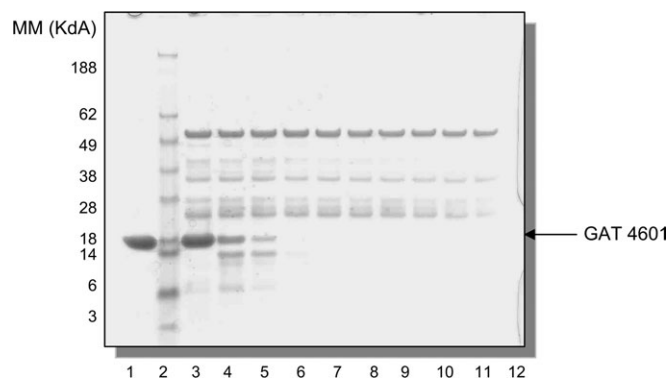


**FIG. 2.** SDS-PAGE of pancreatin digestion of GAT4601 protein. Purified heterologously produced GAT4601 protein was incubated in SIF containing pancreatin for 0–60 min at 37°C. Lane 1 contains GAT4601 protein (3.3 µg); lane 2 contains molecular weight markers; lane 3 contains GAT4601 at “time 0”; lanes 4–11 contain GAT4601 following SIF digestion of 30 s, 1, 2, 5, 10, 20, 30, and 60 min, respectively; lane 12 contains SIF control ~60 min.

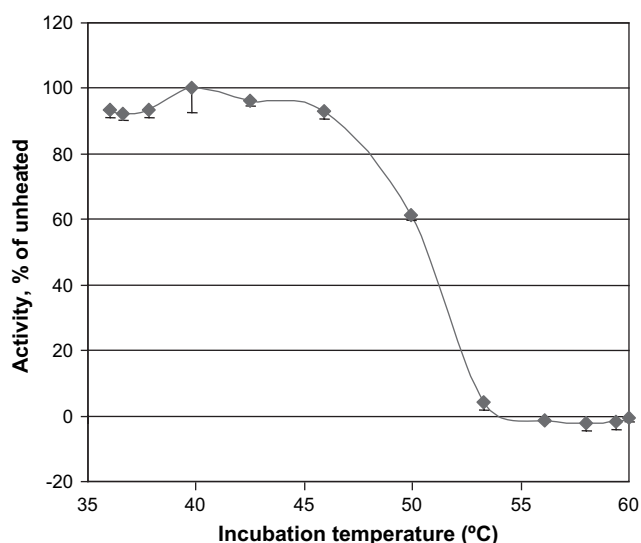


**FIG. 3.** Thermolability of GAT4601 protein. Heterologously produced GAT4601 enzyme was heated for 15 min over a temperature range of 36°C–60°C after which the enzyme activity was measured using the continuous spectrophotometric assay (described in Supplemental Materials and Methods) to monitor cleavage of the thioester bond of acetyl-CoA.

#### Glycosylation Analysis of the GAT4601 Protein

Although the presence of N-glycosylation in proteins does not necessarily indicate that they are allergenic, many allergenic proteins are glycosylated (Huby *et al.*, 2000). Therefore, the glycosylation of GAT4601 protein expressed *in planta* was assessed using protein isolated from 356043 soybean leaf tissue rather than heterologously produced protein. The results demonstrated that the GAT4601 protein expressed in 356043 soybean plants was not glycosylated (Fig. 4).

#### Immunoblot and ELISA Inhibition Analysis of GAT4601 Protein in 356043 Soybeans

Expression of the GAT4601 protein in 356043 soybeans did not affect expression of soybean proteins as observed by comparison of the SDS-PAGE protein profiles for control (Jack) and 356043 soy flour extracts (Supplemental Fig. S3). Similarly, no differences were observed in the IgE binding profiles of control and 356043 soybean proteins using one-dimensional immunoblot analysis with sera from soy allergic humans (Fig. 5).

ELISA inhibition analysis using sera from soy allergic humans showed similar inhibition patterns for control and 356043 soy flour extracts. IgE binding between control (Jack) and 356043 soybean extracts against 50 µg/ml Jack solid-phase with protein concentrations ranging from 50 to 500,000 ng/ml (Supplemental Fig. S4) did not exceed 2.7% standard deviation for 500 ng/ml, approaching 0.8% at the maximum concentration of inhibitor. Together, the SDS-PAGE, immunoblot, and ELISA inhibition analyses indicated that the protein and allergen profiles in flours from control and 356043 (Optimum GAT) soybeans were not different.

#### Acute Toxicity of GAT4601 Protein

The acute toxicity of the GAT4601 protein was assessed in mice by oral administration of 2000 mg/kg of the purified protein (corresponding to approximately 1680 mg/kg pure, full-length GAT4601 protein) via oral gavage. Control groups were administered either vehicle (water) alone or bovine serum albumin at 2000 mg/kg. All mice survived the duration of the study and no clinical signs of systemic toxicity were observed in any of the treatment groups (data not shown). Mice in all

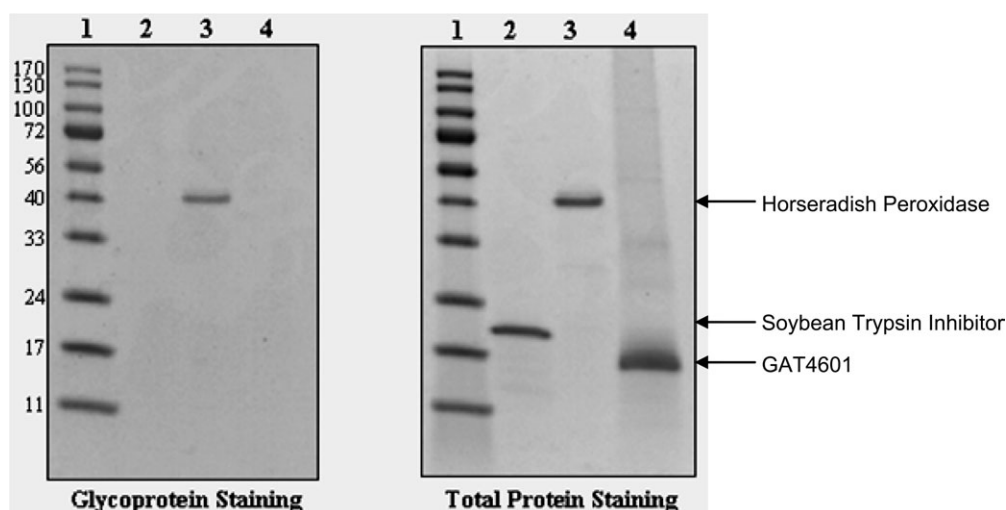


**FIG. 4.** Glycosylation analysis of GAT4601 protein expressed *in planta*. An SDS-PAGE gel was run with GAT4601 protein isolated from the leaves of 356043 soybeans (lane 4), a glycoprotein-positive control (horseradish peroxidase; lane 3), a glycoprotein-negative control (soybean trypsin inhibitor; lane 2), or prestained molecular weight markers (Fermentas; lane 1). In the left figure, the gel was stained for glycoproteins using a glycoprotein staining kit. In the right figure, the same gel was stained with Coomassie Blue to visualize individual protein bands.

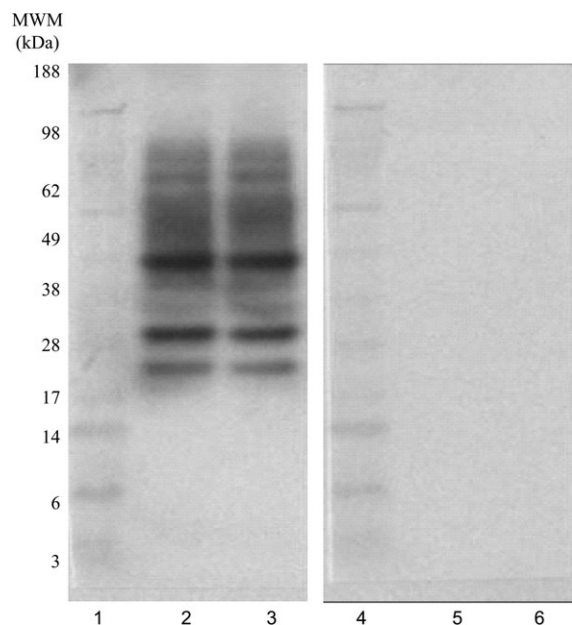


**FIG. 5.** Immunoblot of 356043 and control (Jack) soybean extracts with sera of soybean allergic patients. Western blots were probed with sera from confirmed soy atopic and control patients. Binding was determined using  $^{125}\text{I}$  labeled goat anti-human IgE. Lanes 1 and 4 are molecular weight markers (SeeBlue Plus 2); lanes 2 and 5 are Jack soybeans; and lanes 3 and 6 are 356043 soybeans.

treatment groups gained weight relative to day 0 of dosing (Supplemental Table 1) and no gross lesions were present in any of the mice at necropsy indicating that the GAT4601 protein was not acutely toxic.

#### Repeated Dose Toxicity Assessment of GAT4601 Protein

To assess the potential for adverse effects from repeated exposure, GAT4601 protein was blended into rodent diets and fed to mice for 27 days at three different doses (10, 100, and 1000 mg/kg/day). No significant differences were observed in body weights over the course of the in-life phase of the study (Supplementary Figs. S5 and S6). Similarly, no differences were observed in feed consumption between control groups and those consuming diets containing the GAT4601 protein evaluated in this study (data not shown). Consumed dosages of GAT4601 protein were determined by ELISA analysis of dietary samples collected over the duration of the in-life phase of the study. In males, daily exposure to GAT4601 protein averaged 7.8, 76.7, and 783.1 mg/kg/day for the low, middle, and high dose groups, respectively. Consumed dosages for females averaged 9.2, 94.3, and 926.9 mg/kg/day for the low, middle, and high dose groups, respectively.

No deaths occurred and there were no abnormal clinical or ophthalmological observations in any of the treatment groups (data not shown). No adverse effects in clinical chemistry response (Tables 1 and 2) or hematology (Supplemental Tables 2 and 3) response variables or organ weights (Table 3) were

**TABLE 1**  
Mean Clinical Chemistry Values of Male Mice ( $\pm$  SD) following Repeated Dose Dietary Exposure to GAT4601 Protein

|              | Control<br><i>n</i> = 5 | Low<br><i>n</i> = 5 | Middle<br><i>n</i> = 5 | High<br><i>n</i> = 4 <sup>a</sup> |
|--------------|-------------------------|---------------------|------------------------|-----------------------------------|
| AST (U/l)    | 65.2 $\pm$ 8.3          | 104.0 $\pm$ 104.4   | 68.8 $\pm$ 12.6        | 104.5 $\pm$ 106.2                 |
| ALT (U/l)    | 34.0 $\pm$ 11.6         | 65.6 $\pm$ 67.8     | 32.6 $\pm$ 7.5         | 64.0 $\pm$ 61.0                   |
| ALKP (U/l)   | 57.4 $\pm$ 16.0         | 67.0 $\pm$ 11.7     | 66.0 $\pm$ 21.0        | 64.8 $\pm$ 9.5                    |
| GGT (U/l)    | 0.0 $\pm$ 0.0           | 0.0 $\pm$ 0.0       | 0.0 $\pm$ 0.0          | 0.0 $\pm$ 0.0                     |
| BILI (mg/dl) | 0.16 $\pm$ 0.05         | 0.22 $\pm$ 0.04     | 0.22 $\pm$ 0.4         | 0.20 $\pm$ 0.00                   |
| BUN (mg/dl)  | 24.2 $\pm$ 7.9          | 23.0 $\pm$ 4.1      | 24.4 $\pm$ 3.9         | 22.0 $\pm$ 2.7                    |
| CREA (mg/dl) | 0.18 $\pm$ 0.04         | 0.20 $\pm$ 0.00     | 0.14 $\pm$ 0.05        | 0.13 $\pm$ 0.05                   |
| CHOL (mg/dl) | 94.2 $\pm$ 21.8         | 98.4 $\pm$ 5.9      | 110.4 $\pm$ 14.3       | 114.5 $\pm$ 9.9                   |
| TRIG (mg/dl) | 98.2 $\pm$ 28.7         | 94.2 $\pm$ 18.1     | 125.8 $\pm$ 35.6       | 127.5 $\pm$ 26.6                  |
| GLUC (mg/dl) | 214.6 $\pm$ 20.5        | 219.2 $\pm$ 26.7    | 209.8 $\pm$ 31.2       | 214.5 $\pm$ 25.3                  |
| TP (g/dl)    | 4.40 $\pm$ 0.25         | 4.36 $\pm$ 0.17     | 4.30 $\pm$ 0.07        | 4.40 $\pm$ 0.24                   |
| ALB (g/dl)   | 2.42 $\pm$ 0.23         | 2.40 $\pm$ 0.10     | 2.36 $\pm$ 0.05        | 2.40 $\pm$ 0.08                   |
| GLOB (g/dl)  | 1.98 $\pm$ 0.13         | 1.96 $\pm$ 0.15     | 1.94 $\pm$ 0.05        | 2.00 $\pm$ 0.18                   |
| A/G ratio    | 1.22 $\pm$ 0.15         | 1.26 $\pm$ 0.11     | 1.24 $\pm$ 0.05        | 1.18 $\pm$ 0.10                   |
| CALC (mg/dl) | 8.86 $\pm$ 0.09         | 8.92 $\pm$ 0.16     | 9.00 $\pm$ 0.19        | 9.15 $\pm$ 0.17                   |
| IPHS (mg/dl) | 7.82 $\pm$ 0.51         | 7.90 $\pm$ 0.41     | 8.52 $\pm$ 0.41        | 8.48 $\pm$ 0.38                   |
| NA (mmol/l)  | 151.2 $\pm$ 0.4         | 151.2 $\pm$ 2.4     | 151.2 $\pm$ 1.3        | 151.8 $\pm$ 2.9                   |
| K (mmol/l)   | 4.76 $\pm$ 0.25         | 5.42 $\pm$ 0.18*    | 5.08 $\pm$ 0.23        | 5.40 $\pm$ 0.34*                  |
| CL (mmol/l)  | 114.0 $\pm$ 2.7         | 114.8 $\pm$ 2.3     | 113.8 $\pm$ 1.5        | 115.0 $\pm$ 1.4                   |

Note. \**p* < 0.05 compared with control.

<sup>a</sup>Hemolyzed samples were excluded from analysis.

AST, aspartate aminotransferase; ALT, alanine aminotransferase; ALKP, alkaline phosphatase; GGT, gamma-glutamyltransferase; BILI, total bilirubin; BUN, blood urea nitrogen; CREA, creatinine; CHOL, cholesterol; TRIG, triglycerides; GLUC, glucose; TP, total protein; ALB, albumin (ALB); GLOB, globulin; A/G, albumin/globulin ratio; CALC, calcium; IPHS, inorganic phosphorus; NA, sodium; K, potassium; CL, chloride.

observed in any of the GAT4601 treatment groups that were considered treatment related. Additionally, no gross or microscopic lesions in any tissues that were considered adverse or related to exposure to the GAT4601 protein in the experimental diets (data not shown).

## DISCUSSION

The guidelines and processes used to assess the safety of transgenic proteins used in the context of agricultural biotechnology differ somewhat from those used to assess the safety of other types of food ingredients (Delaney, 2007). The assessments generally focus on evaluating the potential for particular proteins to cause allergenicity and acute toxicity. The potential for allergenicity is conducted using a multiple component weight-of-evidence approach that includes specific elements to compare the physical properties of the particular protein with those of known allergenic proteins (CODEX, 2003). Similarly, a multiple component approach has been applied to assess the toxicity of transgenic proteins (Delaney *et al.*, in press). The studies described in this manuscript were conducted to assess the allergenicity and toxicity of the GAT4601 protein.

TABLE 2

Mean Clinical Chemistry Values of Female Mice ( $\pm$  SD) following Repeated Dose Dietary Exposure to GAT4601 Protein

|              | Control<br><i>n</i> = 4 <sup>a</sup> | Low<br><i>n</i> = 5 | Middle<br><i>n</i> = 4 <sup>a</sup> | High<br><i>n</i> = 5 |
|--------------|--------------------------------------|---------------------|-------------------------------------|----------------------|
| AST (U/l)    | 120.8 $\pm$ 62.9                     | 114.8 $\pm$ 40.6    | 202.8 $\pm$ 274.2                   | 198.0 $\pm$ 184.5    |
| ALT (U/l)    | 73.8 $\pm$ 65.3                      | 56.4 $\pm$ 31.5     | 116.8 $\pm$ 163.1                   | 121.0 $\pm$ 130.0    |
| ALKP (U/l)   | 76.8 $\pm$ 22.7                      | 81.2 $\pm$ 21.4     | 92.8 $\pm$ 28.9                     | 101.4 $\pm$ 34.2     |
| GGT (U/l)    | 0.3 $\pm$ 0.5                        | 0.0 $\pm$ 0.0       | 0.0 $\pm$ 0.0                       | 0.0 $\pm$ 0.0        |
| BILI (mg/dl) | 0.25 $\pm$ 0.10                      | 0.20 $\pm$ 0.00     | 0.23 $\pm$ 0.05                     | 0.26 $\pm$ 0.09      |
| BUN (mg/dl)  | 21.3 $\pm$ 5.7                       | 20.4 $\pm$ 2.2      | 21.3 $\pm$ 2.5                      | 21.4 $\pm$ 6.3       |
| CREA (mg/dl) | 0.18 $\pm$ 0.05                      | 0.22 $\pm$ 0.04     | 0.18 $\pm$ 0.05                     | 0.20 $\pm$ 0.00      |
| CHOL (mg/dl) | 83.0 $\pm$ 5.2                       | 81.4 $\pm$ 18.8     | 72.3 $\pm$ 16.0                     | 69.8 $\pm$ 11.0      |
| TRIG (mg/dl) | 72.3 $\pm$ 10.2                      | 92.6 $\pm$ 50.4     | 98.8 $\pm$ 31.1                     | 129.6 $\pm$ 25.2     |
| GLUC (mg/dl) | 205.8 $\pm$ 19.9                     | 208.6 $\pm$ 21.9    | 213.0 $\pm$ 26.5                    | 176.2 $\pm$ 16.6     |
| TP (g/dl)    | 4.73 $\pm$ 0.13                      | 4.62 $\pm$ 0.24     | 4.68 $\pm$ 0.28                     | 4.24 $\pm$ 0.13*     |
| ALB (g/dl)   | 2.78 $\pm$ 0.10                      | 2.82 $\pm$ 0.04     | 2.75 $\pm$ 0.13                     | 2.58 $\pm$ 0.13*     |
| GLOB (g/dl)  | 1.95 $\pm$ 0.06                      | 1.80 $\pm$ 0.21     | 1.93 $\pm$ 0.17                     | 1.66 $\pm$ 0.15      |
| A/G ratio    | 1.45 $\pm$ 0.06                      | 1.60 $\pm$ 0.19     | 1.45 $\pm$ 0.10                     | 1.56 $\pm$ 0.21      |
| CALC (mg/dl) | 9.18 $\pm$ 0.36                      | 9.16 $\pm$ 0.13     | 9.43 $\pm$ 0.33                     | 9.14 $\pm$ 0.17      |
| IPHS (mg/dl) | 8.75 $\pm$ 0.64                      | 9.00 $\pm$ 1.03     | 8.65 $\pm$ 1.21                     | 8.66 $\pm$ 0.59      |
| NA (mmol/l)  | 149.8 $\pm$ 1.5                      | 150.8 $\pm$ 0.8     | 150.0 $\pm$ 1.2                     | 151.6 $\pm$ 1.1      |
| K (mmol/l)   | 5.23 $\pm$ 0.39                      | 5.12 $\pm$ 0.39     | 5.45 $\pm$ 0.95                     | 5.54 $\pm$ 0.94      |
| CL (mmol/l)  | 114.0 $\pm$ 2.2                      | 115.0 $\pm$ 0.7     | 114.3 $\pm$ 1.3                     | 113.0 $\pm$ 3.4      |

\**p* < 0.05 compared with control.

<sup>a</sup>Hemolyzed samples were excluded from analysis.

AST, aspartate aminotransferase; ALT, alanine aminotransferase; ALKP, alkaline phosphatase; GGT, gamma-glutamyltransferase; BILI, total bilirubin; BUN, blood urea nitrogen; CREA, creatinine; CHOL, cholesterol; TRIG, triglycerides; GLUC, glucose; TP, total protein; ALB, albumin (ALB); GLOB, globulin; A/G, albumin/globulin ratio; CALC, calcium; IPHS, inorganic phosphorus; NA, sodium; K, potassium; CL, chloride.

The *gat4601* gene was produced from multiple rounds of gene shuffling of *gat* genes from *B. licheniformis* to increase the efficiency by which the protein acetylates the herbicidal ingredient glyphosate (Castle *et al.*, 2004; Siehl *et al.*, 2005). This gene was transferred into the germplasm of soybean seeds to produce Optimum GAT soybeans (356043) that are tolerant to field application of herbicides containing glyphosate.

The *gat4601* gene was obtained from an organism (*B. licheniformis*) with a documented history of safe use in the human food supply, does not have amino acid sequence similarity to known allergenic proteins, is not resistant to *in vitro* gastric and intestinal digestion models, and is heat labile. Additionally, the GAT4601 protein expressed in 356043 soybeans was not glycosylated, a property that has been reported for many allergenic proteins (Lack, 2002; Taylor and Lehrer, 1996). When considered in the context of the weight-of-evidence approach, these results demonstrated that the GAT4601 protein bears no similarity to allergenic proteins and is therefore does not represent a risk for allergenicity (CODEX, 2003; Taylor, 2006).

Soybeans are one of the eight major foods known to cause allergic effects and a number of allergenic proteins have been

TABLE 3

Mean Organ:Body Weight Ratios of Mice ( $\pm$  SD) following Repeated Dose Dietary Exposure to GAT4601 Protein

|                    | Control<br><i>n</i> = 5 | Low<br><i>n</i> = 5 | Middle<br><i>n</i> = 5 | High<br><i>n</i> = 5 |
|--------------------|-------------------------|---------------------|------------------------|----------------------|
| <b>Males</b>       |                         |                     |                        |                      |
| Body wt (g, final) | 33.6 $\pm$ 3.6          | 33.6 $\pm$ 0.9      | 34.2 $\pm$ 1.1         | 34.6 $\pm$ 3.2       |
| Adrenals           | 0.0328 $\pm$ 0.0073     | 0.0270 $\pm$ 0.0058 | 0.0258 $\pm$ 0.0068    | 0.0306 $\pm$ 0.0051  |
| Brain              | 1.505 $\pm$ 0.217       | 1.497 $\pm$ 0.078   | 1.493 $\pm$ 0.049      | 1.457 $\pm$ 0.125    |
| Epididymides       | 0.354 $\pm$ 0.066       | 0.385 $\pm$ 0.129   | 0.394 $\pm$ 0.100      | 0.346 $\pm$ 0.044    |
| Heart              | 0.669 $\pm$ 0.154       | 0.591 $\pm$ 0.054   | 0.637 $\pm$ 0.131      | 0.625 $\pm$ 0.072    |
| Kidneys            | 1.844 $\pm$ 0.226       | 1.768 $\pm$ 0.162   | 1.858 $\pm$ 0.235      | 1.682 $\pm$ 0.186    |
| Liver              | 5.245 $\pm$ 0.523       | 5.429 $\pm$ 0.438   | 5.365 $\pm$ 0.317      | 5.465 $\pm$ 0.658    |
| Lungs              | 0.633 $\pm$ 0.077       | 0.620 $\pm$ 0.076   | 0.617 $\pm$ 0.063      | 0.649 $\pm$ 0.131    |
| Spleen             | 0.330 $\pm$ 0.077       | 0.301 $\pm$ 0.066   | 0.284 $\pm$ 0.025      | 0.281 $\pm$ 0.031    |
| Testes             | 0.752 $\pm$ 0.115       | 0.728 $\pm$ 0.080   | 0.728 $\pm$ 0.126      | 0.711 $\pm$ 0.051    |
| Thymus             | 0.138 $\pm$ 0.041       | 0.140 $\pm$ 0.033   | 0.134 $\pm$ 0.061      | 0.148 $\pm$ 0.040    |
| <b>Females</b>     |                         |                     |                        |                      |
| Body wt (g, final) | 28.4 $\pm$ 1.1          | 27.8 $\pm$ 1.6      | 27.0 $\pm$ 2.4         | 28.4 $\pm$ 1.5       |
| Adrenals           | 0.055 $\pm$ 0.015       | 0.055 $\pm$ 0.014   | 0.052 $\pm$ 0.014      | 0.048 $\pm$ 0.006    |
| Brain              | 1.803 $\pm$ 0.131       | 1.839 $\pm$ 0.160   | 1.806 $\pm$ 0.238      | 1.872 $\pm$ 0.169    |
| Heart              | 0.566 $\pm$ 0.037       | 0.642 $\pm$ 0.108   | 0.602 $\pm$ 0.057      | 0.565 $\pm$ 0.037    |
| Kidneys            | 1.416 $\pm$ 0.132       | 1.577 $\pm$ 0.146   | 1.473 $\pm$ 0.221      | 1.587 $\pm$ 0.117    |
| Liver              | 5.298 $\pm$ 0.332       | 5.215 $\pm$ 0.467   | 5.229 $\pm$ 0.366      | 5.780 $\pm$ 0.504    |
| Lungs              | 0.698 $\pm$ 0.068       | 0.716 $\pm$ 0.081   | 0.723 $\pm$ 0.082      | 0.786 $\pm$ 0.129    |
| Ovary              | 0.0808 $\pm$ 0.0362     | 0.0690 $\pm$ 0.0081 | 0.0846 $\pm$ 0.0238    | 0.0834 $\pm$ 0.0198  |
| Spleen             | 0.424 $\pm$ 0.063       | 0.341 $\pm$ 0.064   | 0.415 $\pm$ 0.051      | 0.337 $\pm$ 0.039    |
| Thymus             | 0.182 $\pm$ 0.054       | 0.201 $\pm$ 0.049   | 0.219 $\pm$ 0.028      | 0.184 $\pm$ 0.047    |

identified (Cordle, 2004; Metcalfe *et al.*, 1996; Ogawa *et al.*, 1995, 2000). The serum of soy allergic patients contains IgE antibodies specific to one or more of these allergenic soy proteins. In the current report, serum from clinically reactive soy allergic humans was used to demonstrate that the genetic modification used to produce 356043 soybeans did not alter the allergen content of the seed crop compared with nontransgenic soybeans (see Goodman and Leach, 2004). From these studies it was concluded that no new allergenic proteins are present in 356043 soybeans.

Although the overwhelming majority of dietary proteins are degraded to constituent amino acids and small peptides that are absorbed for nutritive purposes in the mammalian gastrointestinal system, some proteins possess the potential to cause toxicity. Recommendations have recently been developed to assess the toxicity of transgenic proteins using a tiered approach (Delaney *et al.*, in press). Application of the elements from the first tier to the GAT4601 protein in the current paper did not identify evidence for potential toxicity of the GAT4601 protein. It was obtained from a source with a history of safe use, does not have amino acid sequence similarity to proteins known to be toxic, and was readily degraded in the presence of digestive enzymes. The mechanism of action of the GAT4601 protein is enzymatic acetylation of glyphosate. Amino acid substitutions in GAT4601 relative to the native GAT protein to favor acetylation of glyphosate do not result in significant alterations in the structure or enzymatic mechanism of action. Analysis of the crystal structure demonstrated that only 4 of the

21 amino acid substitutions are in the active site, which is defined as lying within 5 Å of either the acetyl-CoA or glyphosate substrates (Siehl *et al.*, 2007). The crystal structure, combined with kinetic characterization of site-directed mutants, defined structural elements that are involved in substrate binding and catalysis, all of which are retained in GAT4601, indicating that its catalytic mechanism is unchanged from the native GAT protein expressed in *B. licheniformis*.

Although there was no indication from the first tier of analysis that the GAT4601 presented any potential for adverse effects, additional studies were conducted with this protein because it is the first known gene-shuffled protein to be used in the context of agricultural biotechnology. With regard to the second tier of analysis, no evidence of toxicity was observed in mice following acute or repeated oral exposure to heterologously produced GAT4601. The overall nutritional contribution of the GAT4601 protein was not substantial as even the highest dose of GAT4601 administered in the diets accounted for less than 2% of total dietary protein intake. Accordingly, the No-Observed-Adverse-Effect-Level was determined to be greater than the high dose group (target of 1000 mg/kg/day) evaluated in this study. The actual exposure to pure GAT4601 protein was adjusted to 783 mg/kg/day for males and 927 mg/kg/day for females based on variability of body weights and feed consumption in the repeated dose study.

The doses of GAT4601 protein that were evaluated in the repeated dose toxicity study can be placed into context by comparison with the concentrations of this protein that have been detected in Optimum GAT soybeans using antibody specific ELISA (0.26 mg of GAT4601 protein/kg on a dry weight basis). Assuming an average body weight of 70 kg, humans would need to consume 54,810 mg/day of the GAT4601 protein in the diet to match the same exposure of the high dose male mice in the repeated dose study, which would require consumption of greater than 210,000 kg of raw soybeans per day. Even this is a very conservative estimate of human exposure to the GAT4601 protein because soybeans fractions that are consumed by humans, such as soybean oil, are processed to produce edible fractions. The processing of soybeans to edible fractions will further decrease protein activity because the GAT4601 protein is not stable to processing temperatures above 56°C. Additionally, the *in vitro* digestion studies conducted with the GAT4601 protein demonstrated that it is likely to be metabolized within the gut as occurs with most other dietary proteins and it therefore will not be absorbed intact.

In conclusion, the results in this paper describe the process that was followed to assess the potential allergenicity and toxicity of the GAT4601 protein expressed in Optimum GAT soybeans. Because GAT4601 was derived from a source that has no history of causing allergic effects in humans and did not present any similarity to known allergenic proteins using the step-wise, weight-of-evidence approach it was concluded that it was not allergenic. There was no evidence indicating that the GAT4601 protein is similar in sequence to known protein

toxins and studies in laboratory animals did not identify any evidence of toxicity following single or repeated dose exposure. Further, potential human exposure to the GAT4601 protein is extremely low. These data support the food and feed safety of the GAT4601 protein and indicate that the GAT4601 protein expressed in Optimum GAT soybeans presents no risk for adverse effects to humans when used in the context of agricultural biotechnology.

## SUPPLEMENTARY DATA

Supplementary data are available online at <http://toxsci.oxfordjournals.org/>.

## FUNDING

Korea Research Foundation Grant funded by the Korean Government (MOEHRD, KRF-2006-214-C00099) to Y.H.; and Pioneer Hi-Bred, International, Inc.

## REFERENCES

- Alexopoulos, C., Georgoulakis, I. E., Tzivara, A., Kritas, S. K., Siochu, A., and Kyriakis, S. C. (2004a). Field evaluation of the efficacy of a probiotic containing *Bacillus licheniformis* and *Bacillus subtilis* spores, on the health status and performance of sows and their litters. *J. Anim. Physiol. Anim. Nutr.* **88**, 381–392.
- Alexopoulos, C., Georgoulakis, I. E., Tzivara, A., Kyriakis, C. S., Govaris, A., and Kyriakis, S. C. (2004b). Field evaluation of the efficacy of a probiotic containing *Bacillus licheniformis* and *Bacillus subtilis* spores, on the health status and performance of grower and finisher pigs. *J. Vet. Med. A Physiol. Clin. Med.* **51**, 306–312.
- Astwood, J. D., and Fuchs, R. L. (1996). Allergenicity of foods derived from transgenic plants. *Monogr. Allergy* **32**, 105–120.
- Biagini, R. E., Driscoll, R. J., Bernstein, D. I., Wilcox, T. G., Henningsen, G. M., MacKenzie, B. A., Burr, G. A., Scinto, J. D., and Baumgardner, E. S. (1996). Hypersensitivity reactions and specific antibodies in workers exposed to industrial enzymes at a biotechnology plant. *J. Appl. Toxicol.* **16**, 139–145.
- Castle, L. A., Siehl, D. L., Gorton, R., Patten, P. A., Chen, Y. H., Bertain, S., Cho, H. J., Duck, N., Wong, J., Liu, D., *et al.* (2004). Discovery and directed evolution of a glyphosate tolerance gene. *Science* **304**, 1151–1154.
- CODEx. (2003). Codex Alimentarius Commission, Alinorm 03/34: Joint FAO/WHO Food Standard Programme, Codex Alimentarius Commission, Twenty-Fifth Session, Rome, Italy, June 30–July 5, 2003. Appendix III, Guideline for the conduct of food safety assessment of foods derived from recombinant-DNA plants, and Appendix IV, Annex of the assessment of possible allergenicity. pp. 47–60.
- Cordle, C. T. (2004). Soy protein allergy: Incidence and relative severity. *J. Nutr.* **134**, 1213S–1219S.
- Delaney, B. (2007). Strategies to evaluate the safety of bioengineered foods. *Int. J. Toxicol.* **26**, 389–399.
- Delaney, B., Astwood, J. D., Cunney, H., Eichen-Conn, R., Herouet-Guicheney, C., MacIntosh, S., Meyer, L. S., Privalle, L., Gao, Y.,

- Mattsson, J., *et al.* Evaluation of protein safety in the context of agricultural biotechnology. ILSI International Food Biotechnology Committee Task Force on Protein Safety. *Food Chem. Toxicol.* (in press).
- Dyda, F., Klein, D. C., and Hickman, A. B. (2000). GCN5-related N-acetyltransferases: A structural overview. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 81–103.
- European Commission. (2000). Opinion of the Scientific Committee on Food on  $\beta$ -Cyclodextrin Produced Using Cycloglycosyltransferase from a Recombinant *Bacillus licheniformis*. SCF/CS/ADD/AMI 52 Final. Available at: [http://ec.europa.eu/food/fs/sc/scf/out58\\_en.pdf](http://ec.europa.eu/food/fs/sc/scf/out58_en.pdf).
- European Commission. (2003). Guidance document for the risk assessment of genetically modified plants and derived food and feed. *Prepared for the Scientific Steering Committee by the Joint Working Group on Novel Foods and GMOs*.
- FAO/WHO. (2001). Evaluation of allergenicity of genetically modified foods. Report of a Joint FAO/WHO Expert Consultation on Allergenicity of Foods Derived from Biotechnology, January 22–25, 2001. Rome, Italy. Available at: [http://www.who.int/foodsafety/publications/biotech/en/ec\\_jan2001.pdf](http://www.who.int/foodsafety/publications/biotech/en/ec_jan2001.pdf).
- Goodman, R. E., and Leach, J. N. (2004). Assessing the allergenicity of proteins introduced into genetically modified crops using specific human IgE assays. *J. AOAC Int.* **87**, 1423–1432.
- Hole, A. M., Draper, A., Joliffe, G., Cullinan, P., Jones, M., and Taylor, A. J. (2000). Occupational asthma caused by bacillary amylase used in the detergent industry. *Occup. Environ. Med.* **57**, 840–842.
- Huby, R. D., Dearman, R. J., and Kimber, I. (2000). Why are some proteins allergens? *Toxicol. Sci.* **55**, 235–246.
- Kritas, S. K., Govaris, A., Christodouloupolous, G., and Burriel, A. R. (2006). Effect of *Bacillus licheniformis* and *Bacillus subtilis* supplementation of ewe's feed on sheep milk production and young lamb mortality. *J. Vet. Med. A Physiol. Pathol. Clin. Med.* **53**, 170–173.
- Lack, G. (2002). Clinical risk assessment of GM foods. *Toxicol. Lett.* **127**, 337–340.
- Ladics, G. S., Bardina, L., Cressman, R. F., Mattsson, J. L., and Sampson, H. A. (2006). Lack of cross-reactivity between the *Bacillus thuringiensis* derived protein Cry1F in maize grain and dust mite Der p 7 protein with human sera positive for Der p 7-IgE. *Regul. Toxicol. Pharmacol.* **44**, 136–143.
- Marchler-Bauer, A., Anderson, J. B., Cherukuri, P. F., DeWeese-Scott, C., Geer, L. Y., Gwadz, M., He, S., Hurwitz, D. I., Jackson, J. D., Ke, Z., *et al.* (2005). CDD: A conserved domain database for protein classification. *Nucleic Acids Res.* **33**, D192–D196.
- Metcalfe, D. D., Astwood, J. D., Townsend, R., Sampson, H. A., Taylor, S. L., and Fuchs, R. L. (1996). Assessment of the allergenic potential of foods from genetically engineered crop plants. *Crit. Rev. Food Sci. Nutr.* **36S**, S165–S186.
- Neuwald, A. F., and Landsman, D. (1997). GCN5-related histone N-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. *Trends Biochem. Sci.* **5**, 154–155.
- OECD. (1995). Guideline for the Testing of Chemicals. Repeated Dose 28-day Oral Toxicity Study in Rodents (407). Available at: [www.oecd.org/dataoecd/50/41/37477972.pdf](http://www.oecd.org/dataoecd/50/41/37477972.pdf). Accessed March 10, 2004.
- OECD. (2001). Acute oral toxicity—Acute toxic class method (423). Available at: [www.oecd.org/dataoecd/17/50/1948370.pdf](http://www.oecd.org/dataoecd/17/50/1948370.pdf). Accessed February 16, 2005.
- Ogawa, T., Bando, N., Tsuji, H., Nishikawa, K., and Kitamura, K. (1995). Alpha-subunit of beta-conglycinin, an allergenic protein recognized by IgE antibodies of soybean sensitive patients with atopic dermatitis. *Biosci. Biotechnol. Biochem.* **59**, 831–833.
- Ogawa, T., Samoto, M., and Takahashi, K. (2000). Soybean allergens and hypoallergenic soybean products. *J. Nutr. Sci. Vitaminol.* **46**, 271–279.
- Sarlo, K., Fletcher, E. R., Gaines, W. G., and Ritz, H. L. (1997). Respiratory allergenicity of detergent enzymes in the guinea pig intratracheal test: Association with sensitization of occupationally exposed individuals. *Fundam. Appl. Toxicol.* **39**, 44–55.
- Siehl, D., Castle, L., Gorton, R., Chen, Y. H., Bertain, S., Cho, H.-J., Keenan, R., Liu, D., and Lassner, M. (2005). Evolution of a microbial acetyltransferase for modification of glyphosate: A novel tolerance strategy. *Pest. Manage. Sci.* **61**, 235–240.
- Siehl, D., Castle, L., Gorton, R., and Keenan, R. (2007). The molecular basis of glyphosate resistance by an optimized microbial acetyltransferase. *J. Biol. Chem.* **282**, 11446–11455.
- Sjogblad, R. D., McClintock, J. T., and Engler, R. (1992). Toxicological considerations for protein components of biological pesticide products. *Regul. Toxicol. Pharmacol.* **15**, 3–9.
- Taylor, S. L. (2006). Review of the development of methodology for evaluating the human allergenic potential of novel proteins. *Mol. Nutr. Food Res.* **50**, 604–609.
- Taylor, S. L., and Lehrer, S. B. (1996). Principles and characteristics of food allergens. *Crit. Rev. Food Sci. Nutr.* **36**(Suppl), S91–S118.
- Thomas, K., Aalbers, M., Bannon, G. A., Bartels, M., Dearman, R. J., Esdaile, D. J., Fu, T. J., Glatt, C. M., Hadfield, N., Hatzos, C., *et al.* (2004). A multi-laboratory evaluation of a common in vitro pepsin digestion assay protocol used in assessing the safety of novel proteins. *Regul. Toxicol. Pharmacol.* **39**, 87–98.
- U.S. EPA. (1996). *Bacillus licheniformis* TSCA Section 5(h)(4) Exemption: Final Decision Document. Available at: [http://earth1.epa.gov/biotech\\_rule/pubs/fra/fd005.htm](http://earth1.epa.gov/biotech_rule/pubs/fra/fd005.htm). Accessed August 4, 2005.
- U.S. FDA. (2001). Partial List of Enzyme Preparations that are Used in Foods. Available at: <http://www.cfsan.fda.gov/~dms/opa-enzy.html>. Accessed August 4, 2005.
- Vetting, M. W., S de Carvalho, L. P., Yu, M., Hegde, S. S., Magnet, S., Roderick, S. L., and Blanchard, J. S. (2005). Structure and functions of the GNAT superfamily of acetyltransferases. *Arch. Biochem. Biophys.* **433**, 212–226.

## SUPPLEMENTAL MATERIALS AND METHODS

### ***Bioinformatic analysis of the amino acid sequence of the GAT4601 protein for similarity to allergens.***

The amino acid sequence of the GAT4601 protein was compared to the Food Allergy Research and Resource Program (FARRP), University of Nebraska, Allergen Database (Version 6.0 [January 2006] of <http://www.allergenonline.com/about.asp>), which contains the amino acid sequences of known and putative allergenic proteins. Potential identities between the GAT4601 amino acid sequence and those of proteins in the allergen database were evaluated using the FASTA34 sequence alignment program (Pearson and Lipman, 1988) set to predetermined default parameters (word size = 2, scoring matrix = BLOSUM50, gap creation penalty = -10, gap extension penalty = -2, *E* score cutoff = 10). Using FAO/WHO and ILSI-IFBiC criteria, alignments were reviewed for identities greater than or equal to 35% over 80 or greater residues as described previously (Ladics *et al.*, 2007) or for any 8 or greater contiguous identical amino acid matches to the allergens contained in the FARRP Allergen Database, respectively (FAO/WHO, 2001; Metcalfe *et al.*, 1996). The use of 8 contiguous, identical amino acids is based on the proposed minimum peptide length for a B cell-binding epitope (Metcalfe *et al.*, 1996).

### ***BLASTP bioinformatic search of the amino acid sequence of the GAT4601 protein for similarity to***

***protein toxins.*** A sequence similarity search was conducted with the amino acid sequence of the modified GAT4601 protein (Castle *et al.*, 2004) using the BLASTP 2.2.13 algorithm against Release 153.0 (4/15/06) of the NCBI Protein dataset (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=protein>) using an *E* score cutoff of 1.0. Sequence annotations of proteins with significant similarity to the GAT4601 amino acid sequence were manually inspected to identify sequence similarity to known toxic proteins. Although a statistically significant sequence similarity generally requires a match with an *E* score of less than 0.01, a threshold *E* score of 1.0 insures that proteins with even limited similarity will not be overlooked in the search (Pearson, 2000). Because of the large number of uncharacterized conceptual protein sequences derived from automated translations of genomic sequences, protein sequences from the RefSeq collection (<http://www.ncbi.nlm.nih.gov/RefSeq/>) were excluded from the search using the Entrez filtering statement "protein\_all [filter] NOT srddb\_refseq [prop]" to delimit the results returned. The scoring matrix used was

the default (BLOSUM62), low complexity filtering was turned off and the number of alignments returned was set to the maximum value of 2000 proteins.

**Recombinant GAT4601 protein preparation, purification, and characterization.** Recombinant GAT4601 protein was expressed in *E. coli* strain BL21(DE3) as a soluble protein and purified using cation exchange chromatography, anion exchange chromatography, and hydrophobic interaction chromatography, followed by diafiltration into 0.1 M  $\text{NH}_4\text{HCO}_3$ , pH 7.8. The purified protein solution was lyophilized and the resulting powder was stored at  $-80^\circ\text{C}$ . Prior to lyophilization, a small aliquot was tested to determine endotoxin concentration. The endotoxin content in the protein preparation was less than 10 Endotoxin Units per mg of lyophilized GAT4601 protein. Prior to conducting *in vitro* digestion or *in vivo* exposure studies (see below), the equivalency of heterologously produced GAT4601 protein and GAT4601 protein obtained from Optimum<sup>TM</sup> GAT<sup>TM</sup> soybean leaf tissue were investigated using a number of analytical techniques. The bacterial protein preparation was characterized by amino acid analysis, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), Western blot analysis using GAT4601 specific antibodies, electrospray mass spectroscopy, mass determination of tryptic peptides by matrix assisted laser desorption ionization mass spectrometry (MALDI-MS), N-terminal amino acid sequencing, enzyme activity assay, and glycoprotein staining to determine concentration, purity, identity and glycosylation of the GAT4601 protein (data not shown). From these analyses, it was determined that the bacterially produced test substance contained >95% pure GAT4601 protein on a total protein basis at a concentration of 0.84 mg GAT4601 per mg lyophilized powder and that it was equivalent to GAT4601 protein expressed *in planta*. It was therefore determined that heterologously produced GAT4601 protein was appropriate for the *in vivo* and *in vitro* studies described below.

***In vitro* pepsin resistance of GAT4601 and control proteins in simulated gastric fluid (SGF).** *In vitro* pepsin resistance assays were conducted with heterologously produced GAT4601 protein. Bovine serum albumin (BSA) and  $\beta$ -Lactoglobulin proteins (Sigma-Aldrich) were used as controls. Gastric control solution was prepared as per guidelines of US Pharmacopoeia (2000) with final concentrations of 0.084 N HCl and 35 mM NaCl at pH 1.2. The pepsin digestion solution, referred to as simulated gastric fluid

(SGF), was prepared fresh on the day of use by solubilizing high purity and activity pepsin (Sigma-Aldrich) in gastric control solution. The pepsin content was adjusted to approximately 10 units/ $\mu\text{g}$  of the GAT4601 or control proteins in the final digestion mixture based on manufacturer reported pepsin activity of 3,370 units per mg of protein, or approximately 0.74 mg/ml pepsin. SGF solution (1900  $\mu\text{l}$ ) was dispensed into 4 ml glass screw top vials and placed in a 37°C water bath for approximately 2 min prior to addition of 100  $\mu\text{l}$  of GAT4601 protein stock solution. The final concentrations in the SGF reaction mixture were 0.25 mg/ml GAT4601 protein in 0.084 N HCl, 35 mM NaCl, and 0.74 mg/ml pepsin at pH 1.2. This resulted in a molar ratio of enzyme (pepsin) to protein (GAT4601) of 1.4:1. This is equivalent to ~ 3:1 pepsin to GAT4601 ratio on a weight basis as previous studies have recommended (Thomas *et al.*, 2004). The reaction mixture was constantly mixed using a stir bar and submersible stir plate. A 120  $\mu\text{l}$  sub-sample of the digestion reaction was neutralized or stopped by mixing with 48  $\mu\text{l}$  of 200 mM  $\text{Na}_2\text{CO}_3$  (pH 11.0), 65  $\mu\text{l}$  NuPAGE LDS 4X Sample Buffer (Invitrogen Corporation, Carlsbad, CA), and 26  $\mu\text{l}$  of NuPAGE Sample Reducing Agent 10X (Invitrogen) containing 0.5 M dithiothreitol (DTT), in a tube at each of the following analytical time points (+/- 10 sec): 0.5, 1, 2, 5, 10, 20, 30, and 60 min. Time zero for each SGF reaction mixture was prepared by first neutralizing a sample of SGF and then adding test protein. Stopped samples were kept on ice until subjected to electrophoresis. Control tubes consisted of GAT4601 alone (in water, no SGF) prepared at time zero and approximately 60 min prior to electrophoresis. BSA and  $\beta$ -lactoglobulin controls were prepared at the time points 0, 1 (+/- 10 sec), and approximately 60 min. A 120  $\mu\text{l}$  aliquot of the control solutions were neutralized or stopped by mixing with 48  $\mu\text{l}$  of 200 mM  $\text{Na}_2\text{CO}_3$  (pH 11.0), 65  $\mu\text{l}$  NuPAGE LDS 4X Sample Buffer, and 26  $\mu\text{l}$  of NuPAGE Sample Reducing Agent 10X containing 0.5 M DTT. Control tubes consisted of SGF alone (no test protein) prepared at time zero and approximately 60 min.

**SDS-PAGE analysis.** SDS-PAGE was conducted on SGF samples using NuPAGE Novex 4-12% Bis-Tris Gels (Invitrogen) with 12 wells. GAT4601 and control protein samples were prepared for electrophoresis by heating for 10 min at 70°C. Samples were loaded at 20  $\mu\text{l}$ /well and SeeBlue Pre-stained Standard (Invitrogen) were loaded at 7  $\mu\text{l}$ /well. Electrophoresis was conducted with NuPAGE MES SDS running buffer (Invitrogen) at a constant 200 volts for approximately 30 min after which the gels were washed three

times for at least 5 min each with de-ionized water. After washing, gels were stained with RAPIDstain (G-Biosciences, St. Louis, MO) for approximately 1 hr and destained with de-ionized water. Destained gels were then evaluated for extent of digestion of the GAT4601 protein.

***In vitro* pancreatin resistance of GAT4601 and control proteins in simulated intestinal fluid (SIF).**

GAT4601 protein and control proteins (BSA and  $\beta$ -lactoglobulin) used in this assay were obtained from the same lots described above dissolved in water to a concentration of 5 mg/ml. Intestinal control solution contained a final concentration of 50 mM  $\text{KH}_2\text{PO}_4$  at pH 7.5 and was prepared as per US Pharmacopoeia (2000). The pancreatin digestion solution, referred to as simulated intestinal fluid (SIF), was prepared fresh on the day of use by solubilizing pancreatin (Sigma-Aldrich) in intestinal control solution at a final concentration of 1% w/v in reaction mixture. SIF solution (1900  $\mu\text{l}$ ) was dispensed into 4 ml glass screw top vials and placed in a 37°C water bath for approximately 2 min prior to addition of 100  $\mu\text{l}$  of GAT4601 protein stock solution. The final concentrations in the SIF reaction mixture were 0.25 mg/ml GAT4601 protein in 50 mM  $\text{KH}_2\text{PO}_4$  and 1% w/v pancreatin at pH 7.5. The reaction mixture was constantly mixed using a stir bar and submersible stir plate. A 120  $\mu\text{l}$  sub-sample of the digestion reaction was neutralized or stopped by mixing 46  $\mu\text{l}$  of NuPAGE LDS 4X Sample Buffer, and 18  $\mu\text{l}$  of NuPAGE Sample Reducing Agent 10X containing 0.5 M DTT, in a tube at each of the following analytical time points (+/- 10 sec): 0.5, 1, 2, 5, 10, 20, 30, and 60 min. Time zero for each SIF reaction mixture was prepared by first neutralizing a sample of SIF and then adding test protein. Stopped samples were heated for 10 min at 70°C and then kept on ice until preparation for electrophoresis. Control tubes consisting of GAT4601 alone (in water, no SIF), SIF alone (no test protein), and intestinal control solution containing GAT4601 protein (no pancreatin) were digested for 60 min and neutralized or stopped as described above. Also, SIF solution with BSA and SIF solution with  $\beta$ -lactoglobulin were prepared at time 0, 1 (+/- 10 sec), and approximately 60 min and neutralized or stopped as described above. SDS-PAGE analysis of GAT4601 and control proteins were conducted as described above in section 2.3.3.

**Western blot analysis.** Western blot analysis of the SIF digestion reaction samples was conducted to differentiate between intact GAT4601 protein and pancreatin protein bands using SDS-PAGE. Proteins in

the unstained gel were electrophoretically transferred to a polyvinylidene fluoride (PVDF) membrane using a Novex XCell II Mini-Cell blot module (Invitrogen) with NuPAGE transfer buffer (Invitrogen) run at a constant 30 volts for approximately 1 hr. The membrane containing the transferred proteins (blot) was blocked with a 5% low fat milk in phosphate buffered saline solution containing 0.05% Tween-20 (PBST; ICI Americas, Inc., Bridgewater, NJ), followed by an approximately 60 min incubation in a 1:20,000 dilution of a primary rabbit anti-GAT4601 polyclonal antibody (R4265-AP; Promega Corporation, Madison, WI) in 1 % low fat milk/PBST solution. The blot was washed with PBST at least three times for approximately 10 min each wash and incubated for approximately 45 min with a 1:20,000 dilution of a secondary donkey anti-rabbit IgG antibody conjugated to horseradish peroxidase (Promega). Unbound secondary antibody-horseradish peroxidase (HRP) conjugate was removed with at least three washes of approximately 10 min each in PBST. Antibody binding to GAT4601 protein bands was detected using an ECL chemiluminescence detection kit according to the directions of the manufacturer (Amersham Biosciences Corporation, Piscataway, NJ) following exposure to film. The film was developed in a M35A X-Omat developer (Eastman Kodak, Rochester, NY) and the exposed film was scanned and evaluated for apparent disappearance of the GAT4601 protein.

***Thermal treatment of GAT4601 protein.*** The thermal stability of GAT4601 was determined by assessing the enzymatic activity of lyophilized GAT4601 protein following treatment to different thermal conditions. Glyphosate and acetyl-CoA were obtained from Sigma-Aldrich (St. Louis, MO). Five milligrams of GAT4601 lyophilized powder was dissolved in 25 mM HEPES, pH 7.0, 100 mM KCl, 20% glycerol (HKG) to a concentration of 5 mg enzyme/ml in the stock solution. A working solution was prepared by diluting the stock solution five-fold in HKG to a concentration of 1 mg enzyme/ml. GAT4601 protein concentrations were determined from measurements of absorbance at 205 nm (extinction,  $\text{mg ml}^{-1} = 30.5 \text{ OD}$ ) after buffer exchange (NAP-5 column, GE Healthcare, Chalfont St. Giles, UK) into 50 mM  $\text{Na}_2\text{SO}_4$ . The extinction coefficient was determined by the procedure of Scopes (1994) where extinction,  $\text{mg ml}^{-1} \text{ cm}^{-1} = 27 + 120(A_{280}/A_{205})$ . A 1 M stock solution of glyphosate was prepared by dissolving the free acid in 1 equivalent of  $\text{NH}_4\text{OH}$  in 0.75x of final volume, titrating to pH 6.8 with 2 M  $\text{NH}_4\text{OH}$ , and adjusting to final volume. Ten microliters ( $\mu\text{l}$ ) of the GAT4601 protein solution was dispensed into PCR microtube strips

(VWR International, West Chester, PA) and diluted to 9  $\mu\text{M}$  with 25 mM HEPES, pH 7.2, 100 mM KCl, 10% methanol (HKM). Individual PCR wells were heated for 15 min at a designated temperature ranging from 36-60°C in a gradient thermocycler (PTC-200 Peltier Thermal Cycler, MJ Research [now Bio-Rad Laboratories], Hercules, CA). Samples were then centrifuged for 5 min at 10000 x g and 20  $\mu\text{l}$  of the supernatant was then diluted five fold with HKM for assay of enzymatic activity.

***Continuous absorbance spectrophotometric GAT4601 enzyme activity assay.*** The enzymatic activity of GAT4601 was determined using a continuous spectrophotometric enzyme activity assay that assessed cleavage of the thioester bond of acetyl-CoA (extinction,  $\text{M}^{-1} \text{cm}^{-1} = 3400$ ) by monitoring at 235 nm with a Spectramax Plus384 plate reader (Molecular Devices Corporation, Sunnyvale, CA). Heat treated GAT4601 enzyme (5  $\mu\text{l}$ ) was added to the wells of the assay plates. Reactions were initiated by addition of 295  $\mu\text{l}$  of the assay mixture consisting of 167  $\mu\text{M}$  acetyl-CoA and 2 mM glyphosate in assay buffer (25 mM HEPES, pH 6.8 and 10% ethylene glycol), resulting in a final enzyme concentration of GAT4601 protein of 0.03  $\mu\text{M}$ . Negative control reactions (Blank) were identical except that enzyme was first inactivated by heating at 95°C for 30 min prior to addition to the assay buffer. Absorbance at 235 nm was read at 5 second intervals for 1.5 min at 21°C with a Spectramax Plus384 plate reader. Activity was quantified from the slope ( $\text{mOD}_{235}/\text{min}$ ) and converted to reaction rate in  $\mu\text{M}$  CoA/min using an extinction coefficient of 3.6  $\text{mOD } \mu\text{M}^{-1} \text{cm}^{-1}$  (2.9 for the relevant 0.8 cm pathlength). Data are shown as reaction rates of heated samples expressed as a percentage of the average values of unheated control reactions. The limit of detection was determined with raw data (negative values), and is the average slope of the blank reactions minus 3x the standard deviation of those slopes.

***Glycosylation analysis of GAT4601 protein isolated from transgenic soybeans.*** GAT4601 protein was isolated from 356043 soybean leaf tissues to evaluate the glycosylation status of the protein expressed *in planta*. The protein was run on SDS-PAGE along with a glycoprotein-positive control (horseradish peroxidase) and glycoprotein-negative control (soybean trypsin inhibitor). Glycoproteins were detected with the GelCode Glycoprotein Staining Kit (Pierce, Rockford, IL) according to the manufacturer's

instructions. Following glycoprotein detection, the gel was stained with Coomassie Blue to visualize total proteins.

***IgE binding with serum from soybean allergic person in flour extracts from Control non-GM soybeans and 356043 soybeans - Immunoblot analysis.*** One dimension IgE immunoblot analysis was conducted using protein extracts from 356043 soybean and soybean flour obtained from control (Jack) soybeans. One gram each of Jack and 356043 soybean flour was extracted by continuous agitation with 9 ml of phosphate buffered saline (PBS) and a protease-inhibitor cocktail without EDTA (Roche, Indianapolis, IN) overnight at 4°C. After centrifugation at 2500g for 15 min at 4°C, the supernatant was collected, and then centrifuged at 12,000g for 3 min. The Coomassie Plus Protein Assay (Pierce, Rockford, IL) was used to determine protein concentrations. The total protein concentration in the soy flour from control (Jack) and 356043 soybeans was 10.0 and 14.4 µg/µl, respectively. Extracts were stored at -80°C. Samples were run on SDS-PAGE after heating at 70°C for 10 min with NuPAGE sample buffer in the presence of 0.05M DTT and proteins were then separated by SDS-PAGE (gradient 4 – 12% gels) following the protocol from the manufacturer. Ten µg of each soy extract were loaded per well. Resolved proteins were transferred to Immobilon-P™ membranes (Millipore, Bedford, MA) and then stained with 2.2% Coomassie Blue for total protein analyses and to confirm adequate protein transfer. For molecular weight determination MultiMark™ Multi-Colored Standard (Invitrogen) were utilized. Sera from 5 soy-sensitive patients with documented clinical sensitivity to soybean obtained from the author's clinical practice (Sampson) were pooled and utilized for this study. Normal (i.e., non-atopic) pooled human serum (Phadia, Uppsala, Sweden) was used as a negative control. Pooled sera from soy allergic and control patients were diluted 1:5 in PBS-Tween plus 1% BSA and 10% normal goat serum, and incubated with non-stained replicate blots on the rocking platform at room temperature. After 2 hr, immunoblots were briefly rinsed with PBS and then labeled with I<sup>125</sup>-goat anti-human IgE diluted 1:5 in PBS-Tween plus 1% BSA with 0.16 µCi/ml of total activity of I<sup>125</sup> as per directions from the manufacturer (DiaMed, Windham, ME). After gentle agitation for 1 hour at room temperature and washing with PBS, immunoblots were mounted on filter paper and exposed to Kodak BioMax Imaging Film (Eastman Kodak) for 1 to 3 days. Protein images were

scanned with CanoScan D125OU2F (Canon, Inc., Lake Success, NY) and composed using Adobe Photoshop and IN Design software (Adobe Systems, Inc., San Jose, CA).

***IgE binding with serum from soybean allergic person in flour extracts from Control non-GM soybeans and 356043 soybeans - ELISA inhibition analysis.*** For the ELISA inhibition study, 96-well plates were coated with 50 µg of non-transgenic (Jack) soy protein extract. Serum from soy-allergic subjects diluted 1:10 in 2% BSA in PBST was pre-incubated with varying concentrations (0, 50-500,000 ng/ml) of extracts from either the 356043 or Jack non-transgenic soybean for 2 hr at 37°C. The inhibition samples were then added to the PBST-washed plate and incubated for 1 hr at room temperature. Plates were washed and incubated with horseradish peroxidase-labeled goat anti-human IgE, developed with tetramethylbenzidine (TMB) peroxidase substrate, and then read on a microplate reader at 595 nm. The percent inhibition values were calculated by the reduction in binding from wells without inhibitor. Inhibitor doses (nanograms) were based on the total protein determination for each extract and adjusted for the dilution of the sample.

***Acute toxicity of GAT4601 protein.*** The acute toxicity study with purified GAT4601 protein obtained from a heterologous bacterial expression system was conducted at DuPont Haskell Laboratory (Newark, DE). The GAT4601 test substance was a lyophilized powder that was shipped on dry ice from Pioneer Hi-Bred, International, Inc. (Pioneer; Johnston, IA) to DuPont Haskell Laboratory (Newark, DE). The study was conducted in accordance with the requirements for the U.S. EPA FIFRA GLP standards which are compatible with OECD standards (OECD, 2001; U.S. EPA, 1989). Male and female (nulliparous and non-pregnant) Crl:CD1(ICR) mice were received from Charles River Laboratories, Inc. (Raleigh, NC). Mice were housed individually in polycarbonate pans containing bedding and held in climate controlled rooms with temperatures ranging from 18-26°C, humidity ranging from 30-70%, and artificial fluorescent illumination on 12 hr light/dark cycles. Feed (Purina Mills International, PMI® LLC Certified Rodent LabDiet® 5002, Richmond, IN) and tap water were available *ad libitum* throughout the study. GAT4601 protein and control protein (BSA) substances were suspended in deionized water at a concentration of approximately 1000 mg per 5 ml of water. Mice were fasted for 5.5 hr after which a single dose of the GAT4601 test substance was administered to mice (5/sex) via oral gavage at a volume of 10 ml/kg of body

weight for a dose of 2000 mg/kg (OECD, 2001). Control animals (5/sex) were administered a single dose of BSA at a dose of 2000 mg/kg of body weight. An additional group of control animals (5/sex) was dosed by oral gavage with vehicle (i.e., deionized water). All animals were dosed with a volume of 10 ml/kg body weight. Individual body weights were determined at the time of dosing (Day 0) and at different time points (Days 1, 2, 4, 7, and 14) following administration of the test or control substances. Animals were observed for clinical signs of toxicity once during the 30 min following dosing and at least 2 times during the 4 hr interval following administration of test or control substances. Animals were observed for signs of illness, injury, or abnormal behavior twice daily until the end of the in-life portion of the study. Fourteen days following administration of the test or control substances all animals were euthanized with carbon dioxide and subjected to a complete gross pathology examination.

***Repeated dose toxicity assessment of GAT4601 protein.*** The repeated dose dietary toxicity assessment of GAT4601 protein was conducted at Charles River Laboratories (Redfield, AR). Healthy male and female Crl:CD1(ICR) mice from Charles River Laboratories (Portage, Michigan) were group housed in stainless steel cages that conformed to standards set forth in the *Guide for the Care and Use of Laboratory Animals* (NRC, 1996). On Study Day 1, the animals were approximately eight weeks old. The males weighed 28 to 32 grams, and the females weighed 22 to 26 grams. Animals consumed the indicated diets for 27 days after which they were switched to control feed (PMI® Certified Rodent LabDiet® 5002) until scheduled sacrifice on study day 29 relative to the first day that they were exposed to the test diets. Filtered tap water was provided *ad libitum* throughout the study. Environmental controls were set to maintain temperatures of 18-26°C with a relative humidity of 30-70%. A 12 hour light/dark cycle was maintained in the animal room. The animal room was ventilated with at least ten air changes per hour with 100% fresh air (no air recirculation). Twenty males and 20 females were randomly assigned to groups by a computer generated weight-ordered distribution such that group mean body weights did not exceed  $\pm 10\%$  of the overall mean weight for each sex.

***Diet formulation.*** Dietary formulations were prepared by Charles River Laboratories (Redfield, AR), stored at room temperature, and used within ten days of preparation. Diets were prepared on Study Days -

3, 5, 12, and 19 based on the overall average male weights recorded on Study Days -4, 4, 11, and 18.

Calculations were based on a purity correction factor based of GAT4601 protein test substance described above (0.84 mg full length GAT4601 protein/mg of test substance) and on average estimated consumption of 7 g of diet per day per mouse. Control mice consumed PMI 5002 diets. For mice consuming diets containing GAT4601, a premix was prepared by mixing GAT4601 test substance and with 200 g of feed in a Hobart mixer for approximately 5 min. The premix was then combined with the PMI 5002 feed in a twin-shell blender with an intensifier bar and mixed for approximately 10 min.

***Concentration, homogeneity, and stability of GAT4601 protein in diets.*** During the course of test diet formulation, quadruplicate samples of at least 1 g were taken from the top, middle, and bottom layers of each diet preparation for homogeneity analysis and stored frozen at -20°C for subsequent determination of GAT4601 protein concentrations. Additional quadruplicate diet samples of approximately 1 g were collected to assess the stability of GAT4601 protein in the diets. The conditions tested were at least 10 days at room temperature and at least 14 days at -20°C. For room temperature stability, duplicate samples were stored at room temperature for at least 10 days and then frozen (-20°C) prior to analysis. A GAT4601 protein specific ELISA was used to determine the concentration, homogeneity, and stability of GAT4601 protein in rodent diets used in this study. The reference substance for the ELISA analysis was purified GAT4601 protein obtained from a heterologous bacterial expression system. Approximately 60 mg of diet samples were weighed into vials and stored at <-10°C until analysis. Prior to extraction, a chilled (2° to 8°C) steel ball (3/8" diameter) was added to each of the tubes and the samples were extracted with 1.8 ml of cold GEB Buffer, 25 mM HEPES, 25 mM CAPS, 150 mM NaCl, 0.1% Tween-20 (ICI Americas, Inc.), 0.5% PVP-40, 10% Ethylene Glycol, and 5% StabilZyme Select (SurModics, Inc., Eden Prairie, MN), homogenized using a GenoGrinder (BT&C/OPS Diagnostics, Metuchen, NJ) and centrifuged. Supernatants were removed and analyzed for GAT4601 protein concentration using a specific ELISA. Criteria for acceptability were a Relative Standard Deviation (RSD) of  $\leq 20\%$ . For frozen stability, duplicate samples were stored at -20°C for at least 14 days. The homogeneity results obtained from the top, middle, and bottom for Week 1 were averaged and utilized as the initial concentration for each group for

the stability analyses. The criterion for acceptability was a concentration after storage that was within  $\pm$  30% of the initial concentration.

***Clinical observations, body weights, and feed consumption.*** Clinical observations were recorded before the start of the study and once daily thereafter. Body weights were recorded on Study Days -4, 1, 4, 8, 11, 15, 18, 22, and 29. Feed consumption was measured on Study Days 1, 8, 15, 17, and once daily thereafter.

***Ophthalmology.*** Ophthalmic examinations were conducted for all animals pretest and on Study Day 22. For all animals, a short-acting, mydriatic solution was instilled into each eye to facilitate the examination. Ophthalmic examinations included macroscopic examination of the anterior portion of the eye, including adnexa, cornea, anterior chamber, iris, and lens. Examination of the posterior segment was conducted by indirect ophthalmoscopy.

***Hematology.*** Animals were anesthetized with isoflurane on Study Day 29 and blood samples for hematological analysis were obtained from animals via cardiac puncture. Blood samples were evaluated with a Bayer (Siemens) Advia 120 Hematology Analyzer (Norwood, MA) for the following response variables: red blood cell count (RBC), hemoglobin concentration (HGB), hematocrit (HCT), mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), reticulocyte percentage (RET), absolute reticulocyte count (ARET), platelet count (PLT), white blood cell count (WBC), neutrophil count (ANEU), lymphocyte count (ALYM), monocyte count (AMON), eosinophil count (AEOS), basophil count (ABAS), and large unstained cells (ALUC).

***Serum chemistry.*** Blood samples obtained from animals on Study Day 29 were evaluated using an Olympus AU 640e Clinical Chemistry Analyzer (Olympus America, Melville, NY) for the following parameters: aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALKP), gamma-glutamyltransferase (GGT), total bilirubin (BILI), blood urea nitrogen (BUN), creatinine (CREA), cholesterol (CHOL), triglycerides (TRIG), glucose (GLUC), total protein (TP), albumin (ALB),

globulin (GLOB), albumin/globulin ratio (A/G), calcium (CALC), inorganic phosphorus (IPHS), sodium (NA), potassium (K), and chloride (CL).

**Organ weights.** Following sacrifice, a gross necropsy (including all external surfaces and orifices, cranial cavity and external surface of the brain, thoracic, abdominal, and pelvic cavities with their associated organs and tissues) was conducted. Organ weights were recorded for the following organs prior to fixation: adrenal gland, brain, epididymides, heart, kidney, liver, lung, ovary, spleen, testis, and thymus. Paired organs were weighed together.

**Tissue collection, preservation, and histopathology.** The following organs and tissues were examined *in situ*, dissected free, and fixed in 10% neutral buffered formalin (except for the testes which were fixed in Modified Davidson's fixative and the eyes with optic nerve which were fixed in Davidson's fixative): adrenal gland, animal identification, aorta, bone (femur), bone (sternum), bone marrow (sternum), brain (brain stem, cerebellum, cerebrum), cervix, epididymus, esophagus, eye, gallbladder, gross lesions, heart, intestine (cecum, colon, rectum, duodenum, ileum, jejunum [with Peyer's patch if present in the routine section]), kidney, liver, lung, lymph node (mandibular, mesenteric), mammary gland, nerve (optic, sciatic), ovary, pancreas, pituitary gland, prostate gland, salivary gland (mandibular), seminal vesicle, skeletal muscle, skin (abdominal), spinal cord (cervical, lumbar, thoracic), spleen, stomach (nonglandular, glandular), testis, thymus, thyroid gland/parathyroid gland (if present in the routine section), tongue, trachea, urinary bladder, uterus, and vagina. Histopathology was conducted on all tissues from animals in the control and high dose groups and all gross lesions. Fixed tissues were shipped to Charles River Laboratories, Pathology Associates, Maryland (Frederick, MD) for trimming, embedding, and sectioning. Slides were stained with hematoxylin and eosin and evaluated microscopically.

**Statistics.** Statistical analysis of data from the repeated dose toxicity study was conducted on body weight, hematology, clinical chemistry, and organ weight data. Means and standard deviations were calculated for all quantitative data. To determine the appropriate statistical test, each data set was subjected to a statistical decision tree. A minimum of three animals per sex per group per interval was required for statistical

analysis. Treated groups of the same sex were compared to the Control group at common time points. Bartlett's Test was used to compare overall variance among all groups. If the Bartlett's Test had a probability of  $p > 0.05$ , a parametric distribution was assumed and a one-way analysis of variance (ANOVA) was conducted. A significant ANOVA result ( $p \leq 0.05$ ) was followed by multiple comparisons using a *post-hoc* Dunnett's *t*-Test to identify differences between control and individual treatment groups. A non-significant ANOVA result ( $p > 0.05$ ) indicated no statistically significant differences between groups for particular response variable. If Bartlett's Test had a probability of  $p \leq 0.05$ , a non-parametric distribution was assumed and a Kruskal-Wallis test for independent groups was conducted. A significant Kruskal-Wallis Test result ( $p \leq 0.05$ ) was followed by multiple comparisons using a *post-hoc* Dunn's procedure to reveal differences between control and individual treatment groups. A non-significant Kruskal-Wallis Test ( $p > 0.05$ ) indicated no statistically significant differences between the particular response variable.

## References:

- Castle, L. A., Siehl, D. L., Gorton, R., Patten, P. A., Chen, Y. H., Bertain, S., Cho, H. J., Duck, N., Wong, J., Liu, D., and Lassner, M.W. (2004). Discovery and directed evolution of a glyphosate tolerance gene. *Science* **304**, 1151-4.
- FAO/WHO. (2001). Evaluation of allergenicity of genetically modified foods. Report of a Joint FAO/WHO Expert Consultation on Allergenicity of Foods Derived from Biotechnology, January 22-25, 2001. Rome, Italy. [http://www.who.int/foodsafety/publications/biotech/en/ec\\_jan2001.pdf](http://www.who.int/foodsafety/publications/biotech/en/ec_jan2001.pdf)
- Ladics, G. S., Bannon, G. A., Silvanovich, A., and Cressman, R. F. (2007). Comparison of conventional FASTA identity searches with the 80 amino acid sliding window FASTA search for the elucidation of potential identities to known allergens. *Mol. Nutr. Food Res.* **51**, 985-998.
- Metcalf, D. D., Astwood, J. D., Townsend, R., Sampson, H. A., Taylor, S. L., and Fuchs, R. L. (1996). Assessment of the allergenic potential of foods from genetically engineered crop plants. *Crit. Rev. Food Sci. Nutr.* **36S**, S165-186.
- NRC (National Research Council). (1996). Guide for the Care and Use of Laboratory Animals. National Academy Press, Washington, D.C.
- OECD. 2001. Acute oral toxicity – Acute toxic class method (423). [www.oecd.org/dataoecd/17/50/1948370.pdf](http://www.oecd.org/dataoecd/17/50/1948370.pdf).
- OECD. 1995. Guideline for the testing of chemicals. Repeated dose 28-day oral toxicity study in rodents (407). [www.oecd.org/dataoecd/50/41/37477972.pdf](http://www.oecd.org/dataoecd/50/41/37477972.pdf).
- Pearson, W. R. (2000). Flexible sequence similarity searching with the FASTA3 program package. *Meth. Mol. Biol.* **132**, 185-219.

Pearson, W. R., and Lipman, D. J. (1988). Improved tools for biological comparison. *Proc. Natl. Acad. Sci. USA* **85**, 2440-2448.

Thomas, K., Aalbers, M., Bannon, G. A., Bartels, M., Dearman, R. J., Esdaile, D. J., Fu, T. J., Glatt, C. M., Hadfield, N., Hatzos, C., Hefle, S. L., Heylings, J. R., Goodman, R. E., Henry, B., Herouet, C., Holsapple, M., Ladics, G. S., Landry, T. D., MacIntosh, S. C., Rice, E. A., Privalle, L. S., Steiner, H. Y., Teshima, R., Van Ree, R., Woolhiser, M., and Zawodny, J. (2004). A multi-laboratory evaluation of a common in vitro pepsin digestion assay protocol used in assessing the safety of novel proteins. *Regul. Toxicol. Pharmacol.* **39**, 87-98.

## Supplemental Tables

**SUPPLEMENTAL TABLE 1**  
**Mean body weights (g  $\pm$  SD) of mice from acute toxicity study with GAT4601 protein**

|         | Day            |                |                |                |                |                |
|---------|----------------|----------------|----------------|----------------|----------------|----------------|
|         | 0              | 1              | 2              | 4              | 7              | 14             |
| Males   |                |                |                |                |                |                |
| Vehicle | 35.0 $\pm$ 1.2 | 35.6 $\pm$ 1.2 | 35.8 $\pm$ 1.1 | 36.1 $\pm$ 1.0 | 37.1 $\pm$ 0.8 | 38.0 $\pm$ 0.6 |
| BSA     | 34.1 $\pm$ 1.0 | 34.4 $\pm$ 0.7 | 34.6 $\pm$ 0.8 | 34.6 $\pm$ 0.8 | 35.1 $\pm$ 1.2 | 35.8 $\pm$ 1.9 |
| GAT4601 | 33.1 $\pm$ 2.7 | 33.7 $\pm$ 1.9 | 34.3 $\pm$ 1.9 | 34.3 $\pm$ 2.3 | 34.7 $\pm$ 2.0 | 35.2 $\pm$ 1.4 |
| Females |                |                |                |                |                |                |
| Vehicle | 26.5 $\pm$ 1.1 | 26.5 $\pm$ 1.6 | 27.1 $\pm$ 1.1 | 27.0 $\pm$ 1.2 | 27.7 $\pm$ 1.6 | 28.6 $\pm$ 2.1 |
| BSA     | 25.7 $\pm$ 1.8 | 25.5 $\pm$ 1.9 | 25.7 $\pm$ 1.8 | 25.7 $\pm$ 2.6 | 25.4 $\pm$ 2.3 | 27.4 $\pm$ 3.1 |
| GAT4601 | 25.3 $\pm$ 1.0 | 25.6 $\pm$ 1.3 | 25.6 $\pm$ 1.3 | 25.9 $\pm$ 1.7 | 26.0 $\pm$ 2.0 | 26.9 $\pm$ 1.4 |

**SUPPLEMENTAL TABLE 2**  
**Mean hematology values of male mice ( $\pm$ SD) following repeated dose dietary exposure to GAT4601 protein**

|                             | Control<br>n = 5   | Low<br>n = 4 <sup>a</sup> | Middle<br>n = 5    | High<br>n = 5                |
|-----------------------------|--------------------|---------------------------|--------------------|------------------------------|
| RBC (x10 <sup>6</sup> /μL)  | 8.46 $\pm$ 0.22    | 8.63 $\pm$ 0.26           | 8.80 $\pm$ 0.44    | 8.70 $\pm$ 0.55              |
| HGB (g/dL)                  | 12.94 $\pm$ 0.25   | 13.38 $\pm$ 0.25          | 13.42 $\pm$ 0.43   | 13.58 $\pm$ 0.69             |
| HCT (%)                     | 39.86 $\pm$ 0.89   | 39.83 $\pm$ 0.59          | 41.04 $\pm$ 1.47   | 40.74 $\pm$ 1.69             |
| MCV (fL)                    | 47.10 $\pm$ 0.73   | 46.23 $\pm$ 1.90          | 46.70 $\pm$ 1.10   | 46.90 $\pm$ 1.85             |
| MCH (pg)                    | 15.30 $\pm$ 0.25   | 15.50 $\pm$ 0.34          | 15.28 $\pm$ 0.39   | 15.64 $\pm$ 0.69             |
| MCHC (g/dL)                 | 32.50 $\pm$ 0.43   | 33.55 $\pm$ 0.87*         | 32.68 $\pm$ 0.49   | 33.30 $\pm$ 0.49             |
| RET (%)                     | 3.84 $\pm$ 0.27    | 3.43 $\pm$ 0.35           | 3.36 $\pm$ 0.46    | 3.68 $\pm$ 0.43              |
| ARET (x10 <sup>3</sup> /μL) | 325.70 $\pm$ 14.97 | 295.30 $\pm$ 30.33        | 293.12 $\pm$ 35.98 | 317.16 $\pm$ 21.41           |
| PLT (x10 <sup>3</sup> /μL)  | 1197.6 $\pm$ 120.4 | 1223.8 $\pm$ 108.5        | 1098.6 $\pm$ 45.1  | 984.2 $\pm$ 216.4            |
| WBC (x10 <sup>3</sup> /μL)  | 2.62 $\pm$ 0.99    | 2.94 $\pm$ 0.92           | 3.58 $\pm$ 1.02    | 3.77 $\pm$ 0.56              |
| ANEU (x10 <sup>3</sup> /μL) | 0.59 $\pm$ 0.25    | 0.45 $\pm$ 0.17           | 0.47 $\pm$ 0.11    | 0.53 $\pm$ 0.23 <sup>b</sup> |
| ALYM (x10 <sup>3</sup> /μL) | 1.84 $\pm$ 0.80    | 2.26 $\pm$ 0.85           | 2.92 $\pm$ 0.89    | 2.92 $\pm$ 0.55              |
| AMON (x10 <sup>3</sup> /μL) | 0.050 $\pm$ 0.025  | 0.040 $\pm$ 0.026         | 0.036 $\pm$ 0.015  | 0.048 $\pm$ 0.022            |
| AEOS (x10 <sup>3</sup> /μL) | 0.094 $\pm$ 0.025  | 0.120 $\pm$ 0.048         | 0.080 $\pm$ 0.023  | 0.106 $\pm$ 0.064            |
| ABAS (x10 <sup>3</sup> /μL) | 0.012 $\pm$ 0.004  | 0.005 $\pm$ 0.010         | 0.016 $\pm$ 0.019  | 0.018 $\pm$ 0.024            |
| ALUC (x10 <sup>3</sup> /μL) | 0.032 $\pm$ 0.015  | 0.063 $\pm$ 0.042         | 0.064 $\pm$ 0.063  | 0.156 $\pm$ 0.208            |

\* p < 0.05 compared to Control.

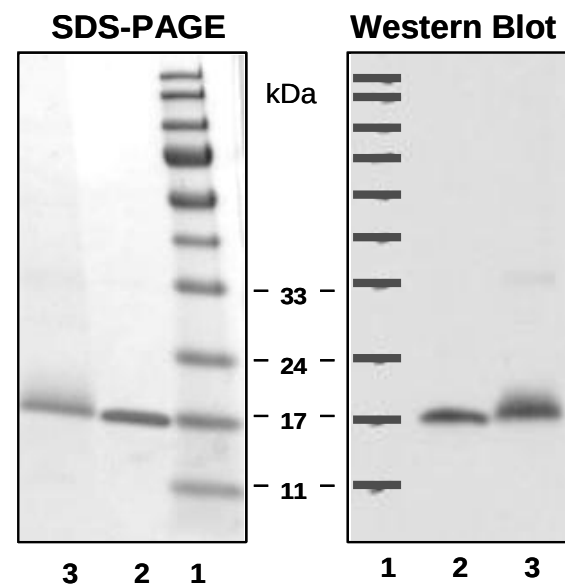
<sup>a</sup> One sample clotted and was excluded from analysis.

<sup>b</sup> n = 4

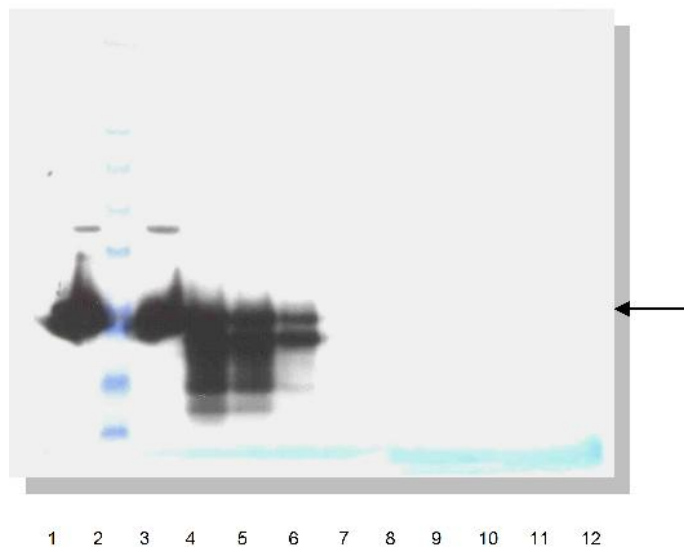
**SUPPLEMENTAL TABLE 3**  
**Mean hematology values of female mice ( $\pm$ SD) following repeated dose dietary exposure to GAT4601 protein**

|                             | Control<br>n = 4 <sup>a</sup> | Low<br>n = 5       | Middle<br>n = 4 <sup>a</sup> | High<br>n = 3 <sup>a</sup> |
|-----------------------------|-------------------------------|--------------------|------------------------------|----------------------------|
| RBC (x10 <sup>6</sup> /μL)  | 8.99 $\pm$ 0.70               | 8.98 $\pm$ 0.61    | 8.92 $\pm$ 0.38              | 8.45 $\pm$ 0.44            |
| HGB (g/dL)                  | 14.50 $\pm$ 0.68              | 14.28 $\pm$ 0.63   | 14.25 $\pm$ 0.82             | 13.87 $\pm$ 0.51           |
| HCT (%)                     | 42.73 $\pm$ 2.37              | 42.56 $\pm$ 2.09   | 42.75 $\pm$ 2.36             | 40.57 $\pm$ 2.08           |
| MCV (fL)                    | 47.65 $\pm$ 2.18              | 47.48 $\pm$ 1.25   | 47.93 $\pm$ 0.92             | 48.03 $\pm$ 0.31           |
| MCH (pg)                    | 16.18 $\pm$ 1.23              | 15.94 $\pm$ 0.56   | 16.00 $\pm$ 0.36             | 16.47 $\pm$ 0.31           |
| MCHC (g/dL)                 | 33.93 $\pm$ 1.15              | 33.56 $\pm$ 0.61   | 33.33 $\pm$ 0.71             | 34.23 $\pm$ 0.58           |
| RET (%)                     | 3.25 $\pm$ 0.85               | 3.18 $\pm$ 0.50    | 3.08 $\pm$ 0.90              | 2.83 $\pm$ 0.55            |
| ARET (x10 <sup>3</sup> /μL) | 291.10 $\pm$ 73.79            | 284.56 $\pm$ 34.98 | 275.43 $\pm$ 80.98           | 235.93 $\pm$ 41.68         |
| PLT (x10 <sup>3</sup> /μL)  | 1134.5 $\pm$ 144.0            | 1071.4 $\pm$ 118.4 | 1117.0 $\pm$ 125.4           | 927.7 $\pm$ 90.8           |
| WBC (x10 <sup>3</sup> /μL)  | 5.72 $\pm$ 3.01               | 3.71 $\pm$ 1.60    | 4.09 $\pm$ 2.15              | 4.77 $\pm$ 2.37            |
| ANEU (x10 <sup>3</sup> /μL) | 0.78 $\pm$ 0.52               | 0.48 $\pm$ 0.24    | 0.70 $\pm$ 0.22              | 0.96 $\pm$ 0.32            |
| ALYM (x10 <sup>3</sup> /μL) | 4.43 $\pm$ 2.20               | 2.73 $\pm$ 1.21    | 2.89 $\pm$ 1.53              | 3.52 $\pm$ 2.18            |
| AMON (x10 <sup>3</sup> /μL) | 0.058 $\pm$ 0.039             | 0.040 $\pm$ 0.024  | 0.075 $\pm$ 0.051            | 0.050 $\pm$ 0.020          |
| AEOS (x10 <sup>3</sup> /μL) | 0.23 $\pm$ 0.19               | 0.12 $\pm$ 0.091   | 0.14 $\pm$ 0.043             | 0.15 $\pm$ 0.015           |
| ABAS (x10 <sup>3</sup> /μL) | 0.025 $\pm$ 0.019             | 0.044 $\pm$ 0.044  | 0.028 $\pm$ 0.029            | 0.013 $\pm$ 0.006          |
| ALUC (x10 <sup>3</sup> /μL) | 0.21 $\pm$ 0.26               | 0.29 $\pm$ 0.26    | 0.27 $\pm$ 0.37              | 0.08 $\pm$ 0.02            |

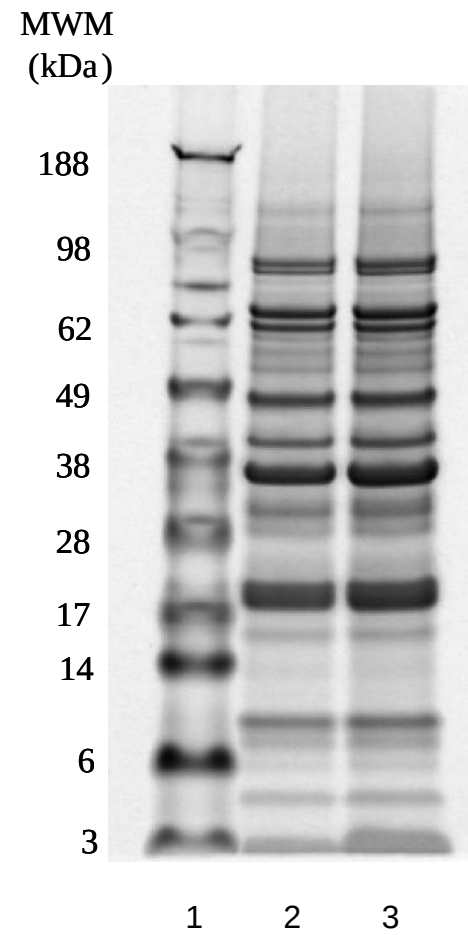
<sup>a</sup> Clotted samples were excluded from analysis.



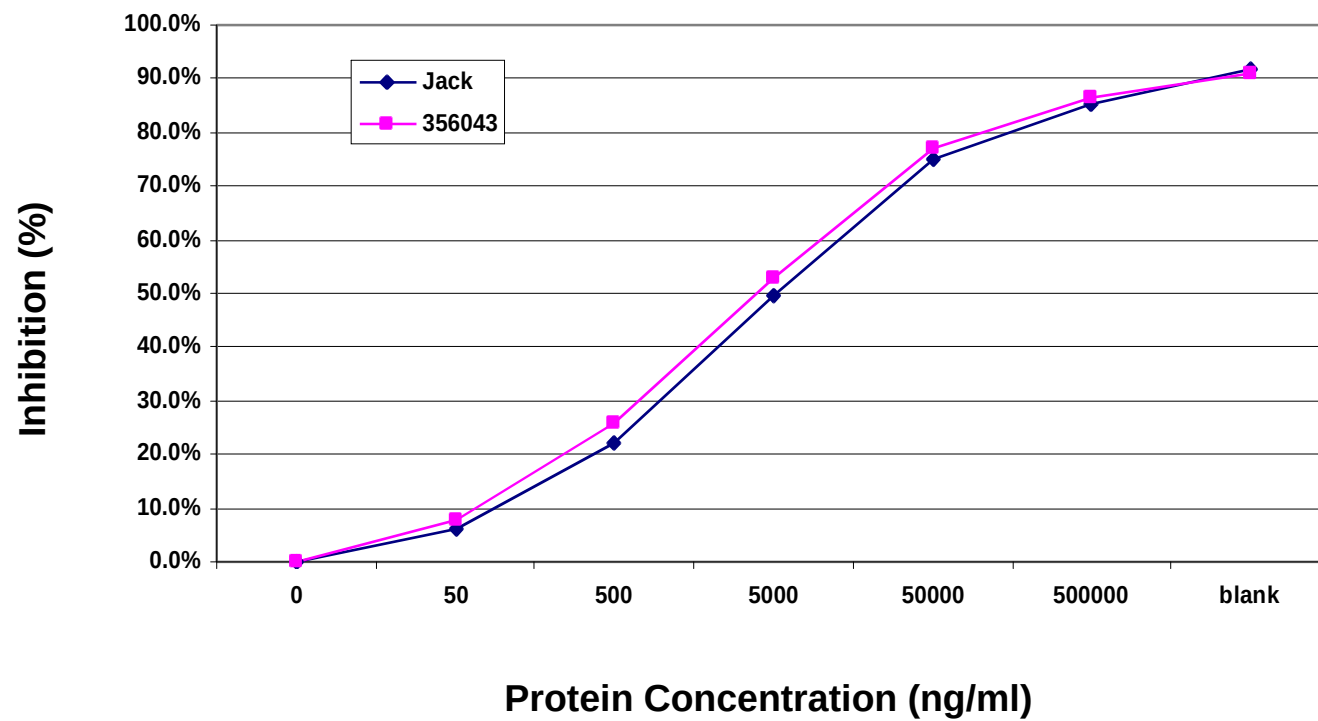
**Supplemental Figure 1**



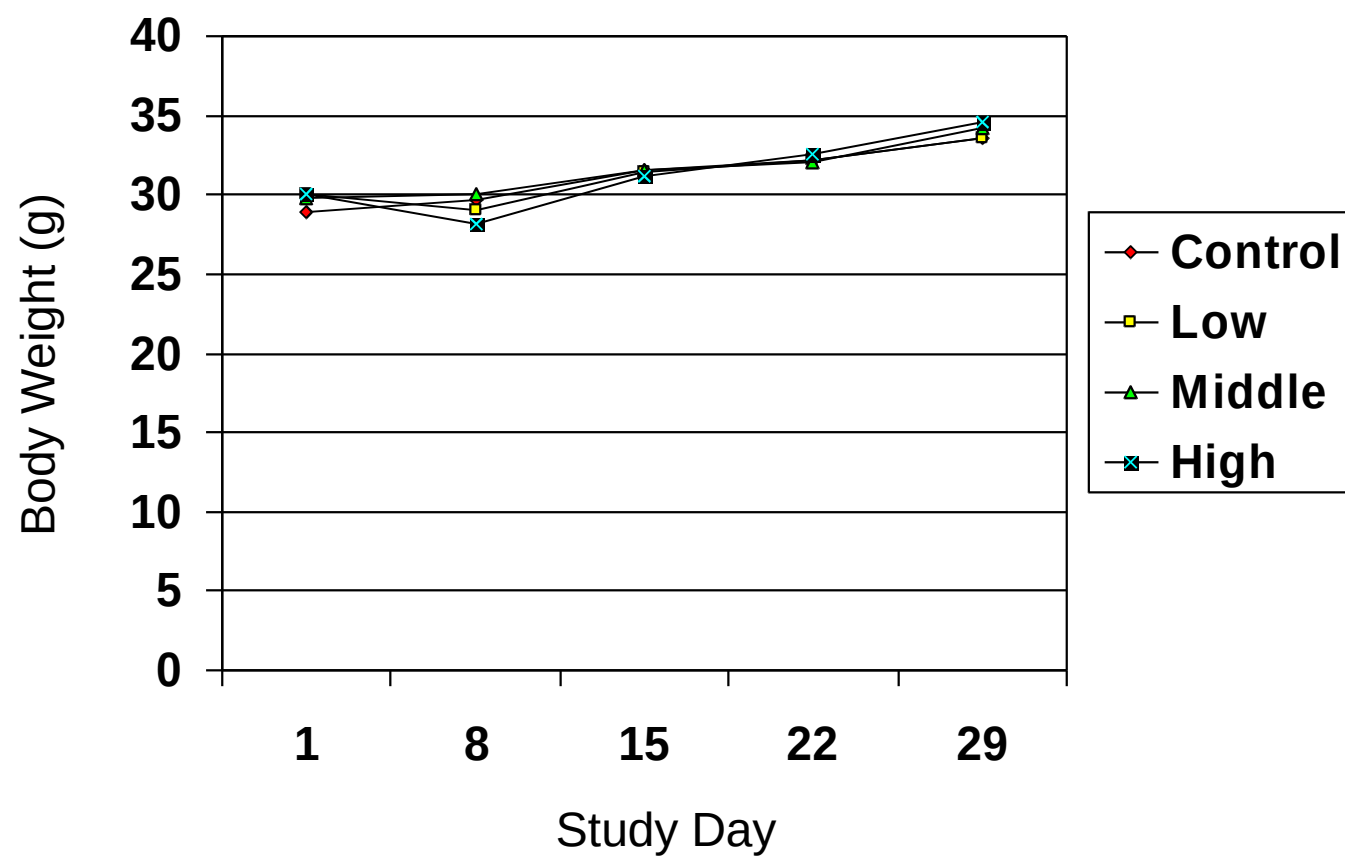
**Supplemental Figure 2**



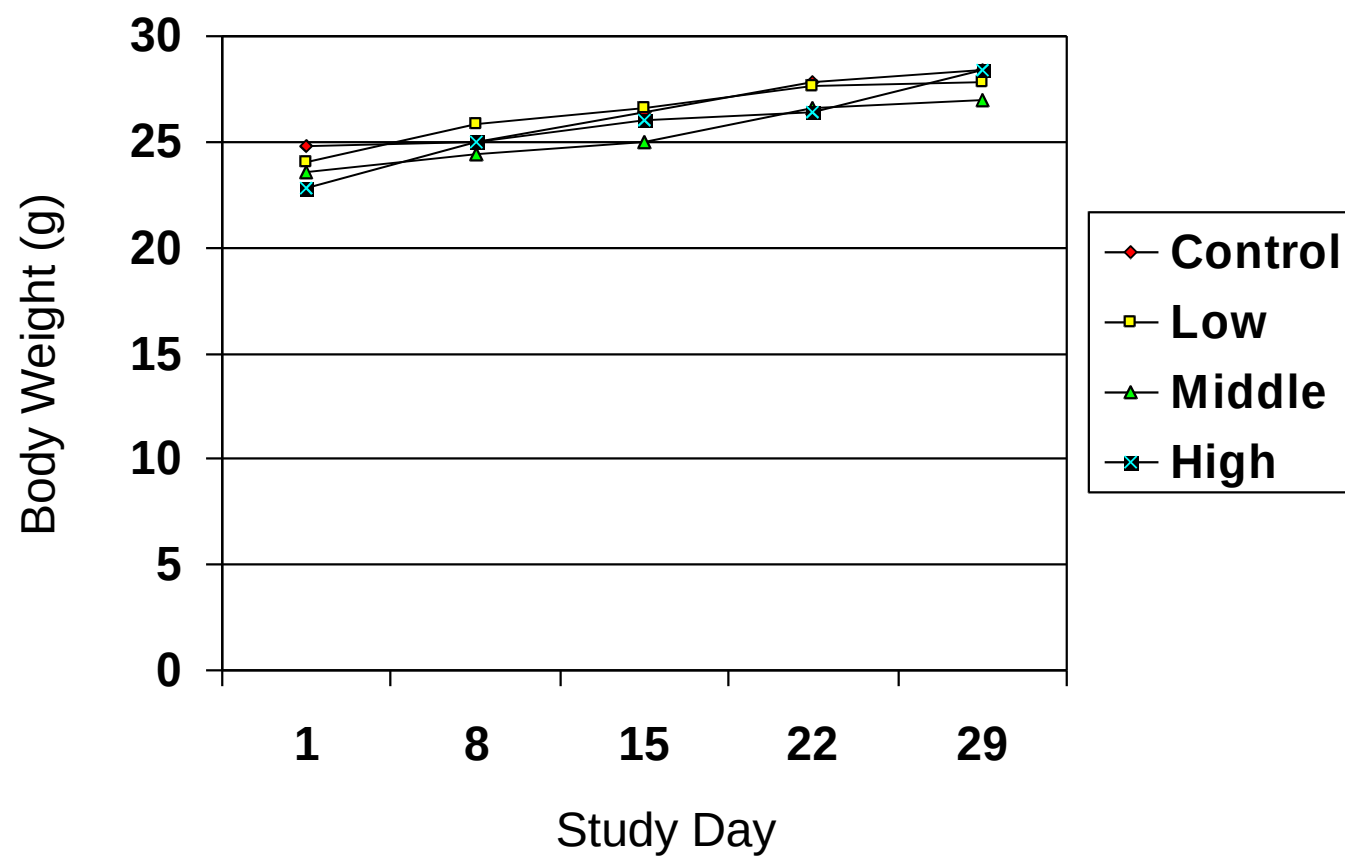
**Supplemental Figure 3**



## Supplemental Figure 4



**Supplemental Figure 5**



**Supplemental Figure 6**