Plant Cell Rep
DOI 10.1007/s00299-011-1044-8

ORIGINAL PAPER

Recombinant expression of homodimeric 660 kDa human thyroglobulin in soybean seeds: an alternative source of human thyroglobulin

Rebecca Powell · Laura C. Hudson · Kevin C. Lambirth · Diane Luth · Kan Wang · Kenneth L. Bost · Kenneth J. Piller

Received: 14 October 2010 / Revised: 4 February 2011 / Accepted: 20 February 2011

Abstract Soybean seeds possess many qualities that make them ideal targets for the production of recombinant proteins. However, one quality often overlooked is their ability to stockpile large amounts of complex storage proteins. Because of this characteristic, we hypothesized that soybean seeds would support recombinant expression of large and complex proteins that are currently difficult or impossible to express using traditional plant and non-plant-based host systems. To test this hypothesis, we transformed soybeans with a synthetic gene encoding human thyroglobulin (hTG)—a 660 kDa homodimeric protein that is widely used in the diagnostic industry for screening and detection of thyroid disease. In the absence of a recombinant system that can produce recombinant hTG, research and diagnostic grade hTG continues to be purified from cadaver and surgically removed thyroid tissue. These less-than-ideal tissue sources lack uniform glycosylation and iodination and therefore introduce variability when purified hTG is used in sensitive ELISA screens. In this study, we report the successful expression of recombinant hTG in soybean seeds. Authenticity of the soy-derived protein was demonstrated using commercial ELISA kits developed specifically for the detection of hTG in patient sera. Western analyses and gel filtration chromatography demonstrated that recombinant hTG and thyroid-purified hTG are biologically similar with respect to size, mass, charge and subunit interaction. The recombinant protein was stable over three generations and accumulated to ~1.5% of total soluble seed protein. These results support our hypothesis that soybeans represent a practical alternative to traditional host systems for the expression of large and complex proteins.

Keywords Recombinant expression · Transgenic soybean · Human thyroglobulin · Soybean seeds · Targeted expression

Communicated by D. Somers.

R. Powell · K. L. Bost · K. J. Piller (✉)
Department of Biology, University of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, NC 28223, USA
e-mail: kjpiller@uncc.edu

L. C. Hudson · K. C. Lambirth · K. L. Bost · K. J. Piller
SoyMeds, Inc., Davidson, NC 28036, USA

D. Luth · K. Wang
Department of Agronomy, Iowa State University, Ames, IA 50011-1010, USA

Introduction

Transgenic plants have been increasing in popularity as bioreactors for the expression of recombinant proteins due in large part to their ability to produce safe, cost-effective and pathogen-free protein-based products (Giddings et al. 2000; Daniell et al. 2009). A number of antigens and pharmaceuticals have already been successfully expressed in plant systems such as tobacco, maize, rice, carrot and tomato (Daniell et al. 2009). Of the many proteins expressed in these systems, some have been approved for human use in Europe and Canada, while others are currently in clinical trials and under review for use in humans (Karg and Kallio 2009). Although these plant host systems exhibit tremendous potential, they also present several challenges that must be considered when choosing a system for recombinant protein expression. For example, many of the plant-based host systems utilized to date have relatively low endogenous protein concentrations at the cellular and subcellular locations where recombinant

Published online: 08 March 2011

proteins are intended to accumulate (Daniell et al. 2009). Typically, expression systems with low levels of endogenous protein require an increase in biomass to help compensate for the low protein content. An increase in biomass usually correlates with an increased footprint for plant growth which can decrease the cost-effectiveness of the platform. While some protein expression issues may be circumvented by planting on increased acreage in open fields such outdoor plantings pose a separate set of risks if containment of the genetically modified host is not meticulously controlled.

Our laboratory has been developing *Glycine max* as a platform for the expression of recombinant proteins. Soybean seeds contain ~40% protein by dry weight (Liu 1999) and represent one of the richest natural sources of protein known. The protein-rich environment of soybeans, coupled with stable transgenic protein accumulations representing 2–4% of total soluble seed protein, underscore the practicality and cost-effectiveness of such a platform. However, one additional characteristic of soybeans often overlooked is their ability to stockpile large amounts of complex storage proteins. While the general function of seeds is to house metabolic and storage proteins required for germination and early seedling growth, soybeans are unique with respect to the level of storage protein that accumulates during seed maturation. This is most evident with conglycinin and glycinin proteins that assemble into 7S and 11S protein complexes, respectively. Both families of proteins are mature with respect to post-translational modification and are capable of forming complexes that comprise 65–80% of the total soybean seed protein which is equivalent to roughly a third of the total weight of a soybean. Such levels of protein are unprecedented for an expression system. This intriguing property led us to question whether the soybean seed environment might represent an ideal location for the production and accumulation of complex proteins that are otherwise difficult or impossible to express using traditional expression systems. If true, then soybeans could be used to generate novel proteins with a variety of applications, including basic science research, diagnosis of disease and possible treatment of disorders and diseases that currently rely on unavailable and/or cost-prohibitive, protein-based reagents.

Human thyroglobulin (hTG) is an important reagent used in the diagnostics industry for the screening and diagnosis of thyroid disease. Functional hTG is a 660-kDa glycoprotein composed of two identical 330 kDa monomers (Marino and McCluskey 2000). hTG is produced in thyrocytes and functions as a hormone storage protein that is later metabolized to produce the thyroid hormones T3 and T4 (Deshpande and Venkatesh 1999). hTG undergoes post translational glycosylation and iodination in vivo, which contributes to its implication as an auto-antigen in hypothyroid disease (Champion et al. 1987; Yang et al. 1996). While glycosylation and disulfide bond formation are necessary for proper folding and final tertiary structure, iodination does not appear to be required for structural integrity (Champion et al. 1987; Dupuy et al. 1988, 1991). Because of its large size and specific requirements for modification and folding, the production of recombinant hTG in traditional expression systems has not been possible. As a result, the only commercially available source of purified hTG is from cadaver or surgically removed thyroid tissue. Unfortunately, the hTG populations in these donor tissues are heterogenous with respect to levels of glycosylation and iodination creating issues with its use in the diagnostics industry. Furthermore, the batch-to-batch variation from different manufacturers also impacts standard ELISA outputs used to screen, detect and monitor thyroid disease (Krahn and Dembinski 2009).

The availability of a universal hTG standard could help remedy existing problems with commercially purified hTG, but to date no traditional expression system has been shown to support recombinant production of this protein. *E. coli* is arguably the most widely understood and simplest recombinant protein expression system to manipulate and has produced commercially relevant proteins such as human insulin and other therapeutic enzymes (Demain and Vaishnav 2009), but would not support expression of hTG due to size and protein modification limitations of the system. Yeast has been utilized as a recombinant protein expression system since it is capable of disulfide bond formation and post-translational modifications; however, variations in glycosylation have been an obstacle in certain species (Gerngross 2004) and have led to diminished yields of recombinant protein (Demain and Vaishnav 2009). Baculovirus and mammalian cell systems have each been used for recombinant protein production since they address limitations of the *E. coli* and yeast systems, but both pose technical hurdles (e.g. maintenance of cell cultures) and do not offer a cost advantage to current purification of hTG from human tissue. To date, there are no reports of hTG produced in any of these systems.

Given that hTG is a large and complex protein, we questioned whether it could be stably expressed within an environment that has evolved to house abundant levels of large protein complexes. We report here that functional hTG was stably expressed in soybean seeds and exhibited molecular, biochemical and immunological properties that were analogous to those exhibited by commercially purified hTG. Soy-derived hTG thus represents a novel source and alternative to hTG purified from human tissue. The expression of hTG in soybeans represents one of the largest functional proteins to be expressed in any plant system underscoring the practicality of this host for the production of novel recombinant proteins that may otherwise be recalcitrant to expression in traditional host systems.

Materials and methods

Vector design and construction

The 8.3 kb human thyroglobulin gene containing sequences encoding an N-terminal signal peptide sequence and C-terminal histidine tag, along with 5′ and 3′ terminal restriction endonuclease sites was optimized for soybean codon usage, and synthesized by GeneArt, Inc. (Burlingame, CA, USA). Following digestion with *Nco*I and *Xba*I, the synthetic gene was isolated from an agarose gel and ligated with linearized pPTN200, a derivative of pPZP202 (Hajdukiewicz et al. 1994) that was generously provided by Dr. Tom Clemente. The resulting product, pPTN-hTG, was confirmed by multiple restriction digestion analyses and integrity of the encoded sequence was confirmed by double strand sequencing.

Soybean transformation

Soybean transformation using the *Agrobacterium*-mediated half-seed method was performed as described (Paz et al. 2006). Briefly, half-seed explants (*G. max* cv Williams 82) were dissected and inoculated with *Agrobacterium* suspension culture (strain EHA101 carrying various binary vectors). The inoculated explants were placed adaxial side down on cocultivation medium at 24°C and under 18:6 photo period for 3–5 days. After cocultivation, explants were cultured for shoot induction and elongation under glufosinate selection (8 mg/L) for 8–12 weeks. Herbicide resistant shoots were harvested, elongated and rooted. Acclimated plantlets were transferred to soil and grown to maturity in the greenhouse. Transformation resulted in a total of five independent glufosinate-resistant events.

Glufosinate screens

To monitor for expression of the herbicide selectable marker, T_1 and T_2 plants were sprayed with Ignite 280 SL herbicide (Bayer CropScience, RTP, NC, USA) at a concentration of 80 mg/L for a total of three sprays every 2 days. Plants with visible chlorosis were scored as negative for resistance to the herbicide and discarded while positive plants were taken to maturity. Plants known to be resistant to phosphinothricin, as well as wild-type plants, were included as controls for spray concentration and application.

Preparation of genomic DNA and PCR

Genomic DNA was prepared from cotyledon tissue using the Maxwell 16 Instrument and Maxwell Tissue DNA Purification Kit (Promega, Madison, WI). PCR reactions were carried out using GoTaq Flexi DNA polymerase (Promega Corp., Madison, WI) with primers: hTG forward (5′-GCTCAACCACTTAGACCATGCGA-3′), hTG reverse (5′-TCAGCGCAGTGGCAATATCCTG-3′), vsp forward (5′-GCTTCCACACATGGGAGCAG-3′) and vsp reverse (5′-CCTCTGTGGTCTCCAAGCAG-3′). Following an initial denaturation cycle (5 min at 94°C) the reactions were subjected to 38 cycles comprised of denaturation (30 s at 94°C), annealing (45 s at 58°C) and extension (60 s at 72°C). PCR products were visualized in 1.0% agarose gels stained with ethidium bromide.

Seed protein extraction

Seed chips (~10 mg of cotyledon tissue) were resuspended in 150 µL of phosphate-buffered saline (PBS) and sonicated for 30 s using a Vibra-Cell ultrasonic processor (Newton, CT, USA). Samples were clarified from soluble debris by centrifugation at 16,000*g* at 4°C. Total soluble protein was quantified with the Bradford Reagent using bovine serum albumin (BSA) as a standard.

Western blot analysis

Protein extracts from transgenic and wild-type seeds were run in 5% native polyacrylamide gels for approximately 2 h at 110 V. Unless noted, SDS sample buffer did not contain β-mercaptoethanol or SDS, and samples were not boiled prior to loading onto the gel. Purified hTG (EMD Chemicals, Gibbstown, NJ, USA) was included as a standard. Following electrophoresis, gels were equilibrated in 1× *N*-cyclohexyl-3-aminopropanesulfonic acid buffer (pH 11) with 10% methanol for 10 min and transferred to Immobilon-P membrane (Millipore, Billerica, MA, USA). Membranes were blocked overnight with 5% nonfat milk in PBS solution at 4°C, incubated with rabbit anti-hTG polyclonal antibody (Gene Tex Inc., Irvine, CA, USA) for 3 h at 23°C and washed three times (10 min each) with PBS containing 0.05% Tween. Membranes were then incubated with goat anti-rabbit HRP-conjugated IgG (Santa Cruz Biotechnology, Santa Cruz, CA, USA) for 30 min at 23°C and washed as described above. Detection was carried out using the SuperSignal West Pico substrate kit (Thermo Scientific, Rockford, IL, USA).

Confocal microscopy

Whole seed tissue was imbibed for 16 h in 1× PBS and the seed coat was removed. Tissue was fixed as described previously by our laboratory (Piller et al. 2005; Oakes et al. 2009). Briefly, sections were permeabilized with 1× PBS containing 0.2% Tween for 10 min, and nonspecific binding was blocked by incubation with 1× PBS supplemented

with 3% BSA for 4 h at 23°C. Tissue was incubated with rabbit anti-hTG serum (1:20 dilution) for 16 h at 4°C followed by incubation with an AlexaFluor 594 goat anti-rabbit IgG-HRP-conjugated secondary antibody (1:200 dilution) for 1 h at 23°C. Finally, tissue was incubated with 4,6-diamidino-2-phenylindole (DAPI; 1:500 dilution) for 5 min. Cover slips were added to the sections using Gel/Mount aqueous mounting media. Images were collected with a LSM 710 Spectral Confocor 3 Confocal Microscope (Carl Zeiss, Inc.) using a 40× objective and a 405-nm laser to visualize DAPI stained nuclei along with a 561-nm laser to collect emitted fluorescence from the Alexafluor 594 antibody. Stacks of images (30 optical sections, 17 nm apart) were collected in the Z plane of the specimens and projected to form a single image. To improve clarity and reproduction quality, image colors were proportionally enhanced using the ZEN 2009 Light Edition software.

Quantification of recombinant protein in seed extracts

For western quantification, known amounts of commercially purified hTG protein and crude seed-extracted protein (line 77-5) were incubated with SDS-sample buffer lacking β-mercaptoethanol and electrophoresed in 5% native polyacrylamide gels. Western blots were performed as described above and X-ray films of the resulting blots were scanned for densitometric analysis. Integrated density was measured using ImageJ software (Abramoff et al. 2004). The image was inverted and background pixel values were subtracted. A standard curve was plotted using integrated density values and known amounts of purified hTG protein which was used to determine an absolute value of hTG in seed samples. For ELISA quantification, known amounts of hTG (0.01–10 ng) and crude seed-extracted protein (10-fold dilutions over four orders of magnitude) were coated onto ELISA plates and processed as described above. Absorbance values from the known concentrations of hTG were used to generate a curve, and the concentration of hTG in seed extracts was determined by extrapolation of hTG concentration for those samples with absorbance values falling within the linear range of the curve. Absolute values were converted to a percentage of total protein.

Gel filtration chromatography

A Sephacryl S-300 HR gel filtration column (bed height 72 cm) was calibrated by determining the peak elution volumes (absorbance at 254 nm, BioLogic LP, BIO-RAD, Inc.) of a set of molecular weight protein standards. Crude, soluble protein was isolated from hTG-positive seeds, applied to the gel filtration column and eluted fractions were collected. Similarly, human thyroid-purified thyroglobulin was applied to the same column and eluted fractions were also collected. Eluted fractions were subjected to ELISA (Orgentec) to detect the presence of immunoreactive thyroglobulin in each fraction.

ELISA

Three different types of capture ELISAs were performed in this study. The first two utilized commercial kits that contained either polyclonal antibody pairs (Orgentec Immunometric Enzyme Immunoassay kit; Orgentec, Germany) or monoclonal antibody pairs (Kronus META-Tg Serum Thyroglobulin assay kit; Kronus, Boise, ID, USA) for capture and detection of hTG. ELISAs using Orgentec and Kronus kits were performed according to the specific manufacturer's directions. A third sandwich-based ELISA was developed in house and utilized a monoclonal antibody for capture and a polyclonal antibody for detection. Briefly, 500 ng of capture antibody (GTX21984, GeneTex, Irvine, CA, USA) was coated onto ELISA plates by incubation at 4°C for 16 h. Unbound antibody was washed with PBS and nonspecific binding sites were blocked by incubation with 1% BSA in PBS for 1 h at 23°C. Soy protein samples and the hTG standard were then loaded onto plates and allowed to complex with the bound antibody for 2 h at 23°C. Unbound products were washed, and a rabbit polyclonal detection antibody (GTX73492, GeneTex, Irvine, CA, USA) was allowed to bind to the antigen for 2 h at 23°C. The secondary antibody was subsequently detected using a goat anti-rabbit IgG-HRP antibody (sc2004, Santa Cruz Biotechnology, Santa Cruz, CA, USA) by incubation for 1 h at 23°C. The antibody–antigen complexes were incubated with TMB substrate and colorimetric reactions were stopped by the addition of 0.6 M sulfuric acid. Absorbance values were read at 450 nm.

Results

Transformation and molecular screening of T_1 events

A synthetic hTG gene was engineered for expression in soybeans. This synthetic gene contained a native signal sequence and an open reading frame encoding the native hTG amino acid sequence. The synthetic gene was also codon-optimized for expression in soybean and the GC content was adjusted to resemble that found in plant systems. The binary vector, pPTN-hTG (Fig. 1a), contains synthetic hTG cloned downstream of the soybean β-conglycinin promoter. *Agrobacterium*-mediated transformation resulted in five independent transformation events designated 77-3, 77-4, 77-5, 77-7 and 77-12. Phenotypically, T_0 parent plants, as well as T_1–T_3 progeny plants, all appeared similar to wild-type nontransgenic control plants with

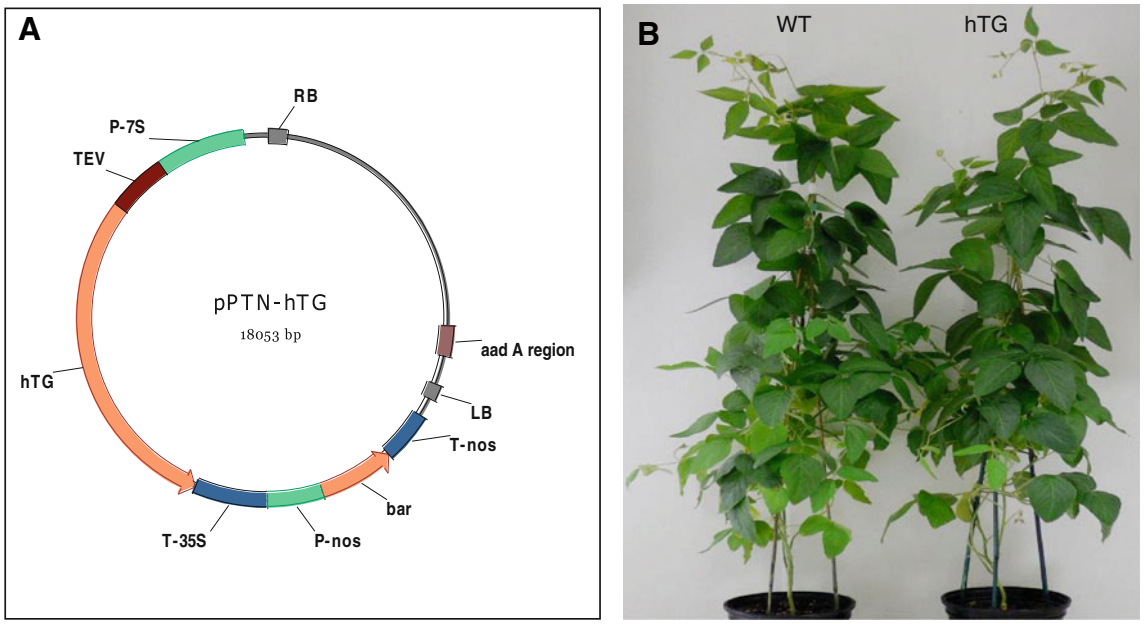

Fig. 1 Gene design and soybean transformation. **a** Binary vector pPTN-hTG used for *Agrobacterium*-mediated transformation. *P-7S* soybean *β*-conglycinin promoter, *TEV* tobacco etch virus translational enhancer element, *hTG* synthetic human thyroglobulin gene, *T-35S* cauliflower mosaic virus terminator element, *P-nos* nopaline synthase promoter, *Bar* phosphinothricin acetyltransferase gene for plant selection, *T-nos* nopaline synthase terminator element, *RB* right border sequence, *LB* left border sequence; *aad A region* streptomycin resistance gene for bacterial selection. **b** 60-day old transgenic (line 77-5) and control (WT) plants

respect to leaf color, growth habit and relative seed yield (Fig. 1b). To screen for the presence of the hTG transgene, six individual T_1 seeds from each parent line were subjected to a duplex PCR using genomic DNA isolated from seed chips harvested prior to planting and germination. The duplex PCR contained primer sets that amplified a 325-bp vegetative storage protein gene fragment (internal control) as well as a 650-bp fragment diagnostic of hTG in transgenic progeny. The PCR results (Fig. 2, top panels) verified that all five of the regenerated lines were transgenic for the hTG gene. Western analyses were performed to characterize the profile of transgenic seed protein. Because of the large size of hTG, crude seed extracts were separated in 5% native polyacrylamide gels. Since the predicted p*I* of hTG is ~5.5, native gel electrophoresis was carried out at pH 8.8 to ensure that the recombinant protein was deprotonated and would migrate into the gels. Commercially purified hTG protein was included as a standard for comparison with soy-derived hTG. It should be noted that while unmodified forms of hTG have a molecular mass of ~302 kDa, commercial preparations of hTG isolated from human thyroids are glycosylated and iodinated and therefore have a molecular mass of ~330 kDa for the monomeric form of protein and ~660 kDa for the dimeric form of protein. Transgenic progeny derived from lines 77-3, 77-5, 77-7 and 77-12 accumulated immunoreactive protein that was detected by rabbit sera containing anti-hTG polyclonal antibodies (Fig. 2, bottom panels). Extracts derived from nontransgenic seeds were used as a negative control and demonstrated specificity of the antibodies for hTG epitopes. Notably, the migration of immunoreactive proteins in crude transgenic seed extracts was similar to that of the 330-kDa monomers and 660-kDa dimers in the commercially purified standard suggesting that these proteins were biologically similar with respect to overall size, charge and other modifications. The slight migration difference observed between the commercial standard and soy immunoreactive protein reflects the absence of iodine residues in the seed proteins, as soybean lacks the enzymes required for protein iodination. In general, there was a direct correlation between the presence of the transgene and expression of hTG, with the exception of line 77-3 and 77-4 in which several of the PCR-positive progeny did not accumulate immunoreactive protein. Progeny derived from events 77-5 and 77-12 consistently expressed the greatest levels of immunoreactive protein and were earmarked as lead events for propagation to second and third generations. Progeny derived from event 77-4 accumulated no detectable hTG and were deprioritized.

Soybean-derived hTG remains intracellular

Although the synthetic hTG gene was designed to mimic the human gene, it was unclear whether the soybean-derived transgenic protein would be secreted into the apoplastic spaces or retained within the cell. Immunohistochemistry

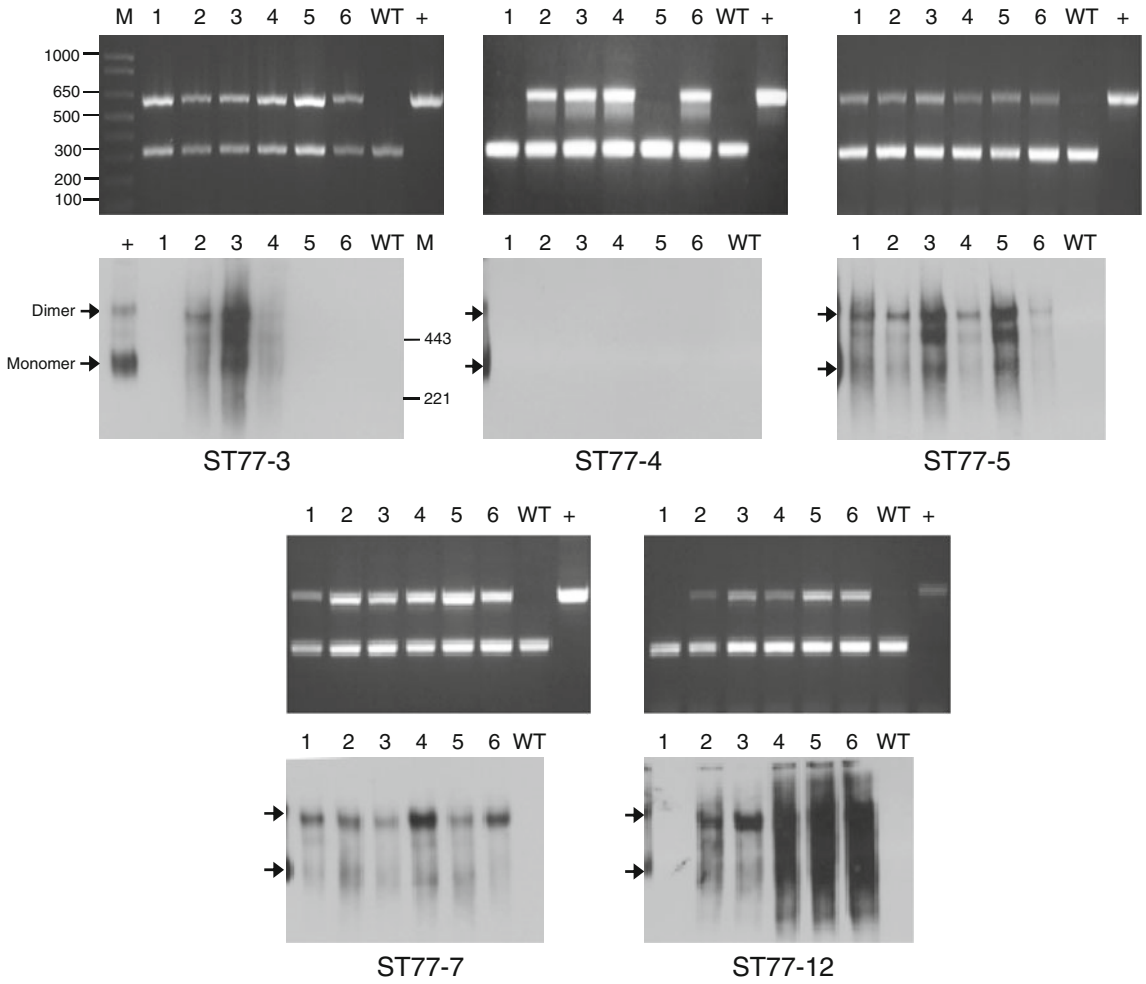

Fig. 2 Molecular characterization of T_1 progeny derived from initial transformation events. *Top panels* agarose gels (1%) showing results from duplex PCR used to detect vegetative storage protein (vsp, internal control) and hTG sequences in genomic DNA isolated from T_1 seed chips. Amplified PCR products diagnostic of vsp and hTG migrate at 325 and 659 bp, respectively. Genomic DNA from control reactions are designated "WT" (wild type, nontransgenic) and "+" (plasmid pPTN-hTG). The designations for molecular weight markers are shown in base pairs. *Bottom panels* western analyses of T_1 crude seed proteins separated in 5% native polyacrylamide gels. The locations of 400 and 250 kDa molecular mass standards, as well as the monomeric (M) and dimeric (D) forms of hTG are indicated

was carried out on cotyledon tissue using fluorescently labeled antibodies. Confocal images (Fig. 3) show that immunoreactive protein was not secreted into apoplastic spaces but instead remained intracellular and localized to cell membranes. Very little transgenic protein was found throughout the cytoplasm, and DAPI staining of nuclear material showed that transgenic protein also appeared to be excluded from the nucleoplasm. Fluorescence was not observed in control (nontransgenic) tissues prepared using identical conditions.

Soybean-derived hTG is detected using commercially available ELISA kits

The Orgentec and Kronus hTG ELISA kits are widely used in the diagnostics industry to measure levels of hTG in patient sera. Both kits utilize pairs of either polyclonal (Orgentec) or monoclonal (Kronus) anti-hTG antibodies in a "sandwich" ELISA format and were therefore used to evaluate recombinant protein expressed in seeds. Crude seed extracts were prepared from six different T_1 progeny derived from each of the five independent transformation events. Absorbance results from the Organtec assay identified seed-specific immunoreactive hTG in cotyledon chips derived from all lines except for line 77-4 (Fig. 4). These results are consistent with previous western results (Fig. 2). The Orgentec kit utilizes polyclonal anti-hTG antibodies to capture and detect hTG, and such antibodies likely bind to both linear and conformational epitopes along the entire length of the thyroglobulin molecule. Thus, the immunoreactivity observed in seed extracts suggested that soy-derived hTG epitopes were also intact. The Kronus

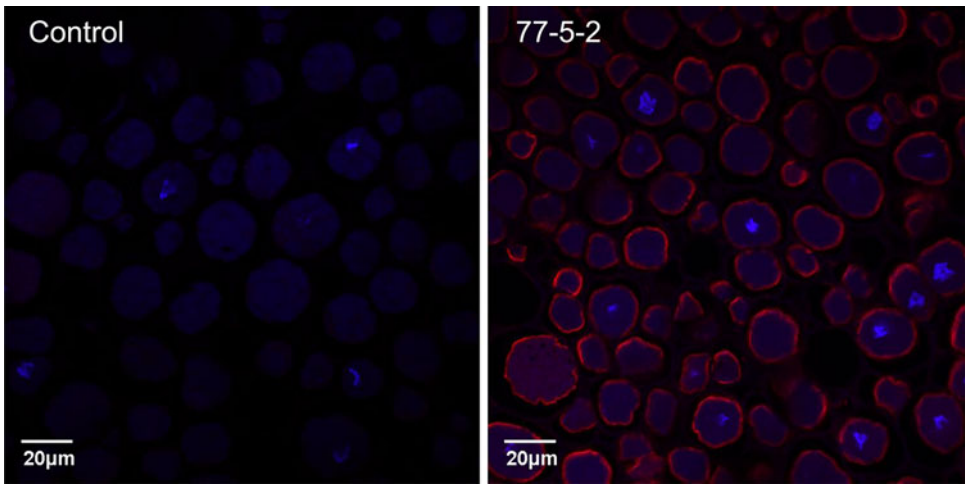

Fig. 3 Confocal microscopy for visualization of hTG in transgenic T_1 cotyledon tissue. The *Blue color* represents DAPI staining while *red color* represents fluorescence from an AlexaFluor secondary antibody recognizing hTG–antibody complexes

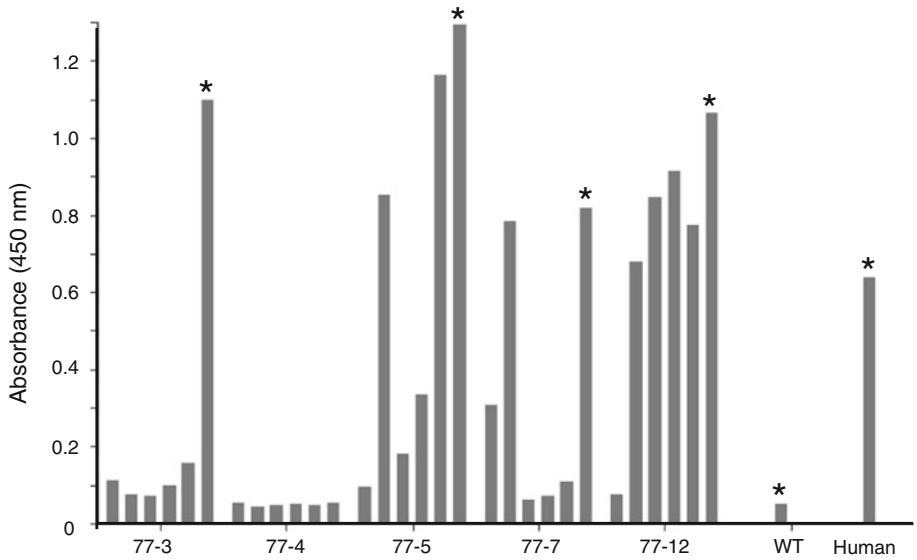

Fig. 4 Orgentec ELISA used for detection of hTG in T_1 seed extracts. Soluble protein from nontransgenic seeds (WT) and commercially purified hTG (human) were used as negative and positive controls, respectively. Values shown represent absorbance values (450 nm). Extracts derived from progeny denoted with an *asterisk* were further characterized in a Kronus assay (see Fig. 5)

ELISA kit allows for a more stringent evaluation of soy-derived hTG since it utilizes two different monoclonal antibodies for capture and detection and can therefore simultaneously recognize two different conformational determinants on the thyroglobulin protein. For the Kronus ELISA screen, one T_1 progeny seed extract was chosen from each positive line (identified by an *asterisk* in Fig. 4), and the assay was carried out using multiple dilutions to ensure that outputs would fall within the linear range of the standard. The ELISA absorbances from two of these dilutions are shown (Fig. 5). The Kronus assay detected seed-specific immunoreactive hTG in each of the crude transgenic seed extracts tested. The fact that separate monoclonal antibodies reacted with the soy-derived transgenic protein, along with the fact that two separate commercial kits detected seed-specific immunoreactive proteins, provided further support for the authenticity of recombinant hTG protein in transgenic seed extracts.

Expression and accumulation of soy-derived hTG is stable over generations

To examine stability of soy-derived hTG over multiple generations, T1 plants from lines 77-5 and 77-12 were taken to maturity and T_2 seeds were screened for hTG expression (data not shown). T_2 events expressing the greatest levels of hTG were taken to maturity and T_3 seed protein extracts were subjected to native polyacrylamide gel electrophoresis and western analysis. Seed extracts derived from original parent lines 77-5 and 77-12 all accumulated immunoreactive hTG protein (Fig. 6a). As observed previously (Fig. 2), the immunoreactive profile of soy-derived hTG was similar to that of commercially purified hTG standard suggesting that both proteins are similar. It is worthwhile to note that while commercial preparations of hTG contained predominantly monomeric protein, soybean extracts contained predominantly dimeric

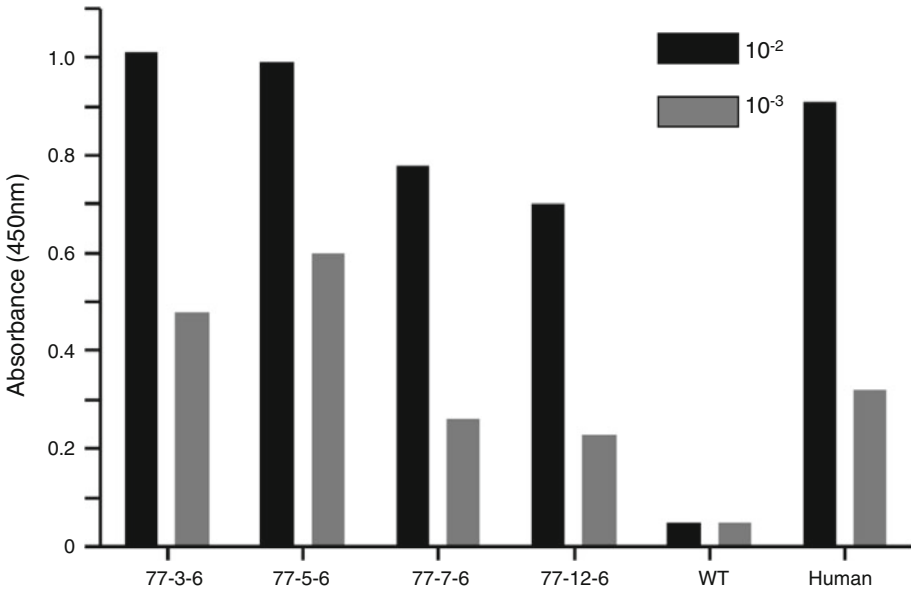

Fig. 5 Kronus ELISA used for detection of hTG in select T_1 seed extracts. Crude seed extract from one representative T_1 progeny (indicated by an asterisk in Fig. 4) was examined along with soluble protein from a nontransgenic seed (WT) and commercially purified hTG. Multiple dilutions of each sample were analyzed using the ELISA, and absorbance values for two of these dilutions (1:100 and 1:1,000) are shown

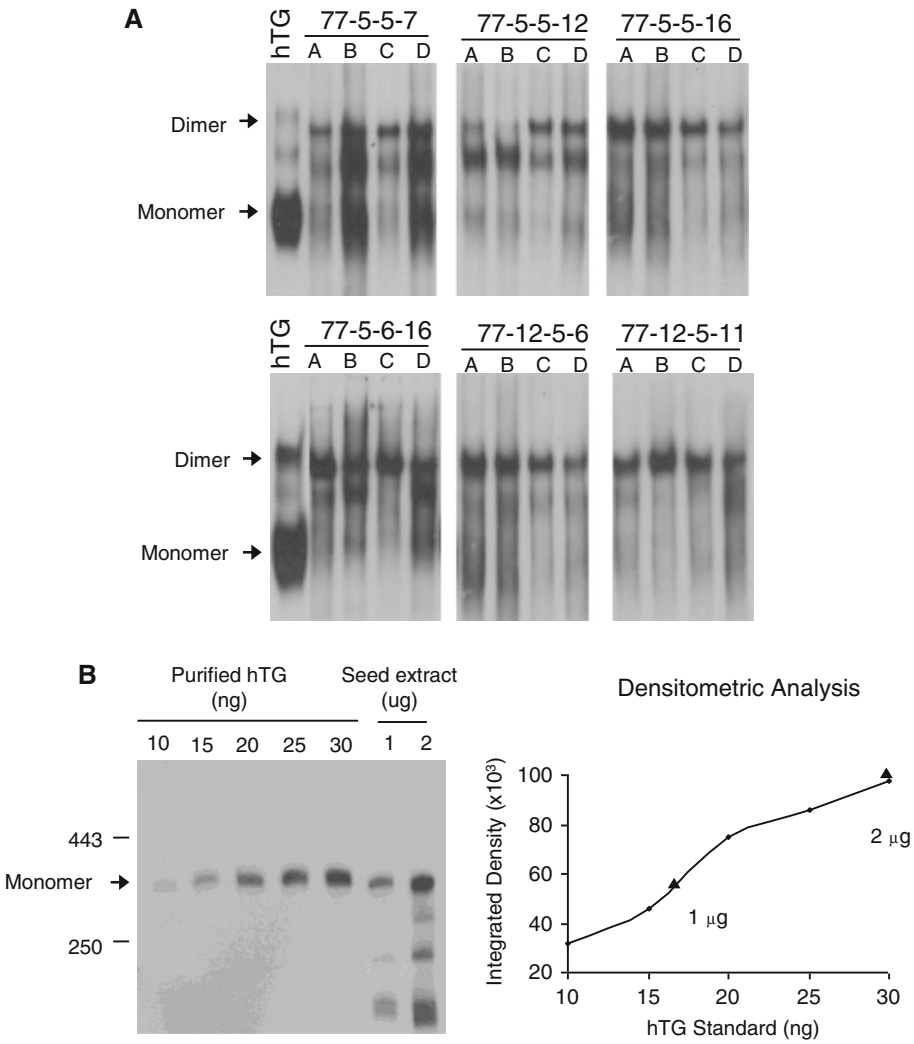

Fig. 6 Western analysis of T_2 seed protein. **a** Crude seed extracts (5 µg) from ten random T_3 progeny were separated in 5% native polyacrylamide gels and screened by western blot analysis for the presence hTG. Commercially purified hTG standard was included to visualize monomeric and dimeric forms of the protein (indicated by *arrows*). **b** Western quantification of recombinant hTG. Known amounts of commercially purified hTG protein and seed protein derived from line 77-5 are indicated. A densitometry curve was generated by scanning the gel image and plotting integrated density of each known standard using ImageJ software. Extrapolation from a best-fit curve revealed 16.8 and 29.2 ng of hTG protein in 1 and 2 µg of seed extract, respectively

hTG. These results clearly demonstrate that hTG is stably expressed over multiple generations. To estimate hTG as a percentage of total seed protein, western blot analyses were carried out using conditions in which hTG existed predominantly as monomeric protein. In separate experiments, we found that this could be accomplished by incubating protein with SDS sample buffer lacking β-mercaptoethanol for 10 min at room temperature prior to electrophoresis. Using these conditions, western analysis was carried out using known amounts of transgenic seed extract and known amounts of commercially purified hTG protein. Following western analysis, densitometry analyses of X-ray films were performed. The results from one such experiment are shown (Fig. 6b). Extrapolation from the hTG standard curve revealed that 1 µg of seed protein contained 16.8 ng of hTG ($\sim$1.7% of TSP) while 2 µg protein contained 29.2 ng of hTG ($\sim$1.5% TSP). The quantification results shown in Fig. 6b are representative of those found in T_3 progeny derived from lines 77-5 and 77-12 and were also verified by ELISA (data not shown).

Partial purification of soybean-derived hTG

Gel filtration chromatography was used to partially purify proteins from crude soluble seed extracts. A Sephacryl S-300 HR gel filtration column was calibrated by determining the peak elution volumes of a commercial set of molecular mass standards ranging in size from 669 to 29 kDa. The largest of these molecular mass standards was bovine thyroglobulin (MW $\sim$669 kDa) and eluted in fraction 20. Following calibration, transgenic seed extract from line 77-5 was applied to the Sephacryl column, and the eluted protein in each fraction was subjected to an ELISA for detection of hTG. The immunoreactive profile for soy-derived hTG is shown as a solid line (Fig. 7a). Fractions 17–23 contained detectable levels of hTG with the peak immunoreactivity localized to fractions 20 and 21. Notably, the elution profile for soy-derived hTG was consistent with the elution of the bovine thyroglobulin standard in fraction 20, suggesting that seed-specific hTG is likely folded and charged in a manner similar to that of the bovine thyroglobulin marker. For comparison, commercially purified

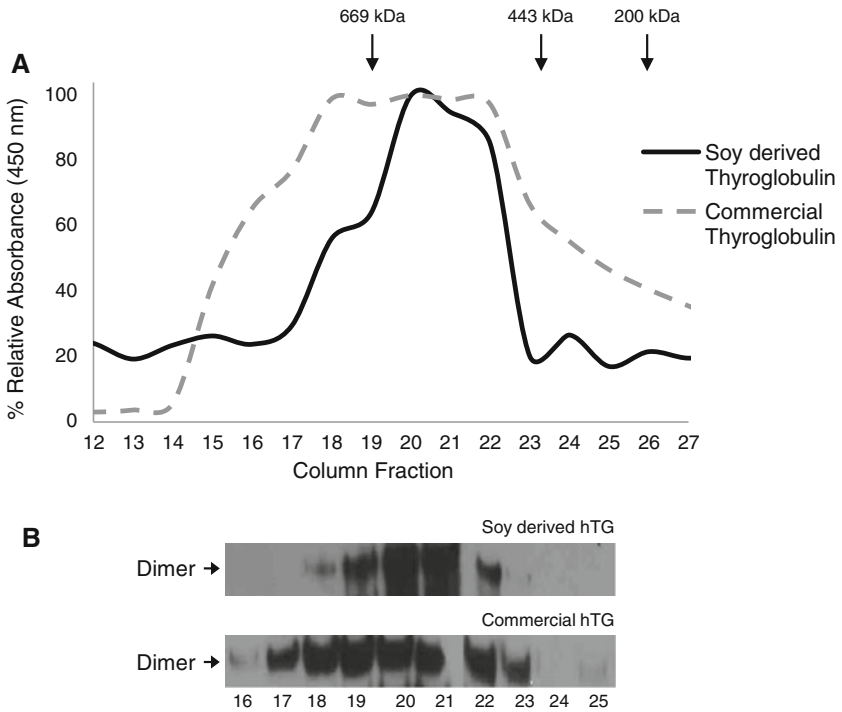

Fig. 7 Fractionation and partial purification soy-derived hTG from crude seed extracts. A calibrated Sephacryl S-300 HR column was used to separate crude transgenic seed protein and commercially purified hTG. **a** Eluted fractions were subjected to a capture ELISA for detection of hTG. The ELISA results are shown as a percentage of the relative absorbance (450 nm) for eluted fractions containing crude soybean seed protein (*solid line*) and human thyroid-purified protein (*dashed line*). The locations of molecular mass standards used for calibration are shown by arrows (669 kDa bovine thyroglobulin; 443 kDa β-amylase; 200 kDa alcohol dehydrogenase). **b** Western analysis of eluted fractions. Equal amounts of protein from the indicated fractions were separated in 5% native gels and subjected to western analysis. The location of the dimeric (D) form of hTG is indicated

hTG was also chromatographed on a Sephacryl column, similarly assayed for immunoreactivity and then plotted as a dashed line (Fig. 7a). The elution profiles of commercially purified hTG suggested that this protein was more heterogeneous than soy-derived hTG since high levels of immunoreactivity were detected in a broad peak throughout fractions 18–22. These results further suggested that commercially purified hTG is slightly heavier than soy-derived hTG which is consistent with the likely iodination of the human sample but not the soy-derived sample. Western analysis was performed to visualize immunoreactive protein in the eluted fractions. Equivalent volumes of either partially purified seed protein or commercially purified hTG were separated in native polyacrylamide gels and subjected to western analysis. The eluted fractions containing peak ELISA immunoreactivity also contained the greatest level of 660 kDa dimer in western experiments (Fig. 7b). As expected, the migration of soy hTG in extracts following partial purification was analogous to that of the commercially purified hTG further demonstrating the molecular similarities of both proteins when characterized under a variety of sizing and separating conditions.

Discussion

To date, the sole source of purified hTG for use as a research and diagnostic reagent is derived from human cadaver or surgically removed thyroid tissue. In this study, we have successfully expressed a recombinant form of hTG in soybean seeds. The expression of hTG in soybeans represents one of the largest recombinant proteins expressed in any plant host system to date and creates an alternative source for this relevant protein. The authenticity of recombinant hTG was confirmed by ELISA using commercially available kits developed specifically for the detection of hTG (Figs. 4, 5). Western analyses and gel filtration chromatography experiments demonstrated that soy-derived hTG and thyroid-purified hTG are biologically similar with respect to size, charge, mass and subunit interaction (Figs. 2, 6, 7). The expression of hTG in transgenic seed extracts was stable over multiple generations and represented ~1.5% of total soluble seed protein.

To our knowledge, there are no reports in the literature describing recombinant expression of hTG in any host system. It seems unlikely that attempts were never made to express this protein in traditional systems given the importance of hTG in the diagnostics industry, the absence of an accepted universal standard and current cost and technical issues associated with purifying thyroglobulin from human tissue. However, with the size and complexity of hTG as well as a stringent requirement for post-translational folding of the protein to avoid degradation, recombinant expression of hTG in any traditional host would seem difficult if not impossible. With the recombinant expression of hTG in soybeans, we have shown that soybeans represent an alternative host to traditional expression systems and may support the expression of other similarly large and complex proteins that are otherwise difficult or impossible to express in traditional systems.

We believe that recombinant expression of hTG in soybeans was possible since soybean seeds are biochemically unique with respect to the accumulation and storage of large protein complexes. This is evident by the abundance of 7S and 11S protein complexes that comprise the majority of total soybean seed protein (Thanh and Shibasaki 1976; Yagasaki et al. 1997). While the individual protein subunits that comprise these complexes are not considered large, fully assembled 7S and 11S complexes have molecular masses analogous to that of hTG (280–350 kDa for 7S and 320–360 kDa for 11S). Thus, it was conceptually logical to postulate that soybeans would support the recombinant expression of large and complex proteins such as hTG. While it is not unusual for a biological system to generate relatively large endogenous proteins, it was intriguing that a search of a soybean proteome database of known proteins (http://proteome.dc.affrc.go.jp/Soybean/) revealed no proteins larger than ~105 kDa. This observation is also consistent with published coomassie-stained gels of separated seed proteins in which proteins >105 kDa are not observed (Liu 1999). With our demonstration of recombinant hTG expression in soybeans, it is clear that cellular components of the seed are capable of producing and storing recombinant proteins that are severalfold larger than endogenous seed proteins with respect to molecular weight.

The accumulation of recombinant hTG in lead lines represented ~1.5% of total soluble protein as measured by western densitometry and ELISA. This observed accumulation of seed-specific hTG is comparable to levels of other recombinant proteins targeted to soybean seeds. For example, it was previously reported that recombinant basic fibroblast growth factor (bFGF) represented ~2.3% of total soluble soybean seed protein (Ding et al. 2006), and recombinant B subunit pentamers of *E. coli* heat labile toxin (LT-B) represented ~2.4% of total soluble soybean seed protein (Moravec et al. 2007). An expression level of 1.5% TSP for hTG represents a significant yield of recombinant protein on a dry weight basis as a single soybean seed will yield ~1 mg of recombinant protein (~150 mg seed weight × ~40% protein content × ~1.5% expression) with an average plant producing >200 mg of transgenic protein. Given previous reports of recombinant protein expression in soybean seeds, it is possible that increased levels of transgenic protein may be

obtained by targeting hTG to the endoplasmic reticulum (ER). However, given that hTG is a secreted protein, along with the fact that misfolded hTG protein complexes are degraded in vivo within the ER lumen of human thyrocytes (van de Graaf et al. 1997), we chose not to include the ER retention signal in our original construct design. During the course of these studies, we found that recombinant hTG was stable in seed extracts for >1 month at 4°C, and stable in seeds for >1 year when stored at ambient temperatures in the laboratory. The stability of recombinant proteins in soybean seeds is not surprising given that the primary purpose of the seed is to store proteins for long periods of time under conditions that may not be optimal for germination. Studies with other recombinant proteins targeted to soybeans have demonstrated stability of transgenic proteins when seeds were stored for up to 4 years under ambient conditions (Moravec et al. 2007; Oakes et al. 2009).

Protein stability and production of a uniform product will increase the means by which an alternative hTG source can be utilized. Commercially purified hTG is derived from human tissue and represents a heterogeneous population comprised of immature and mature forms of hTG. Immature forms of hTG lack carbohydrates and iodine residues while mature forms contain up to 10% carbohydrate and 1% iodine (van de Graaf et al. 1997). Thus, different batches of commercially purified hTG can be quite dissimilar with respect to purity and composition. Our data support this notion as hTG purified from human thyroid tissue eluted as a relatively broad peak on molecular sizing chromatography columns (Fig. 7). We speculate that soybean-derived hTG will be much more homogenous than thyroid-purified hTG. Since soybeans do not have enzymes for iodination, soy-derived hTG will not be iodinated. A non-iodinated, soy-derived hTG standard represents a significant advantage to current thyroid-derived hTG standards since iodination increases the antigenicity of the protein, thereby introducing the possibility of false and/or inconsistent test results when used in ELISA-based screens (Carayanniotis 2007; Nagayama et al. 2007). Similarly, since seed proteins are stored as "mature" forms of protein, it is likely that soy-derived hTG will be uniformly glycosylated. While protein glycosylation in soybean differs from that in humans, the antibodies used in FDA-approved ELISAs utilizing hTG as a reference standard recognize the protein backbone and not the carbohydrate residues (Gentile et al. 2004). We are currently conducting protein analytical experiments to characterize the extent and composition of carbohydrate modification on soybean-derived hTG. We anticipate the lack of iodination, along with uniform glycosylation of proteins within the seed, will contribute significantly to overall hTG protein homogeneity. While future studies will focus on purification and analytical characterization of soy-derived hTG, it will also be important to characterize the recognition of the soybean-derived protein by antibodies found in human sera, particularly from individuals with autoimmune hypothyroid disease. The confirmation of soy-derived hTG as a bio-equivalent to hTG purified from homogenates of human tissue will result in the first alternative sourcing of this protein. It is unclear what other uses may be derived from this never-before-available resource, but possibilities include the production of novel medical assays and devices to improve detection and/or treatment of thyroid cancer and disease. Furthermore, the availability of soy-derived hTG could drastically improve the basic research relating to thyroid dysfunction and disease due to the increased supply and reduced cost of the recombinant molecule. While *G. max* supported the expression and accumulation of hTG targeted to seeds, it is tempting to speculate that this platform could also support the expression of other proteins that to date have remained difficult or recalcitrant to expression using traditional host systems.

Acknowledgments The authors gratefully thank Dr. Scott Olenych and the confocal microscopy group at the David H. Murdock Research Institute for assistance with confocal microscopy, and Dr. Dan Nelson for stimulating discussions and critical reading of the manuscript. This work was supported in part by NIH grant R43CA150324 to K.J.P.

References

Abramoff MD, Magelhaes PJ, Ram SJ (2004) Image processing with ImageJ. Biophoton Int 11:36–42

Carayanniotis G (2007) Recognition of thyroglobulin by T cells: the role of iodine. Thyroid 17:963–973

Champion BR, Rayner DC, Byfield PG, Page KR, Chan CT, Roitt IM (1987) Critical role of iodination for T cell recognition of thyroglobulin in experimental murine thyroid autoimmunity. J Immunol 139:3665–3670

Daniell H, Singh ND, Mason H, Streatfield SJ (2009) Plant-made vaccine antigens and biopharmaceuticals. Trends Plant Sci 14:669–679

Demain AL, Vaishnav P (2009) Production of recombinant proteins by microbes and higher organisms. Biotech Adv 27:297–306

Deshpande V, Venkatesh SG (1999) Thyroglobulin, the prothyroid hormone: chemistry, synthesis and degradation. Biochim Biophys Acta 1430:157–178

Ding SH, Huang LY, Wang YD, Sun HC, Xiang ZH (2006) High-level expression of basic fibroblast growth factor in transgenic soybean seeds and characterization of its biological activity. Biotechnol Lett 28:869–875

Dupuy C, Deme D, Kaniewski J, Pommier J, Virion A (1988) Ca^{2+} regulation of thyroid NADPH-dependent H_2O_2 generation. FEBS Lett 233:74–78

Dupuy C, Virion A, Ohayon R, Kaniewski J, Deme D, Pommier J (1991) Mechanism of hydrogen peroxide formation catalyzed by NADPH oxidase in thyroid plasma membrane. J Biol Chem 266:3739–3743

Gentile F, Conte M, Formisano S (2004) Thyroglobulin as an autoantigen: what can we learn about immunopathogenicity from the correlation of antigenic properties with protein structure? Immunology 112:13–25

Gerngross TU (2004) Advances in the production of human therapeutic proteins in yeasts and filamentous fungi. Nat Biotechnol 22:1409–1414

Giddings G, Allison G, Brooks D, Carter A (2000) Transgenic plants as factories for biopharmaceuticals. Nat Biotechnol 18:1151–1155

Hajdukiewicz P, Svab Z, Maliga P (1994) The small versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. Plant Mol Biol 25:989–994

Karg SR, Kallio PT (2009) The production of biopharmaceuticals in plant systems. Biotechnol Adv 27:879–894

Krahn J, Dembinski T (2009) Thyroglobulin and anti-thyroglobulin assays in thyroid cancer monitoring. Clin Biochem 42:416–419

Liu K (1999) Soybeans: chemistry, technology and utilization. Aspen Publishers, Inc., Gaithersburg

Marino M, McCluskey RT (2000) Role of thyroglobulin endocytic pathways in the control of thyroid hormone release. Am J Physiol 279:1295–1306

Moravec T, Schmidt MA, Herman EM, Woodford-Thomas T (2007) Production of *Escherichia coli* heat labile toxin (LT) B subunit in soybean seed and analysis of its immunogenicity as an oral vaccine. Vaccine 25:1647–1657

Nagayama Y, Horie I, Saitoh O, Nakahara M, Abiru N (2007) CD4+ CD25+ naturally occurring regulatory T cells and not lymphopenia play a role in the pathogenesis of iodide-induced autoimmune thyroiditis in NOD-H2h4 mice. J Autoimmun 29:195–202

Oakes JL, Bost KL, Piller KJ (2009) Stability of a soybean seed-derived vaccine antigen following long-term storage, processing and transport in the absence of a cold chain. J Sci Food Agric 89:2191–2199

Paz MM, Martinez JC, Kalvig AB, Fonger TM, Wang K (2006) Improved cotyledonary node method using an alternative explant derived from mature seed for efficient *Agrobacterium*-mediated soybean transformation. Plant Cell Rep 25:206–213

Piller KJ, Clemente TE, Jun SM, Petty CC, Sato S, Pascual DW, Bost KL (2005) Expression and immunogenicity of an *Escherichia coli* K99 fimbriae subunit antigen in soybean. Planta 222:6–18

Thanh VH, Shibasaki K (1976) Major proteins of soybean seeds. A straight forward fractionation and their characterization. J Agric Food Chem 24:1117–1121

van de Graaf SA, Pauws E, de Vijlder JJ, Ris-Stalpers CR (1997) The revised 8307 base pair coding sequence of human thyroglobulin transiently expressed in eukaryotic cells. Eur J Endocrinol 136:508–515

Yagasaki K, Takagi T, Sakai M, Kitamura K (1997) Biochemical characterization of soybean protein consisting of different subunits of glycinin. J Agric Food Chem 45:656–660

Yang SX, Pollock HG, Rawitch AB (1996) Glycosylation in human thyroglobulin: location of the N-linked oligosaccharide units and comparison with bovine thyroglobulin. Arch Biochem Biophys 327:61–70