

Subcellular targeting is a key condition for high-level accumulation of cellulase protein in transgenic maize seed

Elizabeth E. Hood^{1,*}, Robert Love², Jeff Lane², Jeff Bray², Richard Clough², Kamesh Pappu³, Carol Drees³, Kendall R. Hood¹, Sangwoong Yoon¹, Atta Ahmad³ and John A. Howard³

¹Arkansas State University, PO Box 2760, State University, AR 72467, USA

²Texas A & M University, Food Protein Research Center, Box 2476, College Station, TX 77843, USA

³Applied Biotechnology Institute, Building 36, California Polytechnic State University, San Luis Obispo, CA 93407, USA

Received 14 January 2007;

revised 27 May 2007;

accepted 1 June 2007.

*Correspondence (fax 001-870-972-2336;

e-mail ehoo@astate.edu)

Summary

Ethanol from lignocellulosic biomass is being pursued as an alternative to petroleum-based transportation fuels. To succeed in this endeavour, efficient digestion of cellulose into monomeric sugar streams is a key step. Current production systems for cellulase enzymes, i.e. fungi and bacteria, cannot meet the cost and huge volume requirements of this commodity-based industry. Transgenic maize (*Zea mays* L.) seed containing cellulase protein in embryo tissue, with protein localized to the endoplasmic reticulum, cell wall or vacuole, allows the recovery of commercial amounts of enzyme. E1 cellulase, an endo- β -1,4-glucanase from *Acidothermus cellulolyticus*, was recovered at levels greater than 16% total soluble protein (TSP) in single seed. More significantly, cellobiohydrolase I (CBH I), an exocellulase from *Trichoderma reesei*, also accumulated to levels greater than 16% TSP in single seed, nearly 1000-fold higher than the expression in any other plant reported in the literature. The catalytic domain was the dominant form of E1 that was detected in the endoplasmic reticulum and vacuole, whereas CBH I holoenzyme was present in the cell wall. With one exception, individual transgenic events contained single inserts. Recovery of high levels of enzyme in T₂ ears demonstrated that expression is likely to be stable over multiple generations. The enzymes were active in cleaving soluble substrate.

Keywords: cellulase, transgenic maize, seed protein, biomass, ethanol subcellular targeting.

Introduction

Fossilized hydrocarbon-based energy sources, such as coal, petroleum and natural gas, provide a limited, non-renewable resource pool. Because of the world's increasing population and increasing dependence on energy sources for electricity and heating, transportation fuels and manufacturing processes, energy consumption is rising at an accelerating rate. The US transportation sector alone consumed approximately 180 billion gallons of gasoline in 2006 [Environmental Impact Assessment (EIA) report; http://www.eia.doe.gov/oiaf/aeo/pdf/aeotab_2.pdf]. Most (~60%) of the oil currently used in the USA is imported, creating a precarious situation in today's political climate, because supply disruptions are highly likely and could cripple the ability of the economy to function. Moreover, fossil petroleum resources, on which our standard of living currently depends, will probably be severely limited within the next 50–100 years.

Ethanol is key to the partial replacement of petroleum resources, particularly for transportation fuels. Ethanol is a product of anaerobically fermented sugars and, as such, can be produced from starch or cellulose, each a polymer of glucose. In 2005, the USA manufactured 4.3 billion gallons of ethanol from corn grain-derived starch (American Coalition of Ethanol Production, www.ethanol.org). Although production increases annually, ethanol produced from cornstarch cannot meet gasoline fuel replacement demands because of the volumes required.

The production of ethanol from lignocellulosic biomass can utilize large volumes of agricultural resources that are untapped today. In a recent report (http://www1.eere.energy.gov/biomass/pdfs/final_billionton_vision_report2.pdf), the authors estimated that, with current trends in technology, the USA could produce a billion tons of biomass per year in 20 years. These biomass resources include unharvested residues from

agricultural crops, forest slash, biomass crops, such as switchgrass, and municipal solid waste (MSW). Unharvested crop residues are estimated to have a mass approximately equal to the harvested portion of the crops. Specifically for the corn crop, if half of the residue could be used as a feedstock for the manufacture of ethanol, about 120 million tons of corn stover would be available annually for biomass conversion processes. Assuming that 85 gallons of ethanol can be produced per ton of untreated stover, approximately 10–12 billion gallons of ethanol could be generated from the stover portion of the corn crop today. Although substantial, this also falls short of replacement volumes.

Making ethanol from any source of lignocellulosic biomass is not cost-effective with the current enzyme production systems and the biomass collection and pretreatment technology (Aden *et al.*, 2002). High costs are a result of several factors: (i) the expense of collecting and transporting the feedstock raw material to processing plants; (ii) the expense of producing polysaccharide-degrading enzymes; and (3) the high cost of pretreating the lignocellulosic raw material to facilitate its enzymatic degradation. Efforts are underway in many laboratories to address each of these cost targets. The work described here addresses the issue of enzyme production.

Plant lignocellulosic biomass is a complex matrix of polymers comprising the polysaccharides cellulose and hemicellulose, and a polyphenolic complex, lignin, as the major structural components (Buchanan *et al.*, 2002). Any strategy designed to substitute lignocellulosic feedstocks for petroleum in the manufacture of fuels and chemicals must include the ability to efficiently convert the polysaccharide components of plant cell walls to soluble, monomeric sugar streams. Cellulose, the most abundant biopolymer on earth, is a simple, linear polymer of glucose. However, its semicrystalline structure is notoriously resistant to hydrolysis by both enzymatic and chemical means. Yet, high yields of glucose from cellulose are critical to any economically viable biomass utilization strategy.

Nature has developed effective cellulose hydrolytic machinery, mostly microbial in origin, for the recycling of carbon from plant biomass in the environment. Without it, the global carbon cycle would not function. To date, many cellulase genes have been cloned and sequenced from a wide variety of bacteria, fungi and plants, and many more certainly await discovery and characterization (Tomme *et al.*, 1995; Schulein, 2000). Cellulases are a subset of the glycosyl hydrolase superfamily of enzymes that have been grouped into at least 13 families based on protein sequence similarity, enzyme reaction mechanism and protein fold motif. Cellulose is degraded through the synergistic action of two general types of

cellulase enzyme. Enzymes that cleave the cellulose chain internally are referred to as endo-1,4- β -D-glucanases (E.C. 3.2.1.4), and serve to provide new reducing and non-reducing chain termini on which exo-1,4- β -D-glucanases (cellobiohydrolase, CBH; E.C. 3.2.1.91) can operate (Tomme *et al.*, 1995). Two types of exoglucanase have been described that differ in their approach to the cellulose chain. One type attacks the non-reducing end and the other attacks the reducing end. The product of the exoglucanase reaction is typically cellobiose, but other short-chain cello-oligomers are also produced. A third activity, β -D-glucosidase (E.C. 3.2.1.21), is required to cleave cellobiose and other oligomers to glucose. The β -D-glucosidase activity is required at 100–1000 times lower concentration than the cellulases, and can often be supplied by the fermentative organism.

The plant production system for industrial enzymes is a cost-effective, viable alternative to submerged culture fermentation systems, and has been used to successfully produce several enzymes in transgenic maize (Hood and Woodard, 2002; Hood *et al.*, 2003; Woodard *et al.*, 2003). Previous work to express cellulases in plants has been successful for the thermophilic enzyme E1 endo-1,4- β -D-glucanase, but has failed to produce the enzyme required in the largest amounts, CBH I, at levels enabling a commercial enterprise (Table 1) (Ziegelhoffer *et al.*, 1999; Dai *et al.*, 2000; Ziegler *et al.*, 2000). Thus, the objective of this work was to test the parameters enabling the high-level accumulation of cellulases, particularly CBH I, in a system in which the cost of production could potentially meet the cost target for the enzymes in the biomass-to-ethanol industry. We have chosen to express cellulases in maize as a first step towards developing the maize production system for lignocellulosic degradation.

Results

Choice of enzymes and genes

Converting pretreated lignocellulosic biomass to monomeric sugar streams requires, at a minimum, an endoglucanase and an exoglucanase. An additional consideration is whether these enzymes will produce detrimental effects on plants during their growth and development at ambient temperatures, and thermostable enzymes with high-temperature optima are less likely to do so than enzymes with optima in the 30 °C region. Several enzyme systems have been studied for their activity on crystalline cellulose, not all of which show synergy on this substrate (Baker *et al.*, 1998). Therefore, in choosing enzymes for these experiments, we considered the following:

Table 1 Examples of heterologous cellulase expression in plants, and production considerations

Enzyme	Gene source	Transgenic plant system	Expression level	Scalability§
Endo-1,4-β-D-glucanase	Bacterial (<i>Acidothermus</i>)	<i>Arabidopsis</i> (cell wall targeted)	26%TSP in leaves*	–
Endo-1,4-β-D-glucanase	Bacterial (<i>Acidothermus</i>)	Potato (cell wall or chloroplast targeted)	2.6%TSP† in leaves	+
Endo-1,4-β-D-glucanase	Bacterial (<i>Thermonospora fusca</i>)	Alfalfa (cytosolic localization)	~0.01%TSP‡ in leaves	++
Cellobiohydrolase	Bacterial (<i>T. fusca</i>)	Tobacco (cytosolic localization)	0.1%TSP‡ in leaves	+
		Alfalfa (cytosolic localization)	0.02%TSP‡ in leaves	++
Endo-1,4-β-D-glucanase	Bacterial (<i>Acidothermus</i>)	Tobacco (cytosolic localization)	0.002%TSP‡ in leaves	+
Endo-1,4-β-D-glucanase	Bacterial (<i>Acidothermus</i>)	Maize (cell wall targeted)	2.1% TSP in leaves¶	++

TSP, total soluble protein. Percentage of TSP assumes that 10% of leaf weight is soluble protein.

*Ziegler *et al.* (2000).

†Dai *et al.* (2000).

‡Ziegelhoffer *et al.* (1999).

§Scalability defined by 2002 US crop acreage; scale-up potential: –, unscalable; +, fair; ++, moderate; +++, significant.

¶Biswas *et al.* (2006).

Table 2 Cellulase enzyme characteristics

	E1	CBH I
Family	5 – E.C. 3.2.1.4	7 – E.C. 3.2.1.91
Native source	Bacterial	Fungal
Calculated molecular weight (Da)	56 500	52 500
(catalytic domain) (Da)	40 610	
Molecular weight by SDS-PAGE (Da)	72 000	70 000
(catalytic domain) (Da)	60 000	
Glycosylated native protein	No	Yes, primarily linker region
pI holoenzyme	5.2	4.51
(catalytic domain)	4.87	
pH optimum	5–6	5
Temperature optimum (°C)	81	45–50
Bond cleaved	β-1,4-glycosidic	β-1,4-glycosidic
Mechanism	Retained anomeric configuration*	Retained anomeric configuration*
Substrates	Cellulose fibrils; purified cellulose preparations (Solka-floc, Sigmacell, Avicel); <i>para</i> -nitrophenyl-β-1,4-D-cellobiose (pNPC); methylumbelliferyl-β-1,4-D-cellobioside (MUC)	Cellulose fibrils; purified cellulose preparations (Solka-floc, Sigmacell, Avicel)
Primary reaction products	Decreased degree of polymerization (dp), long-chain, water-insoluble cellulose	Cellobiose (and other water-soluble short-chain cello-oligomers)

CBH I, cellobiohydrolase I from *Trichoderma reesei*; E1, endocellulase from *Acidothermus cellulolyticus*; SDS-PAGE, sodium dodecylsulphate-polyacrylamide gel electrophoresis.

*Source: Schulein (2000); <http://www.expasy.ch/cgi-bin/lists?glycosid.txt>.

(i) enzymes should exhibit synergistic activity on lignocellulosic substrates; (ii) enzymes should be thermostable to at least 45 °C; and (iii) enzymes should have compatible pH optima.

Using these criteria, we chose the E1 endo-1,4-β-D-glucanase from *Acidothermus cellulolyticus* (Mohagheghi *et al.*, 1986; Nieves *et al.*, 1995) (S.R. Thomas *et al.*, US Patent #5 536 655) and CBH I from *Trichoderma reesei* (Shoemaker *et al.*, 1983). E1 and CBH I exhibit synergistic activity on lignocellulosic substrates that have been pretreated with dilute acid and steam (Baker *et al.*, 1998). E1 cuts

cellulose polymers internally to liberate free ends for exoglucanase activity. Although E1 has optimal activity at 81 °C, it is still highly active at 45–50 °C, compatible with the CBH I enzyme. CBH I processively degrades cellulose from non-reducing free ends. CBH I shows optimal and sustained activity at temperatures up to 50 °C. These two enzymes have compatible pH optima at pH 5–6. The actual enzymes chosen for this project are key to demonstrating this production system technology. Some physical characteristics of these proteins are presented in Table 2.

Vectors for transformation

To our knowledge, no-one has attempted to express cellulases in seed, specifically maize seed. Therefore, it was necessary to test multiple parameters to determine the best condition for high-level accumulation of each of these enzymes in our system. Several factors have an impact on our ability to achieve high concentrations of target proteins in tissues of interest, in particular the codon usage in the gene, promoter choice for the generation of messenger RNA (mRNA) and subcellular location of the protein (Streatfield, 2007). Thus, the first 40 codons of each coding sequence, including the signal sequence, were optimized for maize expression. Moreover, the embryo-preferred globulin-1 promoter from maize (Belanger and Kriz, 1991) was used to express the cellulase genes, as this promoter has been shown in several previous studies to drive high levels of protein expression (Streatfield *et al.*, 2002; Hood *et al.*, 2003). Subcellular locations, including the endoplasmic reticulum (ER), vacuole and cell wall, were chosen on the basis of successful protein accumulation in previous studies (Larrick *et al.*, 2002; Streatfield *et al.*, 2002; Hood *et al.*, 2003). The ER and cell wall constructs contained an N-terminal fusion with the barley α -amylase signal sequence (BAASS; Figure 1b) (Rogers, 1985). Further targeting to the ER was achieved by an additional KDEL sequence (lysine, aspartate, glutamate, leucine) on the 3' end. The targeting sequence for the vacuole was derived from the maize lytic vacuole (Figure 1b) (Holwerda *et al.*, 1992). The vectors also included a herbicide resistance gene (*35S:pat*) for the selection of transgenic events in tissue culture and in the field (White *et al.*, 1990). The selectable marker and the gene of interest were located between the left and right borders of *Agrobacterium tumefaciens* T-DNA (see Figure 1) to facilitate integration into plant chromosomes.

(a)

LB

RB

p35S moPAT t35S	pGlob-1 BAASS E1 cellulase pin II	BCD
p35S moPAT t35S	pGlob-1 BAASS E1 cellulase KDEL pin II	BCF
p35S moPAT t35S	pGlob-1 vacuolar seq E1 cellulase pin II	BCH
p35S moPAT t35S	pGlob-1 BAASS CBH I pin II	BCC
p35S moPAT t35S	pGlob-1 BAASS CBH I KDEL pin II	BCE
p35S moPAT t35S	pGlob-1 vacuolar seq CBH I pin II	BCL

(b)

Vacuole Targeting Sequence:

ATGGCCCACGCCCGCGTCTCCTCCTCCTGCGCTCGCCGTCCTGGCCACGGCC
GCGTTCGCGCTCGCCTCCTCCTCCTCCTTCGCGGACTCCAACCCGATCCGGC
CGGTACCCGACCGCGCCGCGTCCACC

Barley Alpha Amylase Sequence:

ATGGCGAACAAAGCACCTGAGCCTTAGCCTTCTCCTCGTGCTCCTGGGCCTCT
CCGCTCCCTCGCCTCCGGC

Generation of transgenic plants

Transgenic maize plants were generated using *A. tumefaciens* co-cultivation with immature zygotic embryos from Hi-II maize. Multiple independent transgenic events (ITEs) were recovered for each vector constructed to express either E1 endocellulase or CBH I exocellulase in transgenic maize seed (Table 3). Herbicide-resistant callus tissue from each of the independent events yielded multiple clonal plants from somatic embryos. The tissue culture-derived plantlets were transplanted to soil-less medium in the glasshouse and allowed to flower. Female flowers from the transgenic plants were pollinated with pollen from an elite inbred and allowed to set seed.

Six ITEs were recovered for ER-localized E1 cellulase and three for vacuolar-targeted enzyme; 7–10 plants per ITE produced seed for analysis. The construct for cell wall-localized E1 enzyme generated 16 robust ITEs. Transformation vectors containing the CBH I gene produced 8–13 ITEs that generated 8–10 clonal plants each, which flowered and produced seed for analysis (Table 3).

Characterization of ITEs

E1 endocellulase

Six seeds from each ear produced by a transgenic plant were individually analysed in order to assess the degree of protein accumulation. Two ITEs designed to express E1 in the ER were null (data not shown). Values of E1 accumulation for all positive seeds for the top three ITEs are shown in Figure 2. ITE BCF06, exhibiting the highest single seed value of 17.9% total soluble protein (TSP), comprised five plants with 8–17% TSP in positive seeds. All positive seeds in the three highest expressing ITEs showed E1 values of $\geq 3\%$ TSP (BCF03, 05, 06; Figure 2). The average protein accumulation for all positive seeds was 6.1% TSP (Table 3).

Figure 1 (a) Vectors for expression of the E1 cellulase and cellobiohydrolase genes in maize, targeting the protein to the cell wall, endoplasmic reticulum or vacuole. Each expression cassette is transferred to a vector that co-integrates in *Agrobacterium* with a superbinary vector with a homologous *cos* site for this co-integration. (b) Vacuole targeting sequence and barley α -amylase signal sequence (BAASS).

Table 3 Recovery and expression of independent transgenic events in maize. Events and plants that produced seed from vectors designed to express cellulose-degrading enzymes in transgenic maize. The value for the highest single seed expression level is shown, as well as the mean of all positive seeds for a particular vector. Values for second-generation, 50-seed bulk samples (T_2) are also shown

Gene	Target organelle	Plasmid name	Number of events	Total number of plants	T_1 high seed TSP (%)	T_1 mean TSP (%)	T_2 TSP (%)
E1	Cell wall	BCD	16	154	ND	0.5	ND
	ER	BCF	6	42	17.9	6.1 ± 4.8	3.5–4.5
	Vacuole	BCH	3	31	16	5.6 ± 2.7	8–9
CBH I	Cell wall	BCC	8	69	17.8	3.2 ± 1.8	3.7–5.1
	ER	BCE	13	107	16.3	4.1 ± 2.6	3.2–3.8
	Vacuole	BCL	10	96	0	0	ND

CBH I, cellobiohydrolase I from *Trichoderma reesei*; ER, endoplasmic reticulum; E1, endocellulase from *Acidothermus cellulolyticus*; ND, not determined; TSP, total soluble protein.

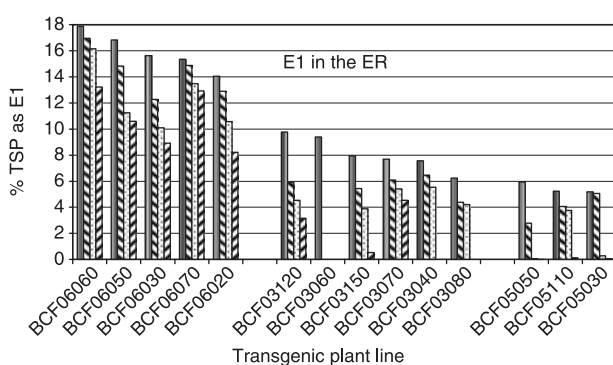


Figure 2 Endoplasmic reticulum (ER)-localized E1 cellulase expression levels in all T_1 positive seeds analysed in the three highest expressing individual transgenic events. Data are arranged by highest expressing individuals within each of the independent transgenic events (ITEs), and by expression amongst the three highest ITEs. TSP, total soluble protein.

In anticipation of creating future industrial production lines from these ITEs, seeds from three plants in this group, BCF0605, BCF0312 and BCF0307, were planted in the glasshouse. Plants were crossed with elite inbred maize varieties for agronomic trait improvement. Seeds from T_2 ears were analysed in 50-seed bulk quantities. E1 accumulation in these ears was in the range 3.5–4.5% TSP, excellent results considering these are averages of 1 : 1 segregating positive and negative seeds with multiple expression levels (Table 3).

The three ITEs from the vector expressing E1 targeted to the vacuole produced seed that contained large quantities of the enzyme (Figure 3). Individual positive seed ranged in expression from 1.5% to 16% TSP. Each of these lines produced ample seed, and three lines from two of the ITEs, BCH0206, BCH0207 and BCH0101, were planted in the glasshouse to produce T_2 seed through crosses with elite inbred maize varieties. T_2 ears ranged in expression from 8%

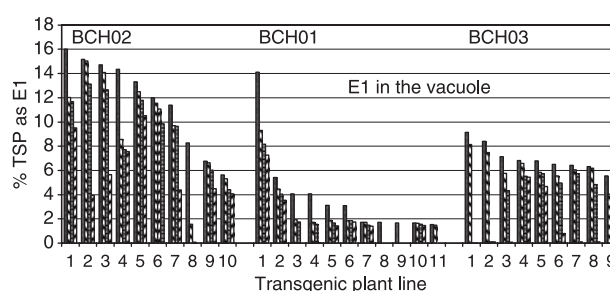


Figure 3 Vacuolar-localized E1 cellulase expression levels in all positive seeds analysed in the three highest expressing individual transgenic events. Data are arranged by highest expressing individuals within each of the independent transgenic events (ITEs), and by expression amongst the three highest ITEs. TSP, total soluble protein.

to 9% TSP in 50-seed bulk quantities (Table 3), even higher than the lines from the ER-targeted enzyme.

Initial screening of bulk seed samples of E1 targeted to the cell wall showed expression levels below those of the vacuole and ER, and further screening of individual seed was not performed (Table 3).

The full-length E1 gene, as transformed into maize in these experiments, should produce a protein of approximately 70 kDa, as estimated by sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE) (Table 1). However, the E1 cellulase recovered in extracts from transgenic maize seed was truncated to ~40 kDa, consistent with the size of the catalytic domain of this enzyme (Figure 4a). Importantly, the E1 cellulase is enzymatically active, meeting a major goal for an industrial enzyme. The amounts of protein estimated from the activity assays and by Western blot are similar, indicating that the protein is predominantly in the active form.

Wide variation in protein accumulation occurred in the various ITEs (Figures 2 and 3). Moreover, within an ITE, plant-to-plant variation was observed for the clonal plants, particularly

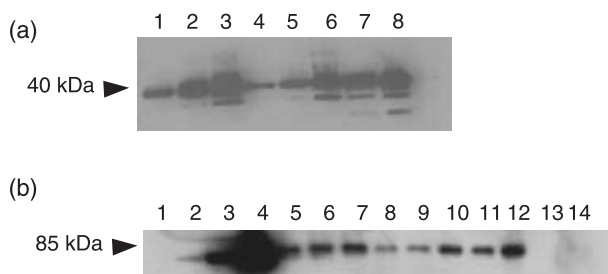


Figure 4 (a) Western blot of single seed extracts from E1 cellulase individual transgenic events. Lanes: 1–3, 100, 200 and 400 ng of extracts from BCH01; 4–6, 100, 200 and 400 ng of extracts from BCF06; 7–8, 15 or 30 ng of E1 control protein; 9, control corn extract. Arrowhead marks the 40-kDa molecular weight standard. (b) Western blot of single seed extracts from cellobiohydrolase I (CBH I) transgenic events. Lanes: 1, 1 µg corn protein only; 2–4, 1, 10 and 100 ng CBH I control protein; 5–9, BCC01 individual seed extracts; 10–12, BCC02 individual seed extracts; 13, 14, BCE02 individual seed extracts. Arrowhead marks the 85-kDa molecular weight standard.

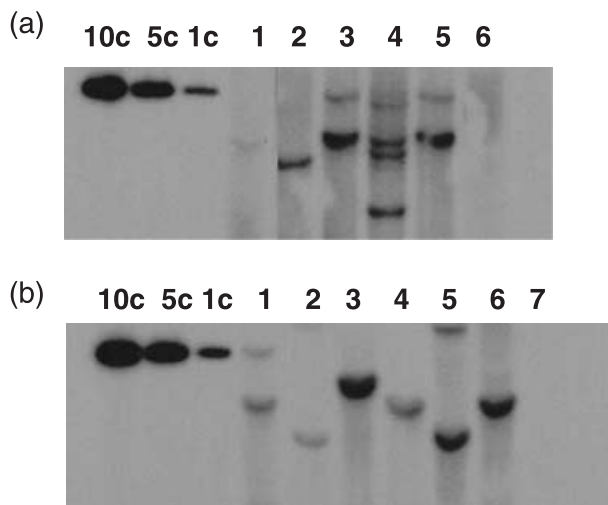


Figure 5 Southern blots of E1 cellulase (BCF and BCH) lines (a) and cellobiohydrolase I (CBH I) (BCC and BCE) lines (b). (a) 1, 5 and 10 copy reconstructions of BCH transformation plasmid (11.1 kb) per 10 µg of corn DNA. Lanes: 1, BCH03; 2, BCH01; 3, BCF06; 4, BCF05; 5, BCF03; 6, control DNA from untransformed corn. (b) 1, 5 and 10 copy reconstructions of BCC transformation plasmid (11.1 kb) per 10 µg of corn DNA. Lanes (10 µg per lane): 1, BCE16; 2, BCE07; 3, BCE09; 4, BCC08; 5, BCC07; 6, BCC02; 7, control DNA from untransformed corn.

for BCH01 (Figure 3). Thus, screening of individual seed from each plant is critical for understanding the range of protein accumulation possible in T_1 seed. Each of the E1 lines for which DNA hybridization experiments were performed had single insertions, with the exception of BCF05 which showed three insertions (Figure 5a). However, as seed from this line was not planted in subsequent experiments, the stability of expression in the next generation was not determined.

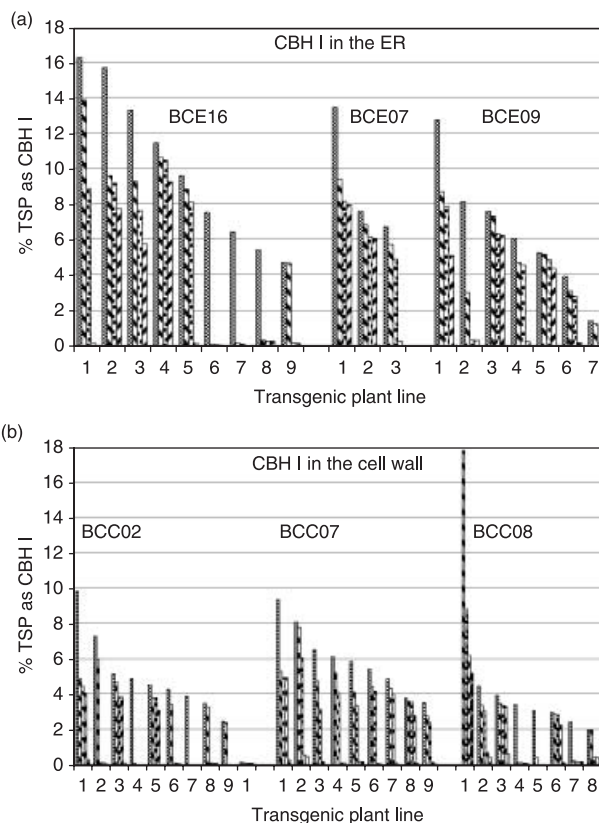


Figure 6 (a) Endoplasmic reticulum (ER)-localized cellobiohydrolase I (CBH I) expression levels in all T_1 positive seeds analysed in the three highest-expressing independent transgenic events (ITEs). Data are arranged by highest expressing individuals within each of the ITEs, and by expression amongst the three highest ITEs. (b) Cell wall-localized CBH I expression levels in all T_1 positive seeds analysed in the three highest expressing ITEs. Data are arranged by highest expressing individuals within each of the ITEs, and by expression amongst the three highest ITEs.

CBH I exocellulase

CBH I expression was analysed in all plants from each vector designed for expression in the cell wall, ER or vacuole (Figure 6). Event-to-event variation was observed, as was variation amongst clonal plants from each ITE. However, for CBH I, the overall best expression was observed in lines that contained the ER-targeted enzyme (Figure 6; Table 3), although cell wall-localized enzyme had the highest accumulating single seed (17.9% TSP; Figure 6). Three cell wall- and ER-targeted CBH I ITEs had essentially no positive seed – BCE06 and BCE08 (ER, data not shown) and BCC06 (cell wall; data not shown). The average for all positive seed for cell wall- and ER-localized enzymes was 3.2% and 4.1% TSP, respectively. Seeds from five lines of cell wall-targeted ITEs and from two lines of ER-targeted ITEs were planted in the glasshouse, and crossed with elite germplasm to begin agronomic trait and protein accumulation improvement. CBH I accumulated in T_2 seed at 3.2%–5.2% TSP in the 50-seed bulk quantities analysed (Table 3).

Western blots of ITEs showing CBH I activity yielded two results (Figure 4b). Protein detected on Western blots correlated with enzyme detected through activity assays for CBH I targeted to the cell wall (BCC events). The enzyme had approximately the same molecular weight as the wild-type protein (70 kDa). However, surprisingly, the ER-targeted version (BCE events) that showed strong activity with substrate showed no detectable band on Western blot (Figure 4b, lanes 13, 14). Further experiments suggested that the protein is truncated and that the epitope recognized by the monoclonal antibody is not present (data not shown). Vacuolar-targeted CBHI (BCL events) showed no enzyme activity, and no protein accumulated in the seed (data not shown). DNA hybridization analysis of six lines indicated that each has a single insert and that all ITEs are truly independently isolated (Figure 5b).

Discussion

The expression of heterologous proteins in transgenic plants is an established technology. The classes of proteins that have been successfully expressed in plant systems at very high levels include industrially useful enzymes, viral proteins, pharmaceutical proteins and peptides (including antibodies) and various structural proteins (Hood *et al.*, 1997, 2003; Witcher *et al.*, 1998; Daniell *et al.*, 2001; Streatfield *et al.*, 2001, 2002; Hood and Woodard, 2002; Woodard *et al.*, 2003). The factors affecting the level of accumulation for each of these protein classes differ. Although success depends on the characteristics of the individual protein, protein accumulation has been particularly successful when targeted to the cell wall, vacuole or ER (Streatfield, 2007).

We chose to target each of the two cellulases to these three organelles and to assess their levels of accumulation, because of previous success in accumulating recombinant protein in these subcellular locations. For two of the three vectors expressing E1 cellulase, levels of enzyme activity in T₁ seed were quite high, the highest at 15–18% TSP in single seed. We found the expected event-to-event variation in expression, in combination with clonal plant-to-plant variation, within an ITE (Figure 2), results observed with several expressed genes and published for laccase (Hood *et al.*, 2003). Many single seeds generated from each vector had values above 10% TSP, indicating that the high-accumulating seeds were not outliers. ER-localized E1 showed low to no expression of the enzyme in two ITEs (BCF01 and BCF04). All other E1 ITEs had plants accumulating protein above the 1.5% TSP level, at the high end of T₁ foreign protein recovery in maize (Hood and Woodard, 2002; Larrick *et al.*, 2002). The vector targeting E1 to the vacuole yielded only three ITEs (Figure 2). However,

all three ITEs produced high-expressing seed. In contrast, the cell wall-localized protein did not accumulate to as large an amount and the lines were not analysed further. This was a surprising result because, in most other cases, enzymes targeted to the cell wall yielded the largest amounts of recombinant protein (Hood *et al.*, 1997, 2003; Woodard *et al.*, 2003) compared with other subcellular locations. This unexpected lower level of expression in the cell wall is not likely to be the result of detrimental enzyme activity, as the optimal temperature for E1 activity is 81 °C. E1 was truncated to its catalytic subunit (Figure 4), similar to the results obtained in other transgenic systems (Ziegler *et al.*, 2000; Biswas *et al.*, 2006). Preliminary experiments indicate that enzyme activity is normal on cellulosic substrates (data not shown).

Presently, most industrial enzymes are produced primarily by submerged culture fermentation. The scale-up of fermentation systems for the large volumes of enzyme required for biomass conversion will be difficult and extremely capital intensive. For purposes of comparison, a single, very large (1 million litre), aerobic fermentation tank could produce approximately 3000 tons of cellulase protein per year in continuous culture. Currently, however, fermentation technology is practised commercially on a significantly smaller scale and in batch mode, so that production capacities are closer to 10% of the theoretical 3000 tons calculated above [S. Thomas, National Renewable Energy Laboratory (NREL), pers. commun.]. Thus, using these assumptions, current practices would yield 3000 times less than the 1.2 million tons of enzyme needed to convert the cellulose content from available corn stover (Walsh, 1999), not even considering the multitude of other feedstocks. The capital and operating costs of such a fermentative approach to producing cellulases are likely to be impractical because of the huge scale and capital investment that would be required.

When maize proteins are extracted in acidic buffer, as performed for the cellulases, levels of 15% TSP in maize seed extract represent approximately 0.04% of dry weight. This is not the target amount required for cost-effective enzymes for ethanol production from biomass. However, we have shown previously that breeding with elite inbred maize lines produces transgenic material, through selection, that shows improved seed yields, field performance and, most importantly, increased accumulation of the target protein to as much as 100-fold over the initial amounts in T₁ seed (Streatfield *et al.*, 2001; Hood *et al.*, 2003; Howard and Hood, 2005). Lines from ER-targeted E1 (ITEs BCF03 and BCF06) and vacuole-targeted E1 (BCH01 and BCH02) showed high single seed expression levels, close to 15% TSP, and therefore these plants were

chosen for the breeding programme for further development. T₂ seed contained E1 at 3.5–4.5% TSP in the ER and at 8–9% TSP in the vacuole. These amounts are for a 50-seed pool that contains approximately equivalent numbers of positive and negative seed, indicating immediately that the average for positive seed in these samples is 7–9% and 16–18% TSP, respectively. Self-pollination to produce homozygous plants and further improvement of agronomic characteristics will ensure increasing accumulation of the recombinant protein in future generations.

Several groups have successfully expressed E1 in transgenic plants to high levels (Table 1 and references cited therein), but we report the first high levels that have been expressed in a monocot seed, a commercially important crop, making our results significant for practical applications. However, what is even more significant and unusual is our results with CBH I. When CBH I was targeted to the cell wall, four plant lines from three ITEs produced seed with protein levels above 8% TSP, with the best expression from this group at 17.9% TSP. In the ER-targeted version, CBH I accumulated to greater than 8% TSP in eight plant lines from three ITEs, with the highest level being 16.3% TSP. These high expression results are extremely important, because the CBH I enzyme has not been recovered previously at high expression levels in any plant or fungal system. The highest expression published to date is 0.02% TSP in tobacco leaves (Table 1) (Ziegelhoffer *et al.*, 1999). Thus, our highest single seed levels at 17.9% TSP are 900-fold higher than the next best tested system. Moreover, because we were able to consistently increase the T₁ expression levels of our proteins up to 100-fold through breeding and selection, we will probably reach cost-effective production levels in a few generations. T₂ expression levels of CBH I indicate that we are well on our way to this goal. Again, the use of several different vectors to express proteins in transgenic plants was of value, because only two of the three constructs used to express CBH I yielded plants with accumulated protein. The cell wall-localized protein had the expected molecular weight (70 kDa) and, in preliminary experiments, showed activity on cellulosic substrates (data not shown).

Why does the maize system show better results than previous attempts? The expression of foreign genes with resultant protein accumulation is affected by numerous factors. Certain factors can be addressed at the molecular level, i.e. the primary molecular sequence of the gene, promoter, leader sequence and terminator. The sequences of the genes in this study were scanned for codon usage and RNA destabilization sequences. The first 40 codons were optimized for expression in maize. This was found to be sufficient for optimal translation. In addition, an embryo-preferred

promoter (from the maize globulin 1 gene) was used on many genes, and worked extremely well for seed expression, our target tissue. We routinely use a transcriptional terminator derived from the potato protease inhibitor II gene (An *et al.*, 1987), and this was also employed in these constructs.

Another consideration for high-level protein accumulation is subcellular targeting of the protein. In many cases, the apoplast (cell wall) is the location of choice for protein accumulation (Hood *et al.*, 1997; Zhong *et al.*, 1999; Ziegler *et al.*, 2000; Hood and Woodard, 2002). However, we have also found that the vacuole (Streatfield *et al.*, 2002) and ER support large-scale accumulation of protein. In the current study, we tested three locations within the cell for targeting the cellulase enzymes – the apoplast, the vacuole and the ER. Each had an impact on protein accumulation, either positive or negative, for each gene. The actual sequences used for targeting the proteins have been tested multiple times in corn and have been shown to be optimal for these locations.

Additional factors that may have an impact on protein accumulation in maize, or any system, may affect protein stability in the target tissue and/or compartment. These factors could be simple proteases that decrease protein accumulation or protease inhibitors that stabilize proteins. However, undoubtedly, there are many cellular processes of which we are, as yet, unaware that have profound influences on foreign protein stability, and can only be empirically selected for. Until such time as we know more about these factors, multiple transgenic plants must be produced and screened for the accumulation of large amounts of the target protein.

Endocellulase activity comprises 20% and exocellulase activity comprises 80% of the enzyme cocktail necessary for efficient digestion of crystalline cellulose (Baker *et al.*, 1998). Thus, the fact that E1 accumulates to higher levels than CBH I in many transgenic systems is problematic for the recovery of cost-effective ratios of CBH I and E1 for saccharification purposes, particularly when previous attempts at the expression of CBH I have shown low recovery (Table 1). The maize seed production system apparently can overcome this difficulty. Expression of these enzymes separately in the maize system allows for the planting of balanced and sufficient acreage to produce the required amount of either enzyme. With improvements in enzyme accumulation through breeding and selection, lines will probably become available that produce sufficient enzyme on 1 acre for approximately 40 acres of non-transgenic corn stover lignocellulosic waste. These calculations are based on today's enzyme specific activity and production capabilities. Because plant biotechnology is a young technology, innovation and improvement in enzyme

specific activity and production capability are likely, thus lowering cost. In addition, these calculations are made for corn stover applications, but the enzymes and technology can be applied to any lignocellulosic feedstocks with minor modifications in enzyme content and ratio. These are important steps in the enablement of the biomass-to-ethanol industry.

Experimental procedures

Vector construction

Six expression vectors were constructed to produce the E1 and CBH I cellulases in maize, three vectors for each gene (Figure 1a). All contained the globulin 1 promoter from maize (Belanger and Kriz, 1991) and three targeting sequences for each protein – apoplast (Figure 1b) (Rogers, 1985), ER (KDEL sequence) and vacuole (Holwerda *et al.*, 1992) – followed in each case by the proteinase inhibitor II (pin II) terminator from potato (An *et al.*, 1987) (Figure 1a).

The E1 cellulase gene (accession number U33212; Table 2) from *Acidothermus cellulolyticus* (Mohagheghi *et al.*, 1986; Nieves *et al.*, 1995) (S.R. Thomas *et al.*, US Patent #5 536 655) was received from NREL. For expression in maize, the first 120 codons were optimized to maize preferred codons. BAASS (for secretion to the apoplast) and KDEL (lysine, aspartate, glutamate, leucine; for retention in the ER) sequences were added to the gene by polymerase chain reaction (PCR), using the NREL clone as template. The PCR product was moved to a PCR-ready cloning vector, then moved to an intermediate vector to add the pin II terminator sequence, and then shuttled into the plant expression vector as a complete unit. The globulin-1 promoter was just upstream of the E1 gene.

For targeting to the vacuole, the vacuole signal sequence (Figure 1b) was substituted for BAASS at the 5' end of the protein coding sequence. The vacuole-targeted version of the E1 cellulase gene was constructed using PCR by adding the vacuole leader to the codon preferred optimized E1 gene generated in the BAASS:E1 construct. This PCR product was cloned into an intermediate vector to add the pin II terminator, and then transferred to the plant expression vector downstream of the globulin-1 promoter.

The CBH I gene (Shoemaker *et al.*, 1983) (accession number X69976; Table 2), also obtained from NREL, most closely matches the CBH I gene from *Trichoderma koningii* at the nucleic acid level. The gene was maize optimized for the codons for the first 40 amino acids using a PCR-based mutagenesis approach, including the 24-amino-acid BAASS. Codons D346 and D386 were also maize codon optimized to remove the potentially destabilizing sequences at these positions. BAASS was added to the optimized CBH I gene by PCR. The PCR product was moved to a PCR-ready cloning vector to add the pin II terminator, and then the entire unit was transferred to the transformation vector. The globulin-1 promoter was used to drive transcription of the CBH I coding sequence. The KDEL sequence was added to this gene, as described above, to target the protein to the ER, and the vacuolar targeting sequence was substituted for BAASS, as described for E1, to target the CBH I protein to the vacuole.

Included in all vectors was the maize-optimized *pat* gene from *Streptococcus hygroscopicus* (White *et al.*, 1990), expressed from the cauliflower mosaic virus (CaMV) 35S promoter. It confers resistance to glufosinate ammonium (bialaphos), and is codon optimized for

expression in corn (maize-optimized *pat*, *mopat*; Anzai *et al.*, 1989; Gordon-Kamm *et al.*, 1990; Uchimiya *et al.*, 1993). The plant expression units were between T-DNA borders in the final cloning vector pSB11 from Japan Tobacco (Hiei *et al.*, 1994; Ishida *et al.*, 1996). The vector was mated into *A. tumefaciens* strain LBA4404 containing the superbinary vector pSB1 (Hiei *et al.*, 1994). After mating, the co-integrated vector pSB111 was isolated and electroporated into EHA101 (Hood *et al.*, 1986), and the resulting strain was used for plant transformation.

Generation and propagation of transgenic plants

The procedure for stable transformation was modified from that of Ishida *et al.* (1996). Immature zygotic embryos of Hi-II kernels were co-cultivated with *A. tumefaciens* strain EHA101 (Hood *et al.*, 1986) containing the vectors described above in the presence of the superbinary vector (Hood *et al.*, 2003). Bacteria were grown in yeast extract/peptone medium to an optical density at 600 nm ($OD_{600\text{ nm}}$) of 0.05. After mixing of embryos and bacteria, embryos were placed on to co-cultivation medium and incubated at 18 °C for 5 days. T_0 plants were regenerated from bialaphos-resistant tissue from culture and pollinated with pollen from elite inbred lines to produce T_1 seeds. Several ITEs produced seed (T_1) and were evaluated at the biochemical and molecular levels (Table 3).

Biochemical assays

Six dry seeds from each ear were individually pulverized and extracted with 1 mL of lysis buffer (50 mM sodium acetate, pH 5.0). Samples were placed in extraction tubes held in a rack, with a ball bearing added to each tube, and were then homogenized in a high-speed shaker for 20 s. Samples were centrifuged, and the supernatants were recovered and stored on ice prior to analysis. The supernatant was transferred to a 96-well plate and assayed for TSP (Bradford, 1976). Extract containing approximately 1 µg TSP was transferred to an assay plate (96-well) and the enzyme was assayed to determine the amount of activity. The E1 enzyme assay utilized the 4-methyl-umbelliferyl-β-1,4-D-cellobioside (MUC) substrate. The sensitivity of the assay ranged from approximately 0.1% TSP to approximately 10% TSP. The assay was performed at 50 °C. The CBH I enzyme assay also utilized the MUC substrate, and was also performed at 50 °C. This assay runs for approximately 2 h before a reading is taken to obtain optimal activity. The linear range was from approximately 0.5% TSP to approximately 10% TSP.

Western blot analysis

Cellulase samples were analysed by Western blot using either anti-E1 or anti-CBH I antibodies (provided by NREL) produced in mouse and rabbit. Two grams of E1 seed from BCH and BCF lines was extracted in 4 mL of 50 mM sodium acetate, pH 5.0. Approximately 100, 200 and 400 ng of total protein from the E1 lines was loaded per lane. E1 standard was loaded at approximately 15 or 30 ng (Figure 4a, lanes 7 and 8). CBH I seed extracts (1 µg protein) from each vector were loaded on to the gel. CBH I standard was loaded at 1, 10 or 100 ng.

Samples were run on a pre-cast Tris-HCl gel (Bio-Rad, Hercules, CA, USA). A BenchMark pre-stained protein marker ladder (Invitrogen, Carlsbad, CA, USA) was used for molecular weight estimates.

SDS-PAGE was run under standard denaturing conditions at a constant voltage of 150 mV in a Mini PROTEAN apparatus (Bio-Rad). The gel was transferred on to Immobilon-P polyvinylidene difluoride (PVDF) membrane (Millipore, Billerica, MA, USA) in an XCELL II Mini Cell blot module (Invitrogen) using tris(hydroxymethyl)aminomethane (Tris) buffer, pH 8.8, with 20% methanol at 35 mV for 1 h. The membrane was treated with 5% milk in Tris-buffered saline (TBS) for 1 h to block nonspecific binding sites, followed by incubation in 1 : 10 000 diluted anti-E1 produced in mouse. The membrane was treated with 5% milk a second time to avoid nonspecific interactions. The membrane was incubated in horseradish peroxidase (HRP)-labelled goat anti-mouse immunoglobulin G (IgG) diluted 1 : 10 000. The membrane was washed three times with TBS before each step, and all antibodies were prepared in TBS with 5% milk. The fluorescence detection dye ECL Plus (Amersham GE Healthcare, Piscataway, NJ, USA) was added according to the manufacturer's protocol. Excessive dye was removed with absorbent paper, and the membrane was placed in an autoradiography cassette and exposed for 5 s. The autoradiography film (Kodak X-OMAT AR, Kodak, Rochester, NY, USA) was developed using a Kodak Film processor 3000RA.

DNA hybridization analysis

Leaves from young plants were harvested, washed and frozen before grinding in liquid nitrogen, and extracted using a cetyltrimethylammonium bromide (CTAB)-based buffer according to <http://gmo-crl.jrc.it/detectionmethods/MON-Art47-dnaextraction.pdf>. Ten micrograms of each DNA sample was digested with *KpnI* [NEB (Ipswich, MA, USA), high concentration]. DNA was separated on a 0.8% agarose gel with Tris/acetate/ethylenediaminetetraacetic acid (EDTA) buffer, and electroblotted on to charged nylon (Millipore). Hybridization was performed in Church buffer (Church and Gilbert, 1985) [6% SDS, 1% bovine serum albumin (BSA), 1 mM EDTA, 0.5 M phosphate buffer] with ³²P random-prime labelled probe. DNA to be labelled was PCR amplified from the transformation plasmids containing the coding region, and included only the cellulase gene sequence. Reconstructions of 1, 5 and 10 copies per cell were generated from linearized plasmid and added to the blots to estimate the copy number. Blots were hybridized for 24 h at 65 °C, washed in 0.1% SDS, 0.1 × standard saline citrate (SSC) and exposed to X-OMAT film for 3 days.

Acknowledgements

We would like to thank Dr Steven R. Thomas, formerly of NREL, now of ceres, for providing E1 and CBH I clones, protein and antibodies, and for advice on enzyme assays. Maureen Dolan, Arkansas State University, is acknowledged for her expert assistance with the DNA hybridization experiments. This work was partially supported by United States Department of Agriculture Small Business Innovation Research (USDA SBIR) grant no. 2003-00561.

References

Aden, A., Ruth, M., Ibsen, K., Jechura, J., Neeves, K., Sheehan, J. and Wallace, B. (2002) *Lignocellulosic Biomass to Ethanol Process*

Design and Economics Utilizing Co-Current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis for Corn Stover. URL <http://www.eere.energy.gov/biomass/progs/searchdb2.cgi?6483> [accessed on 10 December 2006].

- An, G., Thornburg, R., Johnson, R., Hall, G. and Ryan, C.A. (1987) A possible role for 3' sequences of the wound-inducible potato proteinase inhibitor IIK gene in regulating gene expression. *NATO Adv. Sci. Inst. Series, Ser. a Life Sci.* **140**, 377–383.
- Anzai, H., Yoneyama, K. and Yamaguchi, I. (1989) Transgenic tobacco resistant to a bacterial disease by the detoxification of a pathogenic toxin. *Mol. Gen. Genet.* **219**, 492–494.
- Baker, J., Ehrman, C., Adney, W., Thomas, S. and Himmel, M. (1998) Hydrolysis of cellulose using ternary mixtures of purified cellulases. *Appl. Biochem. Biotechnol.* **70–72**, 395–403.
- Belanger, F.C. and Kriz, A.L. (1991) Molecular basis for allelic polymorphism of the maize Globulin-1 gene. *Genetics*, **129**, 863–872.
- Biswas, G.C.G., Ransoma, C. and Sticklen, M. (2006) Expression of biologically active *Acidothermus cellulolyticus* endoglucanase in transgenic maize plants. *Plant Sci.* **171**, 617–623.
- Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Anal. Biochem.* **72**, 248–254.
- Buchanan, B.B., Gruissem, W. and Jones, R.L. (2002) *Biochemistry and Molecular Biology of Plants*. Rockville, MD: American Society of Plant Biologists.
- Church, G.M. and Gilbert, W. (1985) The genomic sequencing technique. *Prog. Clin. Biol. Res.* **177**, 17–21.
- Dai, Z., Hooker, B., Anderson, D.B. and Thomas, S.R. (2000) Expression of *Acidothermus cellulolyticus* endoglucanase E1 in transgenic tobacco: biochemical characteristics and physiological effects. *Transgenic Res.* **9**, 43–54.
- Daniell, H., Streatfield, S.J. and Wycoff, K. (2001) Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. *Trends Plant Sci.* **6**, 219–226.
- Gordon-Kamm, W.J., Spencer, T.M., Mangano, M.L., Adams, T.R., Daines, R.J., Start, W.G., O'Brien, J.V., Chambers, S.A., Adams, W.R. Jr, Willetts, N.G., Rice, T.B., Mackey, C.J., Krueger, R.W., Kausch, A.P. and Lemaux, P.G. (1990) Transformation of maize cells and regeneration of fertile transgenic plants. *Plant Cell*, **2**, 603–618.
- Hiei, Y., Ohta, S., Komari, T. and Kumashiro, T. (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis. *Plant J.* **6**, 271–282.
- Holwerda, B.C., Padgett, H.S. and Rogers, J.C. (1992) Proaleurain vacuolar targeting is mediated by short contiguous peptide interactions. *Plant Cell*, **4**, 307–318.
- Hood, E.E. and Woodard, S. (2002) Industrial proteins produced from plants. In: *Plants as Factories for Protein Production* (Hood, E.E. and Howard, J.A., eds), pp. 119–135. Dordrecht: Kluwer Academic Publishers.
- Hood, E.E., Helmer, G.L., Fraley, R.T. and Chilton, M.D. (1986) The hypervirulence of *Agrobacterium tumefaciens* A281 is encoded in a region of pTiBo542 outside of T-DNA. *J. Bacteriol.* **168**, 1291–1301.
- Hood, E.E., Witcher, D.R., Maddock, S., Meyer, T., Baszczyński, C., Bailey, M.R., Flynn, P., Register, I. James, M.L., Bond, D., Kulisek, E., Kusnadi, A., Evangelista, R., Nikolov, Z., Wooge, C., Mehig, R.J., Hernan, R., Kappel, W.K., Ritland, D., Li, C.-P. and Howard, J.A. (1997) Commercial production of avidin from transgenic

- maize: characterization of transformant, production, processing, extraction and purification. *Mol. Breed*, **3**, 291–306.
- Hood, E.E., Bailey, M.R., Beifuss, K., Magallanes-Lundback, M., Horn, M.E., Callaway, E., Drees, C., Delaney, D.E., Clough, R. and Howard, J.A. (2003) Criteria for high-level expression of a fungal laccase gene in transgenic maize. *Plant Biotechnol. J.* **1**, 129–140.
- Howard, J.A. and Hood, E.E. (2005) Bioindustrial and biopharmaceutical products produced in plants. In: *Advances in Agronomy* (Sparks, D., ed.), pp. 91–124. San Diego, CA: Academic Press.
- Ishida, Y., Saito, H., Ohta, S., Hiei, Y., Komari, T. and Kumashiro, T. (1996) High efficiency transformation of maize (*Zea mays* L.) mediated by *Agrobacterium tumefaciens*. *Nat. Biotechnol.* **14**, 745–750.
- Larrick, J., Yu L., Naftzger, C., Jaiswal, S. and Wycoff, K. (2002) Human pharmaceutical products in plants. In: *Plants as Factories for Protein Production* (Hood, E.E. and Howard, J.A., eds), pp. 79–101. Dordrecht: Kluwer Academic Publishers.
- Mohagheghi, A., Grohmann, K., Himmel, M., Leighton, L. and Updegraff, D.M. (1986) Isolation and characterization of *Acidothermus cellulolyticus* gen.nov., sp.nov., a new genus of thermophilic, acidophilic, cellulolytic bacteria. *Int. J. Syst. Bacteriol.* **36**, 435–443.
- Nieves, R.A., Chou, Y.-C., Himmel, M.E. and Thomas, S.R. (1995) Quantitation of *Acidothermus cellulolyticus* E1 endoglucanase and *Thermonospora fusca* E[3] exoglucanase using enzyme-linked immunosorbent assay (ELISA). In: *Symposium on biotechnology for fuels and chemicals* (Davidson, B.H. and Finkelstein, M., eds), pp. 211–223. Totowa: Humana Press.
- Rogers, J.C. (1985) Two barley alpha-amylase gene families are regulated differently in aleurone cells. *J. Biol. Chem.* **260**, 3731–3738.
- Schulein, M. (2000) Protein engineering of cellulases. *Biochim. Biophys. Acta*, **1543**, 239–252.
- Shoemaker, S., Schweickart, V., Ladner, M., Gelfand, D., Kwok, S., Myambo, K. and Innis, M. (1983) Molecular cloning of exocellobiohydrolase I derived from *Trichoderma reesei* strain L27. *Bio/Technology*, **1**, 691–696.
- Streatfield, S.J. (2007) Approaches to achieve high-level heterologous protein production in plants. *Plant Biotechnol. J.* **5**, 2–15.
- Streatfield, S.J., Jilka, J.M., Hood, E.E., Turner, D.D., Bailey, M.R., Mayor, J.M., Woodard, S.L., Beifuss, K.K., Horn, M.E., Delaney, D.E., Tizard, I.R. and Howard, J.A. (2001) Plant-based vaccines: unique advantages. *Vaccine*, **19**, 2742–2748.
- Streatfield, S.J., Mayor, J.M., Barker, D.K., Brooks, C., Lamphear, B.J., Woodard, S.L., Beifuss, K.K., Vicuna, D., Massey, L.-A., Horn, M.E., Delaney, D.E., Nikolov, Z.L., Hood, E.E., Jilka, J.M. and Howard, J.A. (2002) Development of an edible subunit vaccine in corn against enterotoxigenic strains of *Escherichia coli*. *In Vitro Cell. Dev. Biol. Plant*, **38**, 11–17.
- Tomme, P., Warren, R.A. and Gilkes, N.R. (1995) Cellulose hydrolysis by bacteria and fungi. *Adv. Microb. Physiol.* **37**, 1–81.
- Uchimiya, H., Iwata, M., Nojiri, C., Samarajeewa, P.K., Takamatsu, S., Ooba, S., Ansai, H., Christensen, A.H., Quail, P.H. and Toki, S. (1993) Bialaphos treatment of transgenic rice plants expressing a bar gene prevents infection by the sheath blight pathogen (*Rhizoctonia solani*) *Bio/Technology*, **11**, 835–836.
- Walsh, M.E. (1999) *Biomass Feedstock Availability in the United States, State Level Analysis*. URL <http://bioenergy.ornl.gov/resourcedata/index.html> [accessed on 15 November 2006].
- White, J., Chang, S.Y., Bibb, M.J. and Bibb, M.J. (1990) A cassette containing the bar gene of *Streptomyces hygroscopicus*: a selectable marker for plant transformation. *Nucleic Acids Res.* **18**, 1062.
- Witcher, D.R., Hood, E.E., Peterson, D., Bailey, M., Bond, D., Kusnadi, A., Evangelista, R., Nikolov, Z., Wooge, C., Mehig, R., Kappel, W., Register, I., James, C. and Howard, J.A. (1998) Commercial production of β -glucuronidase (GUS). A model system for the production of proteins in plants. *Mol. Breed.* **4**, 301–312.
- Woodard, S.L., Mayor, J.M., Bailey, M.R., Barker, D.K., Love, R.T., Lane, J.R., Delaney, D.E., McComas-Wagner, J.M., Mallubhotla, H.D., Hood, E.E., Dangott, L.J., Tichy, S.E. and Howard, J.A. (2003) Maize (*Zea mays*)-derived bovine trypsin: characterization of the first large-scale, commercial protein product from transgenic plants. *Biotechnol. Appl. Biochem.* **38**, 123–130.
- Zhong, G.-Y., Peterson, D., Delaney, D.E., Bailey, M.R., Witcher, D.R., Register, J. III, Bond, D., Li, C.-P., Marshall, L., Kulisek, E., Ritland, D., Meyer, T., Hood, E.E. and Howard, J.A. (1999) Commercial production of aprotinin in transgenic maize seeds. *Mol. Breed.* **5**, 345–356.
- Ziegelhoffer, T., Will, J. and Austin-Phillips, S. (1999) Expression of bacterial cellulase genes in transgenic alfalfa (*Medicago sativa* L.), potato (*Solanum tuberosum* L.) and tobacco (*Nicotiana tabacum* L.). *Mol. Breed.* **5**, 309–318.
- Ziegler, M., Thomas, S.R. and Danna, K. (2000) Accumulation of a thermostable endo-1,4- β -D-glucanase in the apoplast of *Arabidopsis thaliana*. *Mol. Breed.* **6**, 37–46.