Journal of Plant Biology, February 2007, 50(1) : 29-37

Aptamers Improve the Expression of a Human Granulocyte-Macrophage Colony Stimulating Factor in Transgenic *Arabidopsis thaliana* Seeds

Biao Wang, Ming Ma, and Tian Long Wu*
Plant Science and Technology Department, Shanghai Jiao Tong University, Shanghai 201101, China

Domain IV nucleotides of human 7SL RNA comprise an approximately 50-nucleotide region with a highly conserved secondary structure consisting of two internal loops. We inserted these into a modified 3'-untranslated region of the human granulocyte-macrophage colony stimulating factor (hGM-CSF) gene. The expression vector, under the control of a tissue-specific promoter derived from the soybean α' subunit of β-conglycinin, was introduced into *Arabidopsis thaliana* by *Agrobacterium*-mediated transformation. Expression of the recombinant hGM-CSF significantly improved, accounting for as much as 0.049% of the total soluble protein in the *Arabidopsis* siliques. Western blots showed that the molecular weight of this recombinant hGM-CSF was approximately 21 kD. In addition, TF-1 cell proliferation data demonstrated that the recombinant hGM-CSF was biologically active. Our results provide evidence that aptamers can strongly enhance gene expression.

Keywords: aptamers, human GM-CSF, TF-1 cell

When genetically engineering plants to produce pharmaceuticals, it is crucial to improve the expression of foreign genes (Kim et al., 2006). Synthesis of non-native proteins may be enhanced by an increase in transgene copies or the insertion of regulated sequences (Roh et al., 2006). The aptamer regulates gene expression at the level of transcription or translation through DNA/RNA aptamer-small molecule interactions (Gossen and Bujard, 1992; Joel et al., 1997; Grate and Wilson, 2001). In plants, these double-stranded DNA or single-stranded RNA molecules are capable of binding specific molecular targets with high affinity and specificity *in vitro* or *in vivo* (Bock et al., 1992). They also generally repress *in vitro* expression in mammalian cells if inserted into the 5'-untranslated region (5'-UTR) of the gene (Werstuck and Green, 1998). In *Arabidopsis thaliana*, the thiamine pyrophosphate-binding motif resides in the 3'-UTR of a thiamine biosynthetic gene, and the isolated RNA domain binds the corresponding coenzyme *in vitro*. These observations suggest that metabolite-binding mRNAs are involved in eukaryotic gene regulation (Sudarsan et al., 2003).

Aptamer behavior can be regulated by specific small molecules, which leads to the development of more sophisticated ligand-regulated aptamers (Sen, 2002). The signal recognition particle (SRP), a well-characterized eukaryotic protein containing two main components -- 7SL RNA and SRP54 -- plays a key role in transporting nascent secretory and membrane proteins. The nucleic acid element of 7SL RNA displays a tight functional relationship between prokaryotes and eukaryotes (Bernstein et al., 1989; Althoff et al., 1994; Keenan et al., 2001). In the latter type, domain IV of human 7SL RNA is an approximately 50-nucleotide (nt) region in which the highly conserved secondary structure consists of two internal loops (Schmitz et al., 1999a, b). Domain IV nucleotides might be a natural aptamer that could be bound to the SRP54 protein *in vivo* (Batey et al., 2000).

Human granulocyte-macrophage colony stimulating factor (hGM-CSF) is a cytokine used in the treatment of neutropenia and aplastic anemia, and greatly reduces the risk of infection after bone marrow transplantation by accelerating the response of neutrophils (Diederichs et al., 1991; Hercus et al., 1994). In the human genome, the 2.5-kb GM-CSF gene is present as a single copy, containing three introns and four exons (Kaushansky et al., 1986). cDNA of human GM-CSF has been isolated from concanavalin A-activated human T-cell clones; the mature human GM-CSF encodes a protein of 144 amino acids with 69% nucleotide homology and 54% amino acid homology to mouse GM-CSF (Lee et al., 1985). This protein has a predicted molecular mass of 16293 daltons, with its actual molecular weight depending upon the level of glycosylation. An adenosine-uridine-rich element in the 3'-UTR may affect the stability of human GM-CSF mRNA (Jarzembowski et al., 1999), which has an expected half-life of approximately 90 min *in vitro* (Rajagopalan and Malter, 1994).

Recombinant human GM-CSF protein has been produced in such plants as tomato (Kwon et al., 2003), tobacco (Sardana et al., 2002), rice (Shin et al., 2003), and sugarcane (Wang et al., 2005). However, the high cost and low deposition of proteins in harvestable organs are still limiting factors in exploiting plants as biofactories for recombinant hGM-CSF. Chung Il et al. (2006) have reported the extensive presence and importance of small noncoding RNAs in eukaryote gene regulation. Insertion of

*Corresponding author; fax +86-21-64785940
e-mail tianlongwu@263.net

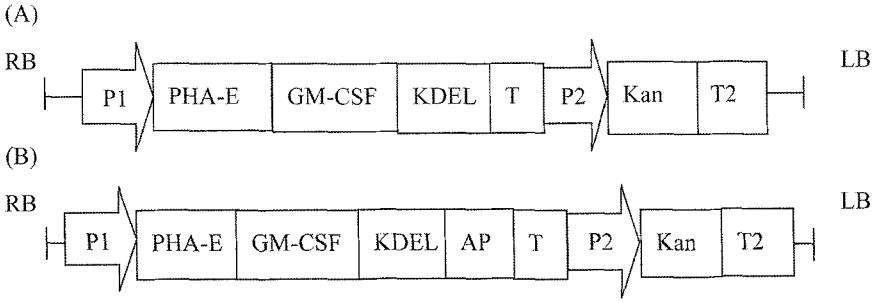
Abbreviations: DAB, 3,3'-diaminobenzidine tetrahydrochloride; DAP, days after pollination; ER, endoplasmic reticulum; hGM-CSF, human granulocyte-macrophage colony stimulating factor; KDEL, lys-asp-glu-leu; NOS, nopaline synthase terminator; SRP, signal recognition particle.

a pre-selected aptamer into the 5′-UTR of an encoding mRNA not only allows its translation to be repressed by the addition of ligands *in vivo* but also reduces the translation efficiency of targets of interest and alters reporter gene expression by riboswitches (Werstuck and Green, 1998; Suess et al., 2003; Hanson et al., 2005). Aptamers also improve translation initiation by introducing a synthetic 15 nt RNA stem-loop downstream from the start codon, thereby enhancing gene expression (Paulus et al., 2004). Here, we added the endoplasmic reticulum (ER) retention signal lys-asp-glu-leu (KDEL) to the C-terminus of human GM-CSF. Nucleotide sequences from domain IV of human 7SL RNA were also introduced into the 3′-UTR of the hGM-CSF gene, and AU-rich elements at its 3′-UTR were replaced by nopaline synthase poly(A) signal sequences. The modified hGM-CSF gene was also cloned into a plant expression vector and transformed into *Arabidopsis thaliana*. Our objective was to investigate the accumulation of recombinant hGM-CSF in transformants.

MATERIALS AND METHODS

DNA Cloning and Plasmid Reconstruction

The phytohemagglutinin (PHA-E) signal peptide was conjugated to the mature hGM-CSF coding region via PCR, using two forward primers: CSF-F1 (5′-CCCTCTTCCTTG-TGCTTCTCACCCACGCAAACTCAGCACCCGCCCGCTCG-3′) and CSF-F (5′-gtgcacAATGGCTTCCTCCAACTTACTCTC-CCTAGCCCTCTTCCTTGTGC-3′), and a reverse primer: CSF-K (5′-TCATCAAAGCTCATCCTTCTCCTGGACTGG-3′). PCR amplification from the plasmid pET28a(+)-CSF, which contained the cDNA sequences of hGM-CSF, was completed with primers CSF-F1 and CSF-K. PCR conditions included initiation at 94°C, 5 min; then 35 cycles of 94°C, 1 min; 65°C, 1 min; 72°C, 35 s; and a final extension at 72°C, 8 min. After cloning into a pMD18-T vector (Takara, Japan) and sequencing, the DNA product was used as template for PCR amplification with primers CSF-F and CSF-K, following



(C)

ATGGCTTCCT CCAACTTACT CTCCCTAGCC CTCTTCCTTG TGCTTCTCAC CCACGCAAAC 60

TCAGCACCCG CCCGCTCGCC CAGCCCCAGC ACGCAGCCCT GGGAGCATGT GAATGCCATC 120

CAGGAGGCCC GGCGTCTCCT GAACCTGAGT AGAGACACTG CTGCTGAGAT GAATGAAACA 180

GTAGAAGTCA TCTCAGAAAT GTTTGACCTC CAGGAGCCGA CCTGCCTACA GACCCGCCTG 240

GAGCTGTACA AGCAGGGCCT GCGGGGCAGC CTCACCAAGC TCAAGGGCCC CTTGACCATG 300

ATAGCCAGCC ACTACAAGCA GCACTGCCCT CCAACCCCGG AAACTTCCTG TGCAACCCAG 360

ATTATCACCT TTGAAAGTTT CAAAGAGAAC CTGAAGGACT TTCTGCTTGT CATCCCCTTT 420

GACTGCTGGG AGCCAGTCCA GGAGAAGGAT GAGCTTTGAT GAGGGGTGAA CCGGCCCAGG 480

TCGGAAACGG AGCAGGTCAA AACTCC 506

Figure 1. Schematic maps of vector constructs and recombinant hGM-CSF nucleotide sequences. (**A**) Map of pCAMBIA2300-α′-CSF-KDEL-NOS, (**B**) map of pCAMBIA2300-α′-CSF-AP-NOS. P1, soybean α′ subunit promoter of β-conglycinin; P2, CaMV35S promoter; PHA-E, phytohemagglutinin signal peptide (63 bp); GM-CSF, mature human granulocyte-macrophage colony stimulating factor coding region (384 bp); KDEL, lysine-aspartic acid-glutamic acid-leucine tetrapeptide; AP, domain IV nucleic acid sequences of human 7SL RNA (44 bp); T, NOS terminator; Kan, Neomycin phosphotransferase gene; T2, CaMV35S poly(A); RB, T-DNA right border; LB, T-DNA left border. (**C**) Recombinant GM-CSF nucleotide sequence from PHA-E to AP in (**B**). Single line, PHA-E sequence; Double line, mature hGM-CSF coding region sequence; Dotted line, KDEL sequences; Bold single line, termination codon sequences; Dashed line, AP sequence.

similar PCR protocols. The PCR fragment was then ligated into a pMD18-T vector to generate the pMD18-CSF-KDEL plasmid, which contained the phytohemagglutinin signal peptide, the mature hGM-CSF coding domain, and KDEL sequences. Domain IV of human 7SL RNA nucleotide sequences was introduced into the 3′-UTR of the hGM-CSF gene by PCR, with the forward primer: CSF-F, and two reverse primers: 5′-CCGACCTGGGCCGGTTCACCCCTCAT-CAAAGCTCATCC-3′ (AP1) and 5′-tggtaccGGAGTTTTGACCT-GCTCCGTTTCCGACCTGGGCCGG-3′ (AP2). Plasmid pMD18-CSF-KDEL, as template, contained a phytohemagglutinin (PHA-E) signal peptide and mature hGM-CSF coding sequences tagged with the ER retention signal, KDEL. PCR was performed with the CSF-F and AP1 primers, under the following conditions: 94°C for 5 min; followed by 35 cycles of 94°C for 1 min, 68°C for 30 s, 72°C for 35 s; with a final extension at 72°C for 8 min. Products were then cloned into a pMD18-T vector and sequenced. The DNA product was used as template for PCR using primers CSF-F and AP2 and similar cycling conditions. Afterward, the PCR fragment was ligated into a pMD18-T vector to generate the plasmid pMD18-CSF-AP, which contained the PHA-E signal peptide, the mature hGM-CSF coding sequences with the KDEL ER retention signal, and 44 nucleotides of the human 7SL RNA domain IV within the hGM-CSF 3′-UTR. Both plasmids were digested with *Eco*RI and *Sal*I, blunted with T4 DNA polymerase, and sub-cloned into the *Sma*I site of the pCAMBIA2300-α′-NOS (unpublished) expression vector. This vector contained a tissue-specific promoter derived from the soybean α′ subunit of β-conglycinin and the nopaline synthase terminator (NOS). Figure 1 shows detailed sequences of the recombinant hGM-CSF gene and schematic constructs of plasmids pCAMBIA2300-α′-CSF-KDEL-NOS and pCAMBIA2300-α′-CSF-AP-NOS.

Transformation and Regeneration

The pCAMBIA2300-α′-CSF-KDEL-NOS and pCAMBIA2300-α′-CSF-AP-NOS expression vectors were introduced into *Agrobacterium tumefaciens* GV3101, and floral-dip protocols were used to transform *A. thaliana* (Clough and Bent, 1998). Sterilized transformed seeds were grown on 0.8% agar plates with a 0.5 Murashige-Skoog medium containing 30 μg mL^{-1} kanamycin for selection. They were first subjected to cold treatment for 2 d, then placed under a 16-h photoperiod (100 μE $m^{-2}s^{-1}$) for 7 to 10 d. As the transformed plants matured, T1 plants were self-pollination and homozygous lines were identified from the progeny for further analysis. In all, 27 KD (from pCAMBIA2300-α′-CSF-KDEL-NOS) and 5 AP (from pCAMBIA2300-α′-CSF-AP-NOS) independently transformed lines were used to analyze the expression of hGM-CSF.

Tissue-Specific Expression and Semi-Quantitative RT-PCR

Reverse transcriptase-PCR (RT-PCR) was used to analyze the expression profile of hGM-CSF in different tissues. Roots, stems, mature leaves, and seeds at 6 d after pollination (DAP) that were derived from KD6 transformants were investigated to confirm that expression was limited to the seed. Total RNA extraction was performed via SDS, with several modifications: the extraction buffer contained 0.18 M Tris, 0.09 M LiCl, 4.5 mM EDTA, 1% (m/V) SDS (pH 8.2), 2% polyvinylpyrrolidone (average mol. wt. 40,000) and 10 mM cysteine. Total RNA was subjected to DNaseI digestion (RNase-free, 30 min, 37°C; Takara, Japan), then purified on a mini-RNA extraction column (Watson Biotechnologies, China) according to the manufacturer's instructions. The *A. thaliana* 2S albumin gene (*at2S2*, Accession number M22034), which is expressed only in the seed (Guerche et al., 1990), served as a control to confirm that no trace genomic DNA contamination was present after the DNaseI digestion. Amplification was performed with a one-step RT-PCR kit (Takara, Japan) and two consensus hGM-CSF primers: CSF-F and CSF-K. RT-PCR cycling parameters included an initial RT reaction (50°C, 30 min) that was stopped by heating (94°C, 2 min). This was followed by *Taq* polymerase activation (95°C, 3 min) and 30 cycles of 95°C for 30 s, 65°C for 30 s, and 72°C for 1 min. The forward primer 5′-AGTAAGAATACAAACACAAATAGC-3′ and reverse primer 5′-CGCTCGAGGGTTTTTGTTTGGAATC-3′ were used to detect *at2S2* expression with a PCR annealing temperature of 55°C. We conducted semi-quantitative RT-PCR to investigate GM-CSF mRNA levels, with *actin2* (At3g18780) as a control, in siliques harvested at 3, 6, 9, and 12 DAP. Total RNA (1 μg, 0.5 μg μL^{-1}) was added to a 50 μL reaction volume containing 5 μL of 10×one step RNA PCR buffer, 10 μL $MgCl_2$ (25 mM), 5 μL dNTP mixture (10 mM), 1 μL RNase inhibitor (40 U μL^{-1}), 1 μL AMV RTase XL (5 U μL^{-1}), 1 μL AMV-Optimized Taq (5 U μL^{-1}), 1 μL forward primer GM-CSF-F (10 μM), 1 μL reverse primer CSF-K (10 μM), and 23 μL RNase-free dH_2O. The mixture was incubated (50°C, 30 min), and RTase activity was stopped by heating (94°C, 2 min). PCR amplification of the hGM-CSF gene was performed for 20 or 25 cycles of 94°C for 30 s, 65°C for 30 s, and 72°C for 1 min. Primers 5′-TAACTCTCCCGCTATG-TATGTC-3′ (ACTIN2F) and 5′-GAGGTAATCAGTAAGGT-CACG-3′ (ACTIN2R) were used to amplify the *actin2* gene for 25 cycles, at an annealing temperature of 55°C. The PCR products were fractionated on a 1.5% agarose gel.

Human GM-CSF-Specific ELISA and Western Blotting

Siliques at 10 to 12 DAP were collected and stored at -70°C. Crude protein extracts were prepared by grinding approximately 100 mg of sample in 0.5 mL extraction buffer containing 50 mM Tris (pH 7.5), 50 mM NaCl, 1 mM EDTA, 1 mM phenylmethylsulfonylfluoride, 1% β-mercaptoethanol, 0.1% Triton X-100, 1% ascorbic acid, and 1% polyvinylpyrrolidone. Extracts were centrifuged (14000*g*, 4°C, 20 min), and the supernatants were transferred to new tubes and centrifuged again (20800*g*, 4°C, 2 min) to remove the remaining debris. Total soluble protein was analyzed by the Bradford (1976) method, with bovine serum albumin as the standard. Enzyme-linked immunosorbent assay (ELISA) protocols have been previously published (Ausubel et al., 2001). Rabbit anti-hGM-CSF antibody (Cytolab, Israel) and goat anti-rabbit-IgG-HRP (Sino-American Biotechnology, China) were used to detect protein levels, with commercial recombinant human GM-CSF produced by *Escherichia coli* (Cytolab, Israel) serving as a standard. Crude protein extracted

from siliques at 10 to 12 DAP were fractionated on 15% SDS polyacrylamide gels and transferred onto polyvinyldifluoride membranes (Osmonics, USA). Proteins were detected with a polyclonal antibody to hGM-CSF and horseradish peroxidase-conjugated goat anti-rabbit IgG, and were developed with a DAB (3,3'-diaminobenzidine tetrahydrochloride) substrate system (Sino-American Biotechnology).

Quantitative Real-Time PCR

A. thaliana siliques were harvested at 10 to 12 DAP to measure the levels of hGM-CSF mRNA. PCR was performed in a Rotor-gene RG-3000A, using SYBR GreenI® dye to monitor DNA synthesis. DNaseI-treated mRNA served as template to synthesize first-strand cDNA. Primer pairs were designed to amplify a 150- to 200-bp fragment, and expression levels were compared against the *actin2* gene standard. A PCR reaction volume of 25 μL contained the first-strand cDNA, DNA polymerase (Takara), the corresponding primer pairs (40 nM each), dNTP mix (200 nM each), $MgCl_2$ (3.0 mM), and 0.8 μL 20× SYBR GreenI® fluorescent dye. After the polymerase was heat-activated at 95°C for 3 min, we performed 35 cycles of denaturation (95°C, 30 s), annealing (60°C, 30 s), and amplification (72°C, 45 s). Melting curves for each PCR reaction were calculated by measuring the decrease in fluorescence while the temperature increased from 55 to 95°C. The specificity of PCR reactions was confirmed by analyzing those curves with Rotor-Gene 6 software, as well as through agarose gel electrophoresis of the products. Threshold cycles at which the fluorescence of the PCR product SYBR® complex first exceeded the background level were determined according to integrated analysis software for each gene. R^2-values were >0.95. The slope (M) of a reaction was used to determine the exponential amplification and efficiency of a reaction, with a value of approx. -3. Relative template concentrations were evaluated based on the standard curve for the hGM-CSF and *actin2* genes. Each PCR run was carried out three times in duplicate.

GM-CSF Biological Activity Assay

Cell proliferation was measured by the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay (Kasugai et al., 1990). To avoid suppression of TF-1 cell proliferation by the protein extraction buffer, we used 1×PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 2 mM KH_2PO_4 pH 7.4), containing 0.1% Tween-20, to extract soluble proteins from the seeds. TF-1 cells were maintained (37°C, 5% CO_2) in an RPMI-1640 medium supplemented with 10% fetal bovine serum and *E. coli* hGM-CSF (2 ng mL^{-1}). Cells were washed twice with 1×PBS, resuspended (2×10^5 mL^{-1}) in an RPMI-1640 medium with fetal bovine serum, then transferred (2×10^4 per well) to tissue culture plates for the proliferation assays. An RPMI-1640 medium (0.5 mL) with 10% fetal bovine serum was added to wells containing one of the following samples: *E. coli*-derived commercial GM-CSF (2 ng mL^{-1}); silique extracts from transgenic KD lines (10 to 12 DAP) containing 2 ng mL^{-1} hGM-CSF; silique extracts from transgenic AP plants (10 to 12 DAP) containing 2 ng mL^{-1} hGM-CSF; silique extracts (10 to 12 DAP) from non-transformed (NT) plants; silique extracts (10 to 12 DAP) from non-transformed (NT) plants with 2 ng mL^{-1} commercial GM-CSF; seed protein extraction buffer (EB); and seed protein EB containing 2 ng mL^{-1} of commercial hGM-CSF. Cells were cultured for 72 h, then incubated with MTT (20 μL per well, 37°C, 4 h) before dissolution in 100 μL lysis buffer that contained 2% SDS and 50% dimethylsulfoxide (pH 4.7). Absorbance (570 nm) was determined on a plate reader after 1 h. All experiments were repeated in triplicate (N=3 per treatment) under sterile conditions.

Data Analysis

F-tests for variance and a student's *t*-test analysis for the means of pooled data were conducted with SAS version 8.0 (SAS, 2000) to analyze our hGM-CSF accumulation data. Analysis of variance within the cell proliferation data utilized *q* tests of the mean values.

RESULTS

Identification of Independent Transgenic Lines

The constructs used to transform *A. thaliana* are shown in Figure 1A and B. Many AP transformants in the T1 generation produced no seeds but were otherwise phenotypically normal, more so than the KD transformants. The GM-CSF gene was amplified from 27 KD and 5 AP independent transgenic plants from the T3 generation, using primers CSF-F and CSF-K. PCR results were also confirmed by Southern analysis of the transformed plants. In all, 5 of 27 KD transgenic lines showed unique hybridization patterns, and all AP transformed plants displayed positive bands.

Tissue-Specific Expression and Semi-Quantitative RT-PCR

To determine whether expression of the transgene was tissue-specific, we analyzed levels of hGM-CSF transcripts in roots, stems, leaves, and seeds. No PCR products were amplified from the first three tissue types, but strong hGM-CSF mRNA expression was noted in the seeds. The *A. thaliana* albumin *at2S2* gene also was expressed only in the seed, with profiles similar to those for the *at2S2* control, suggesting that the total RNA contained no contaminating genomic DNA (Fig. 2A). Our assay data suggested that, in the developing seeds, hGM-CSF was actively transcribed and the soybean α' subunit of the β-conglycinin promoter was preferentially activated. We further investigated hGM-CSF mRNA levels in the early and late stages of seed development, using siliques at 3, 6, 9, and 12 DAP and *actin2* as a control. In general, weaker bands were noted in 3 DAP siliques compared with those from later on, and 12 DAP siliques still displayed strong hGM-CSF expression (Fig. 2B). These results suggest that the soybean α' subunit of the β-conglycinin promoter was activated during early seed development and became stable in later stages (Chen et al., 1986).

ELISA and Western Blots

Recombinant hGM-CSF protein was measured in the KD

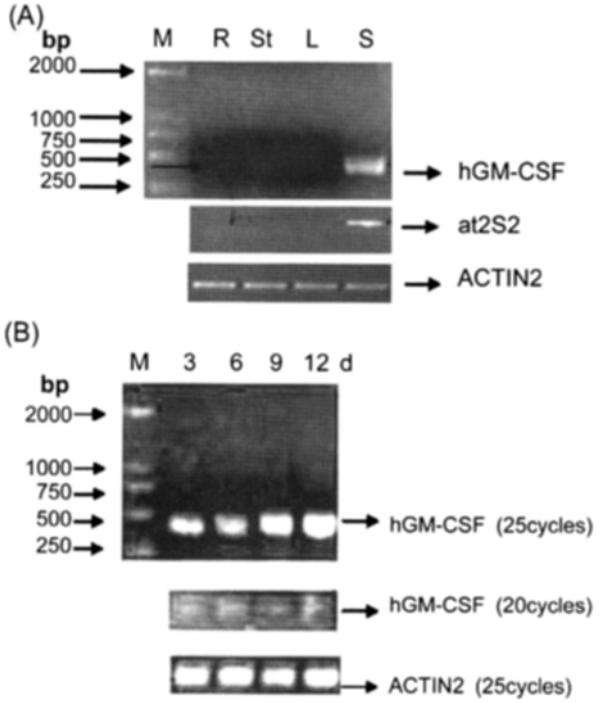


Figure 2. Detection of hGM-CSF mRNA and seed developmental stages. (**A**) KD6 transformants were used to monitor tissue-specific expression of GM-CSF mRNA with gene-specific primers GM-CSF-F and GM-CSF-K. *A. thaliana* seed-specific albumin gene at2S2 was used as control for genomic DNA contamination. M, DL2000 marker; R, St, L, and S, total RNA from roots, stems, leaves and seeds, respectively. (**B**) Total cellular RNA from KD6 siliques at 3, 6, 9, and 12 d DAP was digested with DNaseI and subjected to semi-quantitative RT-PCR and gel electrophoresis. The *actin2* gene served as internal control. M, DL2000 marker; 3, 6, 9, and 12 d, number of days after pollination.

Table 1. Expression levels of recombinant hGM-CSF from developing siliques at 10 to 12 d after pollination (DAP).

Plant	Seed age/DAP	hGM-CSF as % of total soluble protein
KD1	10-12	0.012%
KD2	10-12	<0.01%
KD3	10-12	0.021%
KD5	10-12	<0.01%
KD6	10-12	0.026%
KD7	10-12	<0.01%
KD8	10-12	0.028%
KD11	10-12	<0.01%
KD12	10-12	0.017%
KD13	10-12	<0.01%
KD15	10-12	0.010%
KD16	10-12	<0.01%
KD17	10-12	<0.01%
KD18	10-12	0.019%
KD19	10-12	<0.01%
KD20	10-12	0.026%
KD22	10-12	<0.01%
KD23	10-12	0.020%
KD24	10-12	<0.01%
KD26	10-12	<0.01%
KD27	10-12	0.015%
KD28	10-12	0.011%
KD30	10-12	<0.01%
KD31	10-12	<0.01%
KD32	10-12	0.017%
KD33	10-12	0.016%
KD35	10-12	<0.01%
AP1	10-12	0.047%
AP5	10-12	0.031%
AP8	10-12	0.049%
AP12	10-12	0.020%
AP17	10-12	<0.01%
NT	10-12	0.0

Percentage data are mean values (from 3 independent experiments, each in duplicate). KD, lines transformed with the plasmid pCAMBIA2300-α'-CSF-KDEL-NOS; AP, lines transformed with the plasmid pCAMBIA2300-α'-CSF-AP-NOS; NT, non-transformed plants.

and AP transgenic lines by the ELISA method. KD lines varied in their expression level of hGM-CSF, whereas the amount of recombinant protein was very high in AP transgenic plants, except for the AP17 transformants. Of the five AP independent transgenic lines, four had recombinant protein accumulations far higher than the mean expressed in the KD transformed lines. For example, the AP1, AP5, and AP8 lines showed greater expression than the KD6 transformants (0.026% of the total soluble protein). The highest level accounted for up to 0.049% of the total soluble protein in siliques, and was even greater in the seed (Table 1). Five independent KD transgenic lines containing single copies of the hGM-CSF gene displayed high accumulations of hGM-CSF, with multi-copy KD lines varying in their expression of hGM-CSF and exhibiting far less expression than from AP8. Recombinant hGM-CSF accumulations in our AP-transformed lines were generally more abundant than in the KD lines, indicating that the insertion of aptamers caused hGM-CSF expression to be elevated by mechanisms other than gene copy number, etc.

The molecular weight of GM-CSF from KD- and AP-transformed plants was approximately 21 kD (Fig. 3), which is similar to that of recombinant hGM-CSF produced by transgenic tobacco seeds (Sardana et al., 2002). A main band at the same molecular weight was observed through western blotting in both KD and AP transformants, suggesting that the PHA-E signal peptide was correctly removed and that the recombinant hGM-CSF had been targeted to storage vesicles. Because the molecular mass of natural human GM-CSF is approximately 16 kD (Cantrell et al., 1985), we propose that the high-molecular-weight band observed in our study might represent a glycosylated form of the recombinant hGM-CSF protein.

Quantitative Real-Time PCR

To ascertain the levels of hGM-CSF mRNA expressed in

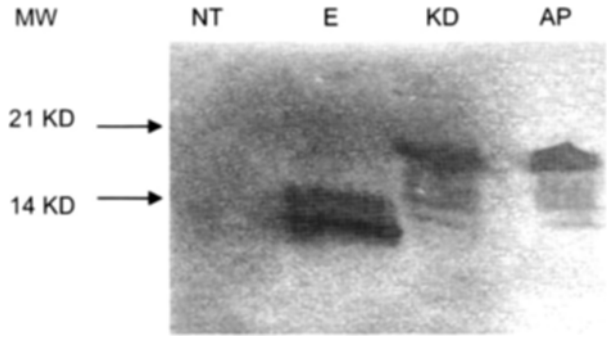


Figure 3. Western blot analysis of transgenic plants. Crude protein extracts (15 μL) from immature siliques were subjected to western blot and probed with rabbit anti-hGM-CSF antibody and goat anti-rabbit IgG-HRP. NT, protein from non-transformed plants; E, 50 ng GM-CSF produced and purified from *E. coli*; KD, silique extracts from KD6 lines; AP, silique extracts from AP8 lines.

transgenic plants, seeds from nine transformants at 10 DAP were analyzed by real-time quantitative SYBR-Green I® RT-PCR, using gene-specific primers. We chose the *A. thaliana actin2* gene as a control to normalize differences in RNA template concentrations. Transformants with increased accumulations of recombinant hGM-CSF also displayed relatively high levels of hGM-CSF mRNA (Fig. 4), with amounts being greater in AP8 than in other transformants. Moreover, the hGM-CSF expression level (hGM-CSF/*actin2*) reached a relative value of 2.4 compared with only 0.8 in KD15 transformants. The greater the accumulation of recombinant hGM-CSF, the higher its mRNA expression level, thereby implying that the insertion of nucleotides from domain IV of human 7SL RNA into the 3′-UTR of our hGM-CSF gene could improve mRNA stability and, therefore, enhance protein accumulation.

Biological Activity of the Seed-Expressed Recombinant GM-CSF

We used an MTT assay to investigate the biological activity of hGM-CSF derived from transformed *A. thaliana* seeds. Interestingly, TF-1 cells grew only in the medium supplemented

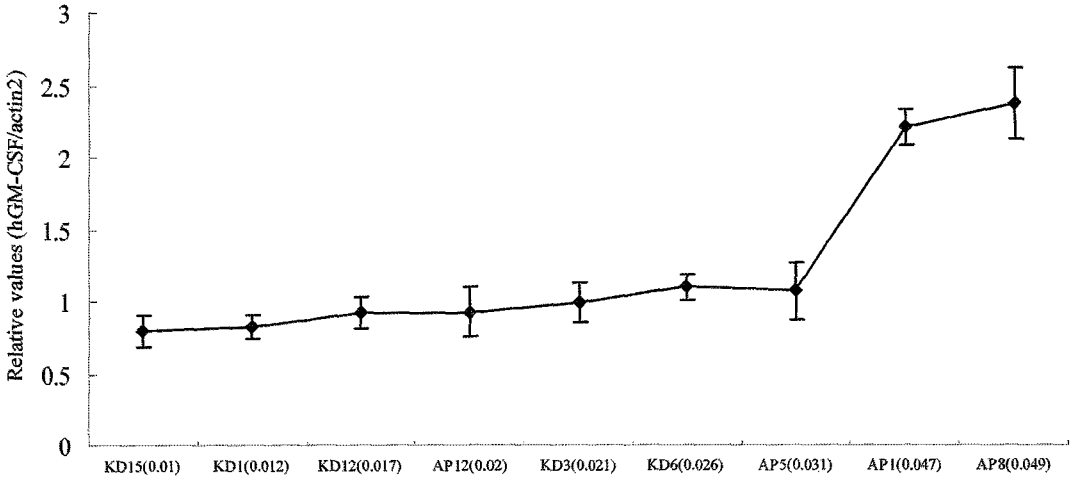


Figure 4. Quantitative real-time PCR analysis. Total RNA was isolated from seeds (10 DAP), reverse-transcribed, and subjected to real-time PCR. Relative amounts were calculated and normalized with respect to *actin2* transcript levels (=100%). Data represent mean values obtained from 3 independent amplification reactions; error bars indicate SE of means. Note that data in parentheses show ELISA values×100.

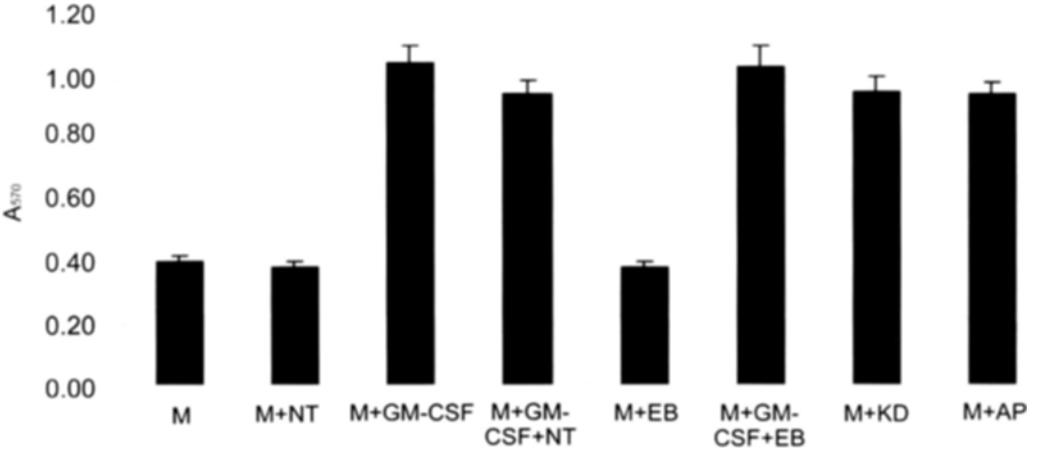


Figure 5. Biological activity of seed-expressed human GM-CSF. Bioassays were performed on TF-1 cells (2×10^4 cells per well), and cell proliferation was analyzed by MTT assay. Cells were incubated in presence or absence of seed protein extracts from KD transformants (KD), AP transformed lines (AP), non-transformed plants (NT), extraction buffer (EB), and with *E. coli*-derived human GM-CSF. All experiments were completed 3 times in triplicate. Standard errors are indicated on each bar.

with GM-CSF. Over 72 h, absorbance values were maintained at approximately 0.4 from all groups, including the RPMI-1640 medium (M) alone, M added to silique extracts from non-transformed plants (NT), and the protein extraction buffers (EB) (Fig. 5). This may have been due to the presence of live cells or background. Both the medium supplemented with extracts from KD transformants (M+KD plus 2 ng mL^{-1} recombinant hGM-CSF) and the medium containing protein extraction buffer from AP-transformed plants (M+AP, plus 2 ng mL^{-1} recombinant hGM-CSF) supported the proliferation of TF-1 cells, similar to yields from the medium supplemented with 2 ng mL^{-1} *E. coli*-produced hGM-CSF and non-transformed extracts (M+GM-CSF+NT). Cell activity from the medium supplemented with 2 ng mL^{-1} commercial hGM-CSF and extraction buffer (M+GM-CSF+EB) was similar to that observed from the RPMI-1640 medium containing 2 ng mL^{-1} *E. coli*-produced human GM-CSF (M+GM-CSF). No significant differences were found among the M+GM-CSF, M+GM-CSF+EB, M+GM-CSF+NT, M+KD, and M+AP treatments. While the MTT assay measured relative TF-1 cell proliferation, its data suggested that hGM-CSF from the seeds of transformed *A. thaliana* was biologically active. This activity was independent of the glycosylation status.

DISCUSSION

We introduced nucleic acid sequences from domain IV of human 7SL RNA into the 3′-UTR of hGM-CSF to allow its secondary loop structure to sense the SRP54 protein, its intrinsic ligand. In doing so, we ensured that the mRNA would be protected by proteins from rapid degradation, thereby prolonging the mRNA half-life and improving the accumulation of recombinant protein. Although the *in vivo* elimination half-life of recombinant human GM-CSF depends on many factors, it is mainly controlled by the 3′-UTR adenosine-uridine-rich element that can bind the RNA binding protein AUF1 to regulate transcript stability (Rajagopalan and Malter, 1994; Jarzembowski et al., 1999). Here, we determined that the 3′-UTR of recombinant human GM-CSF contained domain IV sequences from human 7SL RNA plus the NOS poly(A) signal to improve its mRNA stability. However, the actual mRNA half-life remains unknown in transformed plants. Our AP transformants not only displayed higher mRNA levels of hGM-CSF than did KD transgenic plants, but also showed greater accumulation of recombinant protein than the KD lines. This suggests that the introduction of aptamers into the gene may increase mRNA stability and, therefore, heighten protein accumulation.

The chromatin structure at the site of integration -- euchromatin and heterochromatin – is interspersed throughout the chromosome, and can positively or negatively affect transgene expression (Wakimoto, 1998). Differences in these integration sites may have lead to variations in hGM-CSF gene expression in our AP and KD transformants. Some transformed plants exhibited very low levels of hGM-CSF expression even though the Southern blots confirmed its integration into the plant genome. This may have been the result of post-transcriptional gene silencing, which frequently occurs in transgenic plants but for which the mechanism is not clearly understood (Kooter et al., 1999). Our AP transformants had generally higher levels of expression than did our KD transgenic plants, which indicates that the insertion of aptamers into the 3′-UTR of the gene might function in mRNA stability and protein accumulation, rather than because of any side effect of the integrated position.

In genetic engineering of plants, sterility can be caused by the high accumulation of human cytokines. For example, constitutive expression of the human *epo* gene affects vegetative growth and male sterility of transgenic tobacco (Cheon et al., 2004). In our study, few AP-transformed plants bore seeds, suggesting that high constitutive hGM-CSF expression during early seed development may have harmed fertility. However, recombinant hGM-CSF protein still accumulated in the harvested storage organs from some AP transgenic lines.

Although the predicted molecular mass of natural human GM-CSF is 16 kD, we observed a band at 21 kD from seed-produced hGM-CSF. This may have been due to glycosylation of the protein, although its carbohydrate content has not yet been analyzed (Sardana et al., 2002). Human GM-CSF contains two potential *N*-linked glycosylation sites, at Asn27 and Asn37, as well as two Ser/Thr/Pro rich regions that also may be substrates for *O*-linked glycosylation. In fact, the 21-kD GM-CSF protein may be glycosylated at both sites (Wang et al., 2005). Our results, as well as those of Raemaekers et al. (1999), confirmed that the phytohemagglutinin signal sequences were correctly processed. This might then imply that the PHA-E signal peptide was correctly removed in the mature recombinant hGM-CSF.

While glycosylation does not seem to affect the biological activity of GM-CSF either *in vivo* or *in vitro*, previous research has demonstrated that hGM-CSF purified from yeast, bacteria, or mammalian cells shows variations in its pharmacokinetics and immunogenicity (Burgess et al., 1987; Dorr, 1993). Further research is needed on seed-derived hGM-CSF and its chimeric fusions, as well as analysis of its carbohydrate composition and structure.

Seeds are natural storage sites for edible proteins, and are also harvestable as high-quality products. Pharmaceutical proteins may maintain long-term natural status and biological activity in the seeds, reducing the cost of refrigeration and preservation. For these reasons, seed may provide an effective system for manufacturing therapeutic protein drugs (de Jaeger et al., 2002).

The technique reported here can also be applied in the production of other valuable proteins in plants. Using a combination of multiple mechanisms, e.g., a strong promoter, and increasing the gene-copy numbers may allow for effective exploitation of plants as drug sources.

ACKNOWLEDGMENTS

We thank Professor Qiao for providing the plasmid pET28a(+)-CSF. We are grateful to Jiao Tong-Nottingham University Plant Biotechnology Research Center for the use of those facilities for most of our experiments. This work was supported by a grant from the "Soybean Transgenic Engineering" project in the National High Technology Research

and Development Program of China ("863" Program, No. 2001AA212141).

Received November 14, 2006; accepted January 27, 2007.

LITERAURE CITED

Althoff S, Selinger D, Wise JA (1994) Molecular evolution of SRP cycle components: Functional implications. Nucl Acids Res 22: 1933-1947

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (2001) Current Protocols in Molecular Biology. Ed 4, John Wiley and Sons, New York

Batey RT, Rambo RP, Lucast L, Rha B, Doudna JA (2000) Crystal structure of the ribonucleoprotein core of the signal recognition particle. Science 287: 1232-1239

Bernstein HD, Poritz MA, Strub K, Hoben PJ, Brenner S, Walter P (1989) Model for signal sequence recognition from amino-acid sequence of 54K subunit of signal recognition particle. Nature 340: 482-486

Bock LC, Griffin LC, Latham JA, Vermaas EH, Toole JJ (1992) Selection of single-stranded DNA molecules that bind and inhibit human thrombin. Nature 355: 564-566

Bradford MM (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

Burgess AW, Begley CG, Johnson GR, Lopez AF, Williamson DJ, Mermod JJ (1987) Purification and properties of bacterially synthesized human granulocyte-macrophage colony stimulating factor. Blood 69: 43-51

Cantrell MA, Anderson D, Cerretti DP, Price V, McKereghan K, Tushinski RJ, Mochizuki DY, Larsen A, Grabstein K, Gillis S (1985) Cloning, sequence, and expression of a human granulocyte/macrophage colony-stimulating factor. Proc Natl Acad Sci USA 82: 6250-6254

Chen ZL, Schuler MA, Beachy RN (1986) Functional analysis of regulatory elements in a plant embryo-specific gene. Proc Natl Acad Sci USA 83: 8560-8564

Cheon BY, Kim HJ, Oh KH, Bahn SC, Ahn JH, Choi JW, Ok SH, Bae JM, Shin JS (2004) Overexpression of human erythropoietin (EPO) affects plant morphologies: Retarded vegetative growth in tobacco and male sterility in tobacco and *Arabidopsis*. Transgenic Res 13: 541-549

Chung II A, Trinh VB, Yokobayashi Y (2006) Artificial control of gene expression in mammalian cells by modulating RNA interference through aptamer–small molecule interaction. RNA 12: 710-716

Clough SJ, Bent AF (1998) Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J 16: 735-743

de Jaeger G, Scheffer S, Jacobs A, Zambre M, Zobell O, Goossens A, Depicker A, Angenon G (2002) Boosting heterologous protein production in transgenic dicotyledonous seeds using *Phaseolus vulgaris* regulatory sequences. Nature Biotech 20: 1265-1268

Diederichs K, Boone T, Karplus PA (1991) Novel fold and putative receptor binding site of granulocyte-macrophage colony-stimulating factor. Science 254: 1779-1782

Dorr RT (1993) Clinical properties of yeast-derived versus *Escherichia coli*-derived granulocyte-macrophage colony stimulating factor. Clin Therapeut 15: 19-29

Gossen M, Bujard H (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc Natl Acad Sci USA 89: 5547-5551

Grate D, Wilson C (2001) Inducible regulation of the *S. cerevisiae* cell cycle mediated by an RNA aptamer-ligand complex. Bioorg Med Chem 9: 2565-2570

Guerche P, Tire C, de Sa FG, de Clercq A, van Montagu M, Krebbers E (1990) Differential expression of the *Arabidopsis* 2S albumin genes and the effect of increasing gene family size. Plant Cell 2: 469-478

Hanson S, Bauer G, Fink B, Suess B (2005) Molecular analysis of a synthetic tetracycline-binding riboswitch. RNA 11: 503-511

Hercus TR, Cambareri B, Dottore M, Woodcock J, Bagley CJ, Vadas MA, Shannon MF, Lopez AF (1994) Identification of residues in the first and fourth helices of human granulocyte-macrophage colony-stimulating factor involved in biologic activity and in binding to the alpha- and beta-chains of its receptor. Blood 83: 3500-3508

Jarzembowski JA, Rajagopalan LE, Shin HC, Malter JS (1999) The 5'-untranslated region of GM-CSF mRNA suppresses translational repression mediated by the 3' adenosine-uridine-rich element and the poly(A) tail. Nucl Acids Res 27: 3660-3666

Joel M, Neely GL, Trauger JW, Baird EE, Dervan PB (1997) Regulation of gene expression by small molecules. Nature 387: 202-205

Kasugai S, Hasegawa N, Ogura HA (1990) A simple *in vitro* cytotoxicity test using the MTT (3-(4,5)-dimethylthiazol-2-yl-2,5-diphenyl tetrazolium bromide) colorimetric assay: Analysis of eugenol toxicity on dental pulp cells (RPC-C2A). Jpn J Pharmacol 52: 95-100

Kaushansky K, O'harat PJ, Berknert K, Segal GM, Hagent FS, Adamson JW (1986) Genomic cloning, characterization, and multilineage growth-promoting activity of human granulocyte macrophage colony-stimulating factor. Proc Natl Acad Sci USA 83: 3101-3105

Keenan RJ, Freymann DM, Stroud RM, Walter P (2001) The signal recognition particle. Annu Rev Biochem 70: 755-775

Kim YH, Hur CG, Shin YH, Bae JM, Song YS, Huh GH (2006) Identification and characterization of highly expressed genes in suspension-cultured cells of sweet potato. J Plant Biol 49: 364-370

Kooter JM, Matzke MA, Meyer P (1999) Listening to the silent genes: Transgene silencing, gene regulation and pathogen control. Trends Plant Sci 4: 340-347

Kwon TH, Kim YS, Lee JH, Yang MS (2003) Production and secretion of biologically active human granulocyte-macrophage colony stimulating factor in transgenic tomato suspension cultures. Biotechnol Lett 25: 1571-1574

Lee F, Yokota T, Otsuka T, Gemmell L, Larson N, Luh J, Arai K, Rennick D (1985) Isolation of cDNA for a human granulocyte-macrophage colony-stimulating factor by functional expression in mammalian cells. Proc Natl Acad Sci USA 82: 4360-4364

Paulus M, Haslbeck M, Watzele M (2004) RNA stem–loop enhanced expression of previously non-expressible genes. Nucl Acids Res 32: e78

Raemaekers RJ, de Muro L, Gatehouse JA, Fordham-Skelton AP (1999) Functional phytohemagglutinin (PHA) and *Galanthus nivalis* agglutinin (GNA) expressed in *Pichia pastoris* correct N-terminal processing and secretion of heterologous proteins expressed using the PHA-E signal peptide. Eur J Biochem 265: 394-403

Rajagopalan LE, Malter JS (1994) Modulation of granulocyte-macrophage colony-stimulating factor mRNA stability *in vitro* by the adenosine-uridine binding factor. J Biol Chem 269: 23882-23888

Roh KH, Shin KS, Lee YH, Seo SC, Park HG, Daniell H, Lee SB (2006) Accumulation of sweet protein monellin is regulated by the *psb*A 5'-UTR in tobacco chloroplasts. J Plant Biol 49: 34-43

Sardana RK, Alli Z, Dudani A, Tackaberry E, Panahi M, Narayanan M, Ganz P, Altosaar I (2002) Biological activity of human gran-

ulocyte-macrophage colony stimulating factor is maintained in a fusion with seed glutelin peptide. Transgenic Res 11: 521-531

SAS (2000) Statistical Analysis Systems, Vers 8.0, SAS Institute, Cary

Schmitz U, Behrens S, Freymann DM (1999a) Structure of the phylogenetically most conserved domain of SRP RNA. RNA 5: 1419-1429

Schmitz U, James TL, Lukavsky P, Walter P (1999b) Structure of the most conserved internal loop in SRP RNA. Nat Struct Biol 6: 634-638

Sen D (2002) Aptamer rivalry. Chem Biol 9: 851-852

Shin YJ, Hong SY, Kwon TH, Jang YS, Yang MS (2003) High level of expression of recombinant human granulocyte-macrophage colony stimulating factor in transgenic rice cell suspension culture. Biotechnol Bioengr 82: 778-783

Sudarsan N, Barrick JE, Breaker RR (2003) Metabolite-binding RNA domains are present in the genes of eukaryotes. RNA 9: 644-647

Suess B, Hanson S, Berens C, Fink B, Schroeder R, Hillen W (2003) Conditional gene expression by controlling translation with tetracycline-binding aptamers. Nucl Acids Res 31: 1853-1858

Wakimoto B (1998) Beyond the nucleosome: Epigenetic aspects of position effect variegation in *Drosophila*. Cell 93: 321-324

Wang ML, Goldstein C, Su W, Moore PH, Albert HH (2005) Production of biologically active GM-CSF in sugarcane: A secure biofactory. Transgenic Res 14: 167-178

Werstuck G, Green MR (1998) Controlling gene expression in living cells through small molecule-RNA interactions. Science 282: 296-298