

Utah State University

DigitalCommons@USU

All Graduate Theses and Dissertations

Graduate Studies

5-1998

Protein Production in the Milk of Genetically Engineered Animals

Katherine M. Bates

Utah State University

Follow this and additional works at: <https://digitalcommons.usu.edu/etd>



Part of the [Animal Sciences Commons](#)

Recommended Citation

Bates, Katherine M., "Protein Production in the Milk of Genetically Engineered Animals" (1998). *All Graduate Theses and Dissertations*. 3980.

<https://digitalcommons.usu.edu/etd/3980>

This Thesis is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.



PROTEIN PRODUCTION IN THE MILK OF
GENETICALLY ENGINEERED ANIMALS

by

Katherine M. Bates

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Animal Science
(Molecular Biology)

UTAH STATE UNIVERSITY
Logan, Utah

1998

ABSTRACT

Protein Production in the Milk of Genetically Engineered Animals

by

Katherine M. Bates, Master of Science

Utah State University, 1998

Major Professor: Dr. John D. Morrey

Department: Animal, Dairy and Veterinary Science

There are numerous proteins that have potential uses in commercial and scientific applications that are not utilized to their full potential. This is partly because it is not economically feasible to isolate some of these proteins from their natural sources or to produce them using bacterial fermentation methods. The purpose of this research was to target recombinant protein expression to the mammary glands of genetically engineered or transgenic animals. Foreign protein expression has been achieved in the mammary glands of rabbits, sheep, cows, and swine. By using a strong mammary gland promoter and signal peptide fused to the protein, it was hypothesized at the beginning of the study that the two proteins of this study would be secreted into the milk.

To test this approach for protein production, expression vectors for two different plant proteins were made. The proteins targeted for expression were thaumatin and brazzein, proteins that have sweetener or flavor altering properties. The regulatory

portion of the expression vector used exons and introns from the milk β -casein gene.

Four and a half kilobases of the 5' region of the bovine β -casein gene was isolated, which contained the promoter sequence and other regulatory sequences for gene expression in mammary tissue. A size of 2.2 kilobases of the 3' region of the β -casein gene contained further regulatory sequences as well as a polyadenylation signal. The gene sequence for the protein was modified by using codons commonly used for casein and was generated using synthetic oligonucleotides. Additionally, the signal peptide from the alpha S-1 casein gene was used to transport the protein into the mammary milk vesicle. The DNA expression vectors were subsequently injected into murine and caprine embryos for the production of transgenic animals. Transgenic mice and a goat were identified that contained the thaumatin transgene. Preliminary analysis of mouse milk by capillary gel electrophoresis indicated the expression of thaumatin protein. This protein expression system is intended to utilize large dairy animals as bioreactors for efficient, non-toxic protein production with a view to being applied to different proteins as the technology advances.

(118 pages)

In loving memory and appreciation of my dad, Dr. Micheal P. Bates, C.B.E.,
who has and will always be an inspiration to me.

ACKNOWLEDGMENTS

I would like to thank my advisor, John Morrey, for his support, patience, and guidance in this research project. The technical training and research experience that I have gained are invaluable. The time and advice given by the members of my committee, Drs. Kenneth White and Anne Anderson, are much appreciated.

The many graduate and undergraduate members in the laboratories of Drs. Morrey and White were an invaluable source of assistance and training for the many procedures that were involved in this research project and I am grateful for their friendship and support. Thanks in particular to Dr. Tom Bunch, Charoensri Thonsambulat, Dr. William Reed, Dr. Ken White, and Dr. John Morrey for the work done in murine embryo transfer.

I also would like to thank Yeelan Wang for all the work involved in the DNA sequencing and Dr. Marie Strickland for the capillary gel electrophoresis.

Thanks to my friends and family, in particular my Mum and Dad, who supported me through the past two years and humored me when I got excited about various milestones achieved in my project.

Katherine M. Bates

CONTENTS

	Page
ABSTRACT	i
DEDICATION	iv
ACKNOWLEDGMENTS.....	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
INTRODUCTION	1
LITERATURE REVIEW.....	3
Previous work in generating transgenic animals	3
Incorporation of the transgene and regulation of gene expression.....	3
Previous work done in targeting expression to the mammary tissue	4
Selection of promoter and regulatory elements.....	5
Selection of thaumatin protein for study	6
Characteristics of thaumatin.....	6
Sweetener properties and uses of thaumatin.....	6
Commerce.....	8
Selection of brazzein protein for study.....	10
EXPERIMENTAL DESIGN.....	11
Objective 1: Construction of expression vector	11
Design of synthetic genes.....	15
Construction of the synthetic genes.....	19
Objective 2: Generation of transgenic lines to test the expression vector.....	23
Objective 3: Detection of protein in the milk of the transgenic mice.....	24
EXPERIMENTAL METHODS	25

Objective 1: Construction of expression vector	25
Cloning β -casein regions into pGFP-C1	25
Generating separate segments of the synthetic genes	33
Combining thaumatin segments	37
Inserting flavor or sweetener genes into p5'3' β cas construct	40
Objective 2: Generation of transgenic lines to test the expression vector	42
Preparation of expression vector for microinjection	42
Murine embryo transfer	43
Screening for potential transgenic animals	43
Objective 3: Detection of protein in the milk of the transgenic mice	46
Generation of antibodies to thaumatin and brazzein	46
ELISA	47
Collection of mouse milk	49
Analysis of mouse milk proteins by capillary gel electrophoresis	49
MATERIALS	51
RESULTS	55
Objective 1: Construction of expression vector	55
Cloning β -casein regions into pGFP-C1	56
Generating separate segments of the synthetic genes	67
Combining thaumatin segments	75
Inserting flavor/sweetener genes into p5'3' β cas construct	77
Objective 2: Generation of transgenic lines to test the expression vector	81
Preparation of expression vector for microinjection	81
Screening for potential transgenic animals	85
Objective 3: Detection of protein in the milk of the transgenic mice	90
DISCUSSION	99
Objective 1: Construction of expression vector	99
β casein gene	99
Cloning of the expression vector	99
Synthesis of long oligonucleotides	100

Objective 2: Generation of transgenic lines to test the expression vector.....	100
Identification of transgenic animals.....	100
Objective 3: Detection of protein in the milk of the transgenic mice.....	101
Identification of thaumatin protein in mouse milk.....	101
Immunoassay to detect protein.....	102
CONCLUSION.....	103
LITERATURE CITED.....	105

LIST OF TABLES

Table	Page
1 Applications of thaumatin sweetner protein	7
2 Effect of 1ppm thaumatin on the feed or water intake and performance of 2-week-weaned piglets.....	7
3 Recombinant thaumatin expression attempts	9
4 Comparison of thaumatin and brazzein	10
5 Common codon frequency of leucine.....	16

LIST OF FIGURES

Figure	Page
1	Components of expression vector.....12
2	Features of the 5'β-casein portion of the expression vector13
3	Features of the 3'β-casein portion of the expression vector14
4	Comparison of α S1-casein with β-casein and the synthetic sequence used.....16
5	Sequence of the synthetic thaumatin construct.....17
6	Sequence of the synthetic brazzein construct.....18
7	Design of the genetic expression vector20
8	Construction of the thaumatin synthetic gene21
9	Construction of the brazzein synthetic gene.....21
10	Anticipated transgenic mRNA and protein processing in vivo.....22
11	Design of brazzein oligonucleotides and PCR primers33
12	Design of thaumatin oligonucleotide segments and PCR primers.....34
13	Relation of thaumatin segments with each other.....37
14	Location of probe used in detection of transgene.....45
15	Design of ELISA47
16	Marker size map55
17	5'β-casein DNA region obtained by PCR.....56
18	3'β-casein DNA region obtained by PCR.....57
19	Screening of 5'β-casein into pGFP-C1 (gel #763).....58
20	Screening of putative p5'βcas clones (gel #765)59

21	Sequence comparison between the cloned and published sequences of 5' portion of β -casein.....	60
22	Two possible orientations of cloning 3' β -casein DNA Sequences into pUC-19.....	61
23	Screening of putative 3' β -casein-pUC19 (gel #752).....	62
24	Screening putative 3' β -casein-pUC clones (gel #754).....	63
25	Sequence analysis of 3' portion of β -casein cloned into pUC.....	64
26	Autoradiograph of colony hybridization to detect p5'3' β cas.....	65
27	Screening of p3'5' β -cas #1 (gel #945).....	66
28	Screening of putative p3'5' β -cas #2 (gel #946).....	67
29	Synthetic gene amplification (gel #843).....	68
30	Synthetic gene amplification using optimization buffers #1 (gel #846).....	69
31	Synthetic gene amplification using optimization buffers #2 (gel #847).....	70
32	Screening of thaumatin-pUC (gel #841).....	71
33	Brazzein digested out of pUC (gel #805).....	72
34	Sequence analysis of a thaumatin C clone that contained multiple errors.....	73
35	DNA sequence of synthetic brazzein (rxn: 3786#4, upper row) as compared to the predicted sequence (lower row).....	73
36	DNA sequence of correct thaumatin A (rxn: 3787 #12, upper row) as compared to the predicted sequence (lower row).....	74
37	DNA sequence of correct thaumatin B (rxn: 4059 #8, upper row) as compared to the predicted sequence (lower row).....	74
38	DNA sequence of correct thaumatin C (rxn: 4060 #19, upper row) as compared to the predicted sequence (lower row).....	75

39	DNA sequencing of the correct thaumatin D (rxn: 3989 #26, upper row) as compared to the predicted sequence (lower row).....	76
40	Possible ligation events when joining thaumatin A & B and C & D.....	78
41	10% PAGE with PCR product of thaumatin AB and CD (gel #1012).....	79
42	Screening of thaumatin-pUC (gel #1019).....	79
43	Screening of thaumatin-p5'3' β -cas (gel #1049).....	80
44	Thaumatin expression vector cut for microinjection (gel #1070).....	81
45	Diagnostic agarose gel for brazzein expression vector using restriction enzymes AatII and AgeI (gel #974)	83
46	Pertinent predicted restriction enzyme sites of brazzein expression vector.....	83
47	Diagnostic agarose gel for brazzein expression vector using restriction enzymes BamHI and KpnI (gel #977)	84
48	Brazzein expression vector cut for microinjection (gel #984).....	85
49	Agarose gel prepared for murine genomic Southern blot (gel #1164).....	85
50	Autoradiograph of murine genomic Southern blot (gel #1164)	86
51	Agarose gel prepared for F0 and F1 murine genomic Southern blot (gel #1126)	87
52	Autoradiograph of F0 and F1 murine genomic Southern blot (gel #1226).....	88
53	Agarose gel prepared for genomic caprine Southern blot (gel #1218).....	89
54	Autoradiograph of genomic caprine Southern blot after 3-day exposure (gel #1218).....	89
55	Autoradiograph of genomic caprine Southern blot after 6-week exposure (gel #1218).....	90
56	ELISA with brazzein multiple antigenic peptide diluted in PBS.....	92

57	ELISA with brazzein multiple antigenic peptide diluted in bovine milk	92
58	ELISA with thaumatin plant protein diluted in PBS	93
59	ELISA with thaumatin plant protein diluted in bovine milk	93
60	Capillary gel electrophoresis of normal mouse milk spiked with crude plant thaumatin	94
61	Capillary gel electrophoresis of mouse milk control	96
62	Capillary electrophoresis of crude thaumatin protein	96
63	Capillary gel electrophoresis of milk from mouse #77	97
64	Capillary gel electrophoresis of nontransgenic mouse #8 litter mate milk	97
65	Capillary gel electrophoresis of nontransgenic mouse #19 litter mate milk	98
66	Capillary gel electrophoresis of milk from a nontransgenic mouse	98

INTRODUCTION

The long-term goal of this research was to target foreign protein expression to the milk of genetically engineered livestock animals in order to obtain large-scale protein production of potentially valuable proteins.

The hypothesis is that high levels of functional protein can be produced if the sequence for the foreign protein is placed under the regulatory control of an inducible mammary gland-specific promoter that has been shown to induce gene expression during lactation. By fusing the protein with a secretory signal peptide, the proteolytic enzymes of mammary gland cells can cleave exactly at the carboxy-terminal end of the signal peptide. The functional, native protein is secreted into the milk-producing vesicles of the mammary epithelial cells. The mammary glands in transgenic animals have been shown to be effective for protein production using various combinations of mammary-specific promoters and proteins (4, 32-34). For this study, an expression vector was made which was composed of a mammary gland-specific promoter, regulatory region, and signal peptide sequence such that the gene of interest could be inserted into the expression vector. The vectors were then micro-injected into pronuclear-staged embryos that were implanted into surrogate dams and carried to term. Those offspring that had incorporated the transgene should express the foreign protein in the milk. The protein can then be harvested from the milk.

Two proteins that have flavor-enhancing/sweetener properties were used to test the expression vector. The first protein was thaumatin, which is a sweet-tasting protein from the fruit of a West African perennial herb *Thaumatococcus Daniellii* Benth. Over the years thaumatin has been shown to have multiple applications that are economically important. However, it does not naturally occur in high enough quantities to meet the current market. Various recombinant methods have been, and continue to be, used to attempt to produce recombinant thaumatin to meet the market demands. These methods to date have proven unsuccessful, because they either do not yield functional protein or they do not produce a

large enough thaumatin yield to make the production economically feasible. The second protein was brazzein, a sweetener protein found in plants. The gene was isolated from a wild African plant *Pentadiplandra brazzeana* Baillon (36). Although there is not currently a market for brazzein, it has properties that suggest it is the most superior protein sweetener known (24). It is 2000 times sweeter than sucrose, more water-soluble than any other protein sweetener, is thermostable, and possesses a sucrose-like taste.

The goal of this project was to develop the gene expression vectors containing the genes of interest, generate transgenic mice, and develop assays that could be used to detect the proteins.

LITERATURE REVIEW

Previous work in generating transgenic animals

The first transgenic animal was made in 1980. By injecting pBR322 plasmid containing portions of the herpes simplex and SV40 viruses into the pronucleus of an early mouse embryo, mice were generated that contained the injected plasmid in their genome (12).

Since the first transgenic mice, the use of transgenic animals has been incorporated in a variety of fields, such as for models of human diseases. By using these transgenic mouse models, the disease traits and mechanisms can be better studied in a living system. Another application is using transgenic animals as bioreactors to produce recombinant products by achieving expression in the mammary gland or the blood (16), which is the effort of this research.

Incorporation of the transgene and regulation of gene expression

The DNA sequence to be targeted to the host genome is purified and microinjected into the pronuclei of single-cell fertilized oocytes. The transgene integrates into the genome of the embryo by recombination. If the foreign DNA, or transgene, is integrated into germline cells of the animal, the transgene will be inherited by the transgenic progeny through Mendelian genetics (16).

The genetic expression vector is designed such that the gene of interest will be easily recognized and regulated by the host cellular mechanisms. The DNA sequence consists of regulatory regions that are required for the tissue-specific recognition of the foreign DNA. Mammary gland-specific transcription factors bind to the regulatory sequences of the promoter, which synthesizes the corresponding RNA. Inducible transcription factors are synthesized at specific times during cell cycling and development, in specific tissues, and play a role in the control of transcription (20). These promoter-

associated factors enhance or suppress the efficiency of the RNA polymerase/transcription complex to initiate and transcribe the DNA sequence into messenger RNA (mRNA). The enzyme responsible, RNA polymerase, continues along the sequence in a 5' to 3' direction until it reaches and recognizes the polyadenylation signal - AATAAA. This results in the subsequent addition of a polyadenylic (polyA) tail to the 3' end and cessation of transcription. The mRNA includes both the regulatory intronic and coding exonic regions. To remove the introns, the mRNA is processed or spliced. The processed mRNA, which now has a 5' cap and a 3' polyA tail, is then transported out of the nucleus to the cytoplasm, where it is translated into the corresponding amino acid residues.

The mammary gland proteins found in milk need to be translocated into milk secretory vesicles. This is facilitated by a secretory signal peptide that is flush with the first residue of the protein to be secreted (28). The signal peptide is recognized by factors associated with the membrane of the secretory vesicle and these membrane-associated factors cause the peptide to be transported through the membrane of the secretory vesicle. As it is passing through the membrane, the signal peptide is cleaved away from the remainder of the protein at the signal peptides' carboxy terminal end. Consequently only the remaining protein sequence is secreted in the milk (20).

Previous work in targeting expression to the mammary tissue

Transgenic protein expression in the milk has been achieved by other investigators. Some examples of such proteins are human tissue plasminogen activator (6), human growth factor (33), and whey acidic protein (18). Such research has shown that heterologous mammary-specific promoters can be used to express transgenic proteins in the milk of transgenic animals as illustrated below.

Whey acidic protein (WAP) is a milk protein that is unique to the rodent family. Transgenic swine and sheep were generated using the gene coding for the murine WAP. This resulted in the WAP being expressed at a high level of 3% (0.9 g per liter) of the total

milk proteins (18). Expression of this rodent mammary-specific gene in swine and sheep demonstrates that milk or mammary gland-specific promoters could still be functionally recognized in a heterologous species. It also demonstrates that a murine protein could be expressed at high levels in the milk livestock animals.

The utility of the WAP promoter provided the opportunity to express human proteins in the milk of livestock animals. The cDNA of human protein C (hPC), utilizing the first exon of WAP gene, was expressed at a high concentration of 1 g/L in swine. In addition, the biological activity was comparable to that of hPC obtained from human plasma (32).

Selection of promoter and regulatory elements

The protein concentration of milk is 32 g/L. The caseins, of which there are four types, alpha-s1, alpha-s2, beta and kappa, comprise 76-86% of the milk proteins. Each of the casein genes has been utilized in a transgenic system. The milk protein genes contain short conserved sequences in their 5' flanking regions which contain the binding sites for the tissue-specific transcription factors (16). The casein genes are expressed in the mammary glands of the animals during the late gestation and lactation periods under the influence of hormones.

Of the four caseins, beta-casein is the second most prevalent at 9 g/L. A comparison between the promoters of the beta-casein of different species show that the DNA sequences are conserved. This is indicative of the biological importance of beta-casein as a component of milk. These casein promoters are useful for milk expression of proteins. For example, a 3-kb portion of the beta-casein promoter region was successfully used to achieve expression of a portion of the caprine kappa-casein gene in mice (13). In vitro expression of the bovine beta-casein promoter has been induced in the HC11 mammary epithelial cell line, in the presence of insulin, glucocorticoid, and prolactin (2). The beta-casein promoter region was linked to the CAT reporter gene and expression levels

were measured. As a negative control, the SV40 promoter was fused with the CAT reporter gene and no expression was seen (2). These hormonal conditions resemble those seen during lactation.

Selection of thaumatin protein for study

Thaumatococcus daniellii Benth is a perennial herb which grows predominantly in the West African countries and other humid regions on the African continent. Its fruit is known to be very sweet, and as a result, is used both as a fruit delicacy and a food additive by the indigenous people. *T. daniellii* grows naturally and is also cultivated and farmed due to its economic value. As well as the fruit being used for food, the remaining portions of the plant have a variety of other uses that make it practical for production in these countries (1).

Characteristics of thaumatin

There are five known thaumatin proteins. The predominant proteins are thaumatin I and II. They are both 207 amino acids long and they only differ by 5 amino acid residue changes (10). Thaumatin II is considered to have the strongest role in flavor altering properties and therefore has been the protein of choice for recombinant DNA production. The protein has eight disulfide bonds, which would presumably account for the heat stability observed. After boiling the thaumatin protein for 1 hour, once it has cooled, it regains its sweetness (21). In the plant, thaumatin is secreted as a pre-pro-thaumatin with 22 amino acids on the N-terminus and 6 amino acids on the C-terminus. These sequences are cleaved as the protein is compartmentalized.

Sweetener properties and uses of thaumatin

Thaumatin is 3000 times sweeter than sucrose and has also been shown to enhance the sweet taste of foods and other non-nutritive sweeteners such as saccharin (11). It is

TABLE 1. Applications of thaumatin sweetener protein (11)

Amount (ppm)	Effect	Product
0.2 - 0.5	enhances flavor and masks bitterness	pet food, animal feed
0.5 - 1.0	modifies flavor and enhances aroma	coffee, artificial flavor
30 - 200	prolongs flavor and intensifies sweetness	gum and candy
50 - 150	masks bitterness and contributes sweetness	oral pharmaceuticals

best to use thaumatin for its ability to enhance taste and flavor due to the delay in perception of sweetness and an after-taste of licorice (11). Table 1 summarizes the different applications of the thaumatin protein.

Thaumatococcus has been successfully used in the feed of weaned piglets to cause them to eat more feed so that they would gain weight faster. Adding thaumatin to the water of weaned pigs at 1 part per million (ppm, $\mu\text{g/g}$) caused the pigs to drink more water and eat more feed, which resulted in more rapid weight gain (table 2). It has also been tried in other livestock animal feed applications and has been shown to increase the palatability in both cat and dog foods. For example, the eating preference of cats for thaumatin-treated feed was 82% as compared to 18% for untreated food (14).

TABLE 2. Effect of 1 ppm thaumatin on the feed or water intake and performance of 2-week-weaned piglets (14)

Treatment	Control	Thaumatococcus	Improvement
Initial average weight (kg)	3.45	3.68	-
Final average weight (kg)	10.70	12.64	+ 1.69 kg
Average daily gain (g)	207	256	+ 12.4 %
Average feed intake (kg)	9.36	11.5	+ 2.14 kg
Feed conversion ratio	1.29	1.29	-
Water intake (L) 6 days	2.9	4.06	+ 40 %
Trial duration (days)	35	35	

Commerce

Proteins are advantageous as sweeteners because they are generally not toxic and have low caloric content. However, most protein sweeteners are found in exotic plants, and therefore the production can be limited. If they cannot be obtained in sufficient quantities from the native source, they could either be chemically synthesized or made by recombinant DNA technology. Presently there is not widespread use of thaumatin because there is not a large enough natural supply of the protein and due to political and economic factors the production costs are high. From a political standpoint, the countries that grow *T. daniellii* have an unstable political history. This results in some reluctance on the part of clients to purchase thaumatin due to a perceived unstable supply (5). Attempts have been made to grow *T. daniellii* in greenhouses in the United Kingdom but no fruit was produced (21). Thaumatin has an existing market for use in animal feeds, due to its taste-enhancing properties. However, the high cost of extracting the low level of protein from the fruit and exporting it limit its full commercial potential.

In an attempt to solve the supply problems associated with obtaining the thaumatin protein, extensive research has been undertaken to produce the protein using recombinant DNA technology. An initial attempt to produce recombinant thaumatin (r-thaumatin) in *Eschericia coli* (*E. coli*) resulted in low-level thaumatin that could only be detected using sensitive radiolabeled methods (10). In addition to only obtaining low levels of protein in *E. coli*, the protein was also not processed properly in the cell and therefore was not functional to exhibit sweet taste. Presumably one of the problems encountered with this effort was that the native signal sequence, that allowed secretion of the functional protein, was not recognized in bacteria. Therefore, the thaumatin protein was probably not processed or secreted properly. *Bacillus subtilis* and *Streptomyces lividans* naturally have high levels of enzyme secretion (13, 17). Better yields were obtained when using *B. subtilis* or *S. lividans* than were seen with *E. coli*, but the levels were not high enough for commercialization.

The most promising results were seen using the yeast *Saccharomyces cerevisiae* (21, 38). For this effort, a synthetic gene sequence was used that contained codons that were commonly used in yeast systems. The secreted thaumatin was not sweet unless the extracted protein was chemically denatured and then renatured into the functionally sweet form (19). Investigators eventually produced sweet thaumatin by fusing the invertase signal peptide sequence to the thaumatin protein sequence. The resulting protein possessed sweetness presumably because it was processed into the tertiary structure properly and was secreted into the growth medium. Under direction of the invertase signal sequence, a yield of 300 mg/L, from protein secreted into the growth medium, was obtained. Although this was a high yield, it did not reach the break-even point for commercialization (35). Table 3 summarizes some of the attempts to produce thaumatin by recombinant technology.

These microbial production methods have not yet proven successful for large-scale commercial production of the functional thaumatin protein, because the market for thaumatin is dependent on a high yield and low cost.

TABLE 3. Recombinant thaumatin expression attempts (11)

Host	Gene Description	Activity	Yield
<i>E. coli</i>	pre- and pro-gene	not functional	low (10)
<i>B. subtilis</i>	pre-gene	not functional	low (15)
<i>S. lividans</i>	pre-gene	not functional	low (11)
<i>K. lactis</i>	pre- and pro-gene	not functional	low (10)
<i>A. oryzae</i>	pre- and pro-gene	not functional	low (11)
<i>S. tuberosum</i>	pre- and pro-gene	functionally sweet	low (21)
<i>S. cerevisiae</i>	mature gene	not functional, but was refolded <i>ex vivo</i> to a sweet form	high (19)
	signal peptide linked with mature gene	functionally sweet	high (35)

Selection of brazzein protein for study

Brazzein, like thaumatin, is also a sweetener of plant origin. It was discovered in a wild African plant *Pentadiplandra brazzeana* Baillon (24). Brazzein is the most superior sweetener protein known. It is 2000-fold sweeter than sucrose and is more water soluble than any other protein sweetener. It is thermostable because it contains eight cysteines. The electrophysiological taste profile of brazzein is similar to sucrose (24). The sweetness of brazzein is immediate and short lasting unlike thaumatin. Functionally sweet brazzein was obtained using microbial host systems albeit at low levels (personal communication, Goran Hellekant, University of Wisconsin). This functional production in bacteria increase the likelihood that brazzein secreted in the milk will retain the desirable sweetness properties of the native brazzein.

TABLE 4. Comparison of thaumatin and brazzein (23,27)

Characteristic	Thaumatococcus II	Brazzein
Origin	<i>T. daniellii</i>	<i>P. brazzeana</i>
Sweetness ^a	3,000	2,000
Protein size	207 amino acid residues	54 amino acid residues
Molecular mass	22,293 Da	6,473 Da
Flavor altering property	flavor and aroma enhancer	sucrose-like sweet taste
Water soluble	yes	yes
Heat stability	regains sweetness after boiling for 1 hr when cooled	sweetness retained after 4 hr at 80°C

^a Compared to 2% aqueous sucrose

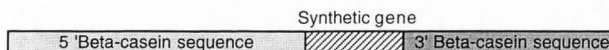
EXPERIMENTAL DESIGN

This Experimental Design section outlines the strategy and justification for the three objectives stated below without elucidating the details of the methods. An Experimental Methods section, following this Experimental Design section, describes the details of the methods.

1. Construct an expression vector containing genes of interest designed to target protein expression to the mammary gland and secretion of the proteins into the milk.
2. Generate a line of transgenic founder animals.
3. Develop an assay to detect and quantify the flavor/sweetener proteins.

Objective 1: Construction of expression vector

The expression vector seen in figure 1 was designed to target expression of the flavor/sweetener proteins, thaumatin and brazzein, to the mammary gland. One major component of the vector was a β -casein promoter anticipated to cause tissue-specific expression of the gene in mammary glands. Another component of the vector was a synthetic oligonucleotide sequence for the genes of the flavor and sweetener proteins targeted for expression. The synthetic oligonucleotides also included the code for the signal peptide required to transport the protein into the vesicles of the mammary glands so that it would be secreted in the milk. The amino acid sequence for the signal peptide from alpha S1-casein was used (accession #M33123). Alpha S1-casein is the predominant milk protein at 10 g/L (31) and has previously been used successfully to cause secretion of proteins into the milk (29). The third component was the 3' portion of the β -casein gene that contained a polyadenylation signal (poly A) from the β -casein gene. In addition to the essential promoter and poly A regions, the 3' and 5' segments of the β -casein gene included several kilobases of intronic and exonic regions. It was hypothesized that greater



□ 5'β-casein regulatory sequence: Contained the promoter and enhancer regions for gene expression in the mammary gland tissue, this included β-casein intron 1, exon 1, and a portion of exon 2, a transcription start codon and the sequence for the first three amino acid residues of the secretory signal peptide.

▨ Synthetic genes were designed for both thaumatin and brazzein, they consisted of the following:

1. Signal peptide sequence from alpha S1-casein(17)
2. Synthetic sequence for flavor proteins
3. Two stop codons at 3' end

■ 3'β-casein regulatory sequence: Contained a portion of β-casein exon 7 as well as intron 7, exon 8, intron 8, exon 9 and the poly-adenylation signal.

FIG. 1. Components of expression vector.

efficiency of expression would be achieved when a high proportion of the native coding sequences were included in the vector (6).

The sequence for the bovine β-casein gene was obtained from the Genbank data base, accession #X14711. Primers were designed, using this sequence, that were complementary to portions of the 5' and 3' regulatory regions of the bovine β-casein gene. The 5' and 3' DNAs were then obtained by performing the polymerase chain reaction (PCR) on bovine white blood cells. The various regulatory regions and pertinent restriction enzyme sites for the 5' promoter region and the 3' polyadenylation region are shown in figures 2 and 3, respectively. Rather than show the entire sequence, the regions have been truncated and the length indicated by the base number at the beginning and end of each row. Exons were identified by underlining. In the 5' β-casein segment, the TATA box was italicized and underlined. The restriction enzyme site, PpumI, was convenient

KpnI

TGGGTACCAGGATAAAATCCACCCC EXON#7 (8028-8525) CCTGTCCGGGGA
 -----+----- - - - - - - - - - -+----- 8510
 ACCCATGGTCCTATTTTAGGTGGGG GGACAGGCCCTT

CCCTTCCCTATTATTGTAAGTCTAAATTTACTAACTGTGCCTGTTAACTTCTGATGTTT
 8511 -----+-----+-----+-----+-----+-----+----- 8570
 GGGGAAGGGATAATAACATTTCAGATTAAATGATTGACACGGACAATGAAGACTACAAA

GTATGATATTTCGA INTRON #7 (8526-9126) TTTTATTTTATTAT
 8571 -----+-----+-----+-----+-----+-----+----- 9110
 CATACTATAAGCT AAAATAAAATAATA

TTTTTTTATTTTAAAGTCTAAGAGGATTTCAAAGTGAATGCCCCCTCCTCACTTTTGGT
 9111 -----+-----+-----+-----+-----+-----+-----+9170
 AAAAAATAAAAAATTCAGATTCTCTAAAGTTTCACTTACGGGGGAGGAGTGAAAACCA

AAGCTTTAGGATAT INTRON #8 (9169-9898) TCTATGTTTGTCTCTA
 9171 -----+-----+-----+-----+-----+-----+-----+9890
 TCGAAATCCTATA AGATACAAACAAGAGAT

TTCCACAGAAATGACTGCGACTGG EXON#9 (9899-10221) AACATATCAAAATA
 9891 -----+-----+-----+-----+-----+-----+-----+10130
 AAGGTGCTTAACTGACGCTGACC TTGTATAGTTTTAT

TGTATAAAAAATAATTTCTGGAATTGTGATTATTATTCTTTAAGAACTATTTCCTAAC
 10131 -----+-----+-----+-----+-----+-----+-----+10190
 ACATATTTTATTAAAGACCTTAACACTAATAATAAGAAATCTTAGATAAAGGATTGG

AGTCATTTCAATAAATAATCCTTAGGCATATTTAAGTTTCTGTCTTTATTATATTTT
 10191 -----+-----+-----+-----+-----+-----+-----+10250
 TCAGTAAAGTTATTTAATTAGGAATCCGTATAAAATTCAAAAGAACAGAAATAATATAAAA

TTTTAATGAAATTGGTCTCTTTATTGTTAACTTAAATTTATCTTTGATGTTAAAAAGAGC
 10251 -----+-----+-----+-----+-----+-----+-----+ 10310
 AAAATTACTTTAACCAGAGAAATAACAATTTGAATTTAAATAGAACTACAATTTTCTCTCG

BamHI

TGTGGAATAATAAATTGGATCCAATTCCC
 10311 -----+-----+-----+-----+-----+-----+-----
 ACACCTTTTAATTTAACCTAGGTTAAGGG

FIG. 3. Features of the 3'-casein portion of the expression vector.

Design of synthetic genes

The native sequences from plants for both thaumatin (9) and brazzein (13) were known. Synthetic oligonucleotides were designed to code for the protein sequences to generate the genes. The DNA sequences found in the plant genes of thaumatin and brazzein were not used due to the frequency that some codons were used in plants as compared to mammals. There were one to four codons that could potentially code for each of the twenty amino acids. Different codons could be used with varying frequencies depending on the type of organism that might be analyzed. Investigators have shown that levels of thaumatin protein can be dramatically improved by using the codons that were commonly used normally in the organism, which was yeast in their case (19). The common codon usage of milk protein genes and plant genes of the family *Marantaceae*, to which *Thaumatococcus* belongs, was determined using Wisconsin GCG Sequence Analysis Software Package (Genetics Computer Inc.). To illustrate this concept, the analysis of codon frequencies of milk proteins for the amino acid leucine was compared to the frequencies of codons used for plant proteins (table 5). TTA, the codon that was not used in the milk proteins, was used frequently in plants. Therefore, if the native sequence that contained the codon TTA were used, the expression system would probably not have sufficient tRNA's capable of recognizing the codon and translating it into the corresponding amino acid, which in this case, was leucine. Therefore, using the published sequence of thaumatin (Genebank accession #A15660), and the sequence contained in the patent (36) for brazzein, the nucleic acid sequences were modified in the synthetic genes as compared to the native plant genes. The gene was designed to contain the common codons used in the mammary gland genes. The codons used in the native plant thaumatin or brazzein genes were preserved unless they were used with frequencies less than 20% in mammary gland genes.

TABLE 5. Common codon frequency of leucine

Plant protein codon usage		Mammalian milk protein codon usage	
CTG	0.03	CTG	0.47
TTG	0.28	TTG	0.11
CTC	0.00	CTC	0.22
CTA	0.19	CTA	0.02
CTT	0.17	CTT	0.18
TTA	0.33	TTA	0.00

The remainder (residues 3 through 15) of the alpha S1-casein secretory signal peptide sequence on the 5' end of the flavor/sweetener gene up to the PpumI restriction site (figure 2) was coded by synthetic sequence. The nucleic acid sequence for the signal peptide was modified slightly due to the construction of a PpumI site for cloning purposes, but the amino acid sequence was preserved. A comparison between the wild-type nucleic acid sequence for alpha S1-casein signal peptide and the sequence that was used in the vector is shown in figure 4.

Synthetic:^{1,2} *ctg* atc *ctg* acc tgc ctg gtg gcc gtc gcc *ctg* gcc
 L I L T C L V A V A L A

Bovine β -casein: ctc atc ctt gcc tgc ctg gtg gct ctg gcc ctt gca
 L I L A C L V A L A L A

Bovine α S1-casein: ctc atc ctt acc tgt ctt gtg gct gtt gct ctt gcc
 L I L T C L V A V A L A

1: underline indicates differences when compared to β -casein

2: italic indicated differences when compared to alpha S1-casein

FIG. 4. Comparison of α -s1-casein with β -casein and the synthetic sequence used.

Two stop codons followed by a SalI restriction enzyme site for thaumatin and brazzein gene sequences were positioned on the 3' end of the synthetic sequence as can be seen in figures 5 and 6, respectively. Once the desired sequence of the genes was known, synthetic oligonucleotides were made, either by National Biosciences, Inc. or the USU Biotechnology Service Laboratory.

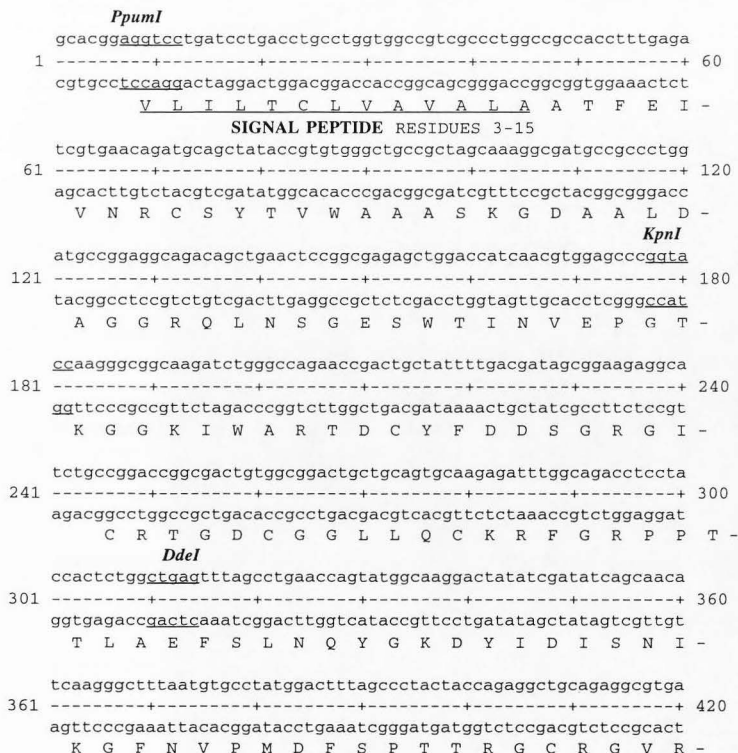


FIG. 5. Sequence of the synthetic thaumatin construct.

gatgcgctgccacatcgctcgccagtcgacctgccaagctgaaggccctggaggcggt
 421 -----+-----+-----+-----+-----+ 480
 ctacgcgacggctgtagcagccgggtcacgggacggttcgacttccggggacctccgcga
 C A A D I V G Q C P A K L K A P G G G C -
SpeI
 gcaatgacgctgcacgctgtttcagactagtgtgagtattgtgcaccactggcaagtgcg
 481 -----+-----+-----+-----+-----+ 540
 cgttactgcggacgtggcacaaagtctgaatcactcataacgacgtggtgacctttcacgc
 N D A C T V F Q T S E Y C C T T G K C G -
 gccctaccgagtatagcagatttttaagagactgtgcacctgacgccttttagctatgtgc
 541 -----+-----+-----+-----+-----+ 600
 cgggatggctcatatcgctctaaaaaattctctgacacgggactgcggaaatcgatacacg
 P T E Y S R F F K R L C P D A F S Y V L -
 tggacaagcctaccacgctgacctgacctggcagcagcaactatagagtaccttttggc
 601 -----+-----+-----+-----+-----+ 660
 acctgttcggatggtggcactggacgggacctgcgtcgttgatatctcactggaaaacgg
 D K P T T V T C P G S S N Y R V T F C P -
Sall
 ctaccgcctagtaggtcagcgtgcg
 661 -----+-----+-----+-----+ 686
 gatggcggatcatccagctgcgacgc
 T A * *

FIG. 5. Continued.

PpumI
 gctcggaggtcctgatacctcacctgcctggctgcgctggctcttgcgcacaaatgaaga
 1 -----+-----+-----+-----+ 60
 cgagcctccaggaactaggtggacggaccagcggcaccgagaacggctgtttacgttct
 V L I L T C L V A V A L A D K C K K -
 SIGNAL PEPTIDE RESIDUES 3-15
 aagtgtacgagaattatcctgtgagcaagtgcgacctggccaaccagtgttaattacgact
 61 -----+-----+-----+-----+ 120
 ttcacatgctcttaataaggacactcgttcacgggtcgacgggttggtcacattaatgctga
 V Y E N Y P V S K C Q L A N Q C N Y D C -
 gcaaaactggacaagcagccagatccggggaatgcttctatgacgaaaagcggaacctgc
 121 -----+-----+-----+-----+ 180
 cgttttgacctgttcgtgcggtctagggcccttacgaagatactgcttttcgccttggacg
 K L D K H A R S G E C F Y D E K R N L Q -
Sall
 agtgtattggcgattatttgcggctagtaagtccagcgcg
 181 -----+-----+-----+-----+ 221
 tcacataaccgctaataacgcgcatcattcagctgcgcgcg
 C I G D Y C G * *

FIG. 6. Sequence of the synthetic brazzein construct.

A general outline of the order in which the different components were cloned into the plasmid backbone was shown in figure 7. AgeI and AatII were used to ligate the 5'β-casein region into the pGFP-C1 vector (figure 7, first and second image). This particular vector was used because there was no PpuMI site in the vector, which was necessary for subsequent manipulations. When the synthetic genes were ligated into the vector (figure 7, fourth image), the AgeI and PpuMI sites were used and the 1.5 kb sequence between PpuMI and AgeI was lost from the vector sequence. The 3'β-casein region was ligated into the vector in a similar manner using the BamHI and KpnI restriction enzyme sites (figure 7, third image).

Construction of the synthetic genes

The oligonucleotides that constituted both brazzein and thaumatin were generated synthetically. Due to the size of the thaumatin gene, as compared to the brazzein synthetic gene, the gene had to be divided into four pieces to make the synthetic thaumatin gene. As diagrammed in figure 8, eight separate thaumatin oligonucleotides were synthesized and cloned into plasmids as four separate groups (second line of figure 8). These four groups were then ligated together as shown in the third line of figure 8. To achieve orderly ligation, restriction enzyme sites were designed into the gene so that once the individual clones were obtained they could be ligated together to form the final synthetic gene product of 686 bp (fourth line, figure 8). The middle enzyme, DdeI, was a non-palindromic enzyme; therefore, the ligation involving that restriction enzyme site could only be achieved in one orientation.

Brazzein was only 220 bp long by comparison, so only two oligonucleotides were used to generate the brazzein gene. Therefore, the intermediate steps required to make the thaumatin gene were not necessary for the construction of the brazzein synthetic gene (figure 9). The details of these steps were described completely in the Methods section.

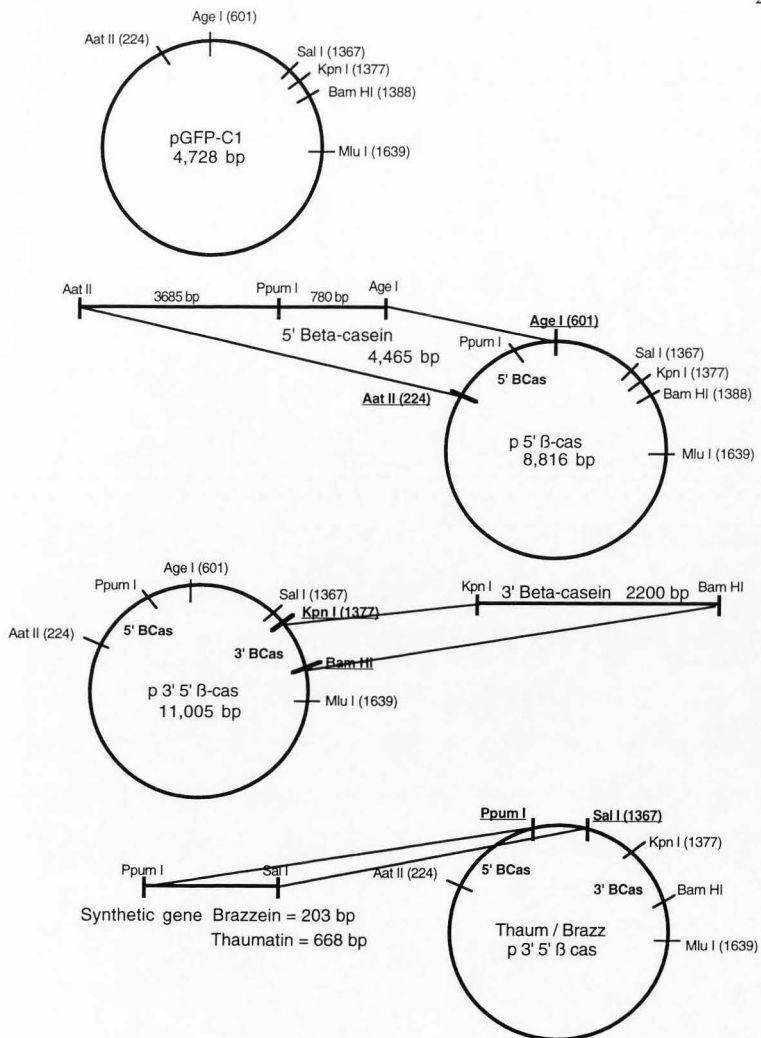


FIG. 7. Design of the genetic expression vector.

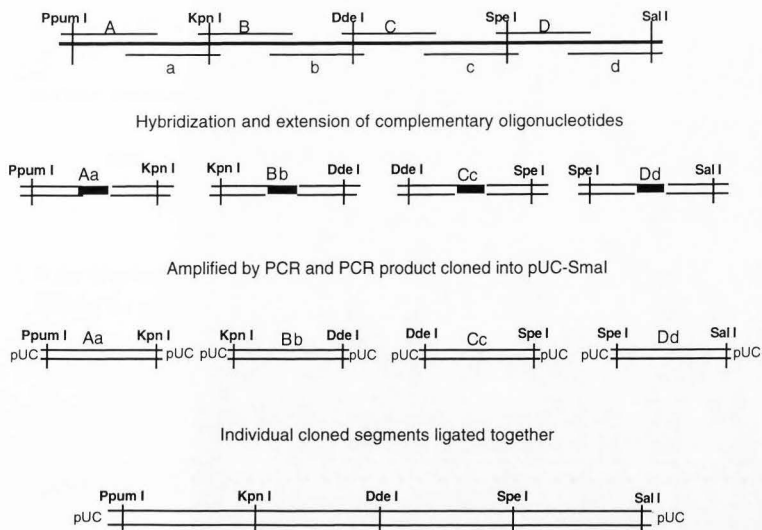


FIG. 8. Construction of the thaumatin synthetic gene.

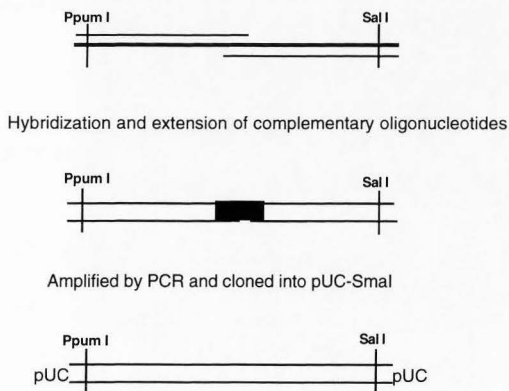
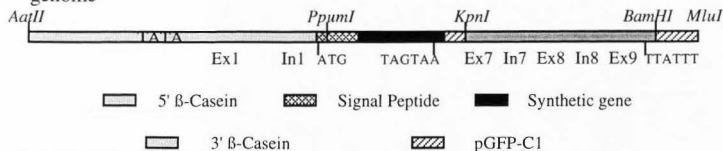
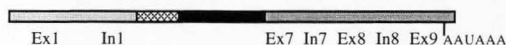


FIG. 9. Construction of the brazzein synthetic gene.

1. Expression vector was microinjected into pronucleus and integrated into the host genome



2. In the mammary tissue, the DNA sequence was transcribed into mRNA. Sequences flanking the promoter and polyA were not transcribed.



3. mRNA was processed, and the regulatory introns were spliced out.



4. Translation of processed mRNA into the corresponding amino acids.



5. Signal peptide was cleaved during transport across the secretory vesicle membrane



FIG. 10. Anticipated transgenic mRNA and protein processing *in vivo*.

Figure 10 illustrates the expected cell processing of gene products once the expression vector had been incorporated into the host genome. The products of the expression vector would undergo a series of modifications that were cell-cycle dependent

that would ultimately result in only the flavor/sweetener proteins, not the fused signal peptides, being present in the mammary milk secretory vesicle and secreted into the milk.

Objective 2: Generation of transgenic lines to test the expression vector

Mice were used as the model to test the expression vector. The synthetic gene was injected into the pronuclei of one-cell stage embryos. Mice were super-ovulated and sacrificed and the reproductive tracts collected to obtain the embryos. The surviving embryos after microinjection were transferred into recipient mice that were pseudo-impregnated with a vasectomized male mouse (27). The offspring were screened (using PCR and Southern blot hybridization) 3 weeks after birth to detect whether they had incorporated the transgene. The founder animals, shown to have the transgene, were then used to establish further generations of transgenic mice. This effort was accomplished by a team of people in Drs. John Morrey's and Kenneth White's laboratories at Utah State University. My involvement was to screen the resulting mice for the transgenes.

Objective 3: Detection of protein in the milk of the transgenic mice

An enzyme-linked immunosorbant assay (ELISA) was developed by using crude native thaumatin plant protein as a standard (Sigma Chemical Co., St. Louis, MO). The assay was developed to detect and quantify the levels of thaumatin and brazzein proteins. Antibodies were generated against the thaumatin and brazzein proteins in both goats and rabbits and the serum was isolated for use in the assay. A goat capture antibody specific for one of the proteins was bound to the ELISA plate. The standards or milk samples, potentially containing the transgenic proteins, were sandwiched between the goat antibody and a rabbit antibody specific to the protein. The detection system made use of an anti-rabbit antibody conjugated to horse radish peroxidase. When the substrate for the enzyme was added, a color change corresponding to the concentration of the antigen was observed.

Additionally, a mouse milk sample from a putative transgenic mouse was analyzed using capillary electrophoresis which separated the proteins according to the charge-to-mass ratio. Migration towards the negative electrode depended on the charge density of the molecule with small, highly positive charged species eluting first, followed by larger, less charged species. The new transgenic proteins should be observed in the milk by this method.

EXPERIMENTAL METHODS

Objective 1: Construction of expression vector

Cloning β -casein regions into pGFP-C1Isolating β -casein DNA fragments

PCR (26) was performed on bovine blood cells using primers designed from the sequence found in the Genbank data base, accession #X14711. The resulting PCR products were fractionated on a 1% Seaplaque GTG agarose gel (FMC Bioproducts, Rockland, ME). The agarose gel was stained by soaking in a 0.5 mg/ml ethidium bromide (Sigma Chemical Co., St. Louis, MO) solution that allowed visualization of the DNA when the gel was viewed under UV light. DNA bands of the predicted sizes of the 5' and 3' β -casein DNA were seen at 4,600 bp and 2,220 bp, respectively, when the correct PCR products were obtained. The previous development and optimization of the PCR was done by Javier Alcazerra and Charoensri Thonabulsombat in Drs. John Morrey's and Ken White's laboratories (Utah State University).

A method was developed to overcome problems encountered with cloning PCR products. PCR products are notoriously difficult to clone into vectors. For example, one problem seen with the use of Taq DNA polymerase is the addition of adenines to the end of each strand that it synthesizes. Various methods have been developed by others to overcome this cloning difficulty, such as 1) using a cloning vector with a thymidine overhang and, 2) using primers with restriction enzyme sites designed into them (7). However, known improvements of PCR cloning were not successful to clone the PCR products of this study.

The following method was successful in cloning the PCR products for this experiment. The product was first digested with proteinase K (8) to presumably remove the residual Taq polymerase and any other proteins that could be interfering with subsequent manipulations. The PCR samples were digested with 50 μ g/ml proteinase K

(Sigma Chemical Co., St. Louis, MO), in 1X proteinase K buffer (0.5% SDS, 5 mM EDTA, 1 mM Tris) at 37°C for 30 min. The proteinase K was then heat-inactivated at 65°C for 20 min so that the enzymes used in the next reactions were not destroyed by any residual proteinase K that was still active. Next, the samples were purified by phenol:chloroform extraction and ethanol precipitation (PC/EtOH)(3).

Cloning of 5' β -casein DNA into pGFP-C1

After suspending the treated PCR product of the 5' β -casein region in TE buffer (10 mM Tris, 1 mM EDTA pH 8.0), it was then cloned into the vector pGFP-C1 (Clontech, Palo Alto, CA). This particular vector was not used because of the presence of green fluorescing protein gene; rather, it was used because of convenient restriction enzyme sites present in the plasmid. During the cloning process, the green fluorescent protein gene was lost from the plasmid. The cloning was achieved by digesting both the vector pGFP-C1 and the purified 5' β -casein PCR product with the restriction enzymes AatII and AgeI (New England Biolabs, Beverly, MA. All enzymes were from New England Biolabs unless otherwise stated.). They were digested with AatII in 1X NEBuffer#4 at 37°C for approximately 2 hours, purified by PC/EtOH, and digested with AgeI with 1X NEBuffer#1 at 37°C for 2 hours. After being suspended in TE buffer, the DNA concentration was measured on both the pGFP-C1 that had been cut with AatII and AgeI and the 5' β -casein product that had been cut with AatII and AgeI. The 4,478 bp 5' β -casein fragment was ligated into the 4,728 bp pGFP-C1 using a one-to-one molar ratio of insert to plasmid, a total DNA concentration of 50 ng/ μ l, 1X T4 DNA polymerase buffer, and 1 unit (U) of T4 DNA polymerase. The ligation mixture was incubated for at least 8 hours at 16°C. An amount of 5 ng of the ligation reaction was transformed into DH5-alpha *E. coli* transformation-competent cells using a Biorad Gene Pulser™ and the protocol accompanying the instrument's instructions (Bio-Rad, Hercules, CA.). The transformation reaction was then plated out onto Luria-Bertani (LB) (1% bacto-tryptone, 0.5% bacto-yeast

extract, 1% NaCl) agar plates containing 30 µg/ml kanamycin (Sigma Chemical Co., St. Louis, MO). Kanamycin antibiotic was used because pGFP-C1 contains a kanamycin-resistance gene. The presence of this gene ensured that only those colonies that contained the plasmid survived in the presence of kanamycin. The plates were incubated overnight at 37°C. The following day, single bacterial colonies were picked from the plates under sterile conditions and inoculated in 2 ml of LB media that contained 30 µg/ml of kanamycin. After incubation of individually picked colonies in LB media at 37°C for 12 - 16 hours, the DNA was isolated using an alkaline lysis miniprep method (30) which purified the circularized plasmid away from the bacterial genome.

To screen the DNA obtained from the colonies for plasmids that contained the 5'β-casein DNA, the uncut plasmids were fractionated on a 1% Seakem LE gel (FMC Bioproducts, Rockland, ME). Those plasmids that migrated differently on the gel were plasmids that possibly contained an insert. Each putative clone was subjected to restriction enzyme analysis and fractionated on 1% Seakem LE agarose to diagnose which clone had an insert in the correct orientation (see Experimental Design, figure 7). A correct clone would result in the following size DNA fragments when cut with the restriction enzymes as indicated; AatII and AgeI: 4.4 kb & 4.7 kb, SpeI: 7.3 kb & 1.8 kb. The clone that exhibited the correct diagnostic bands was then sequenced by the USU Biotechnology Laboratory. Once the correct clone had been identified, it was then used as the backbone into which the remaining DNA fragments were ligated. The correct clone identified at this step, that consisted of 5'β-casein DNA and pGFP-C1, was referred to throughout the text as p5'β-cas.

Cloning 3' β-casein DNA into pUC-19

The transformation of ligation products of gel-purified fragments with p5'β-cas was inefficiently low, if not nonexistent, i.e., no correctly transformed bacteria were obtained. For this reason, it was necessary to use a circuitous approach to clone the remaining

segments into p5'β-cas. This was accomplished by first cloning the PCR products into a simple plasmid, pUC19, and then subcloning the fragment from pUC19 into p5'β-cas.

The 3'β-casein proteinase K-treated PCR product obtained in step #1 was first cloned into the vector pUC19 using blunt end ligation, which does not involve sequence-specific overhanging sticky ends. The PCR product was blunt-ended and phosphorylated using the following reaction conditions: 1 X T4 DNA polymerase buffer, 1 mM ATP, 0.8 mM dNTP's, 50 µg/ml BSA, 20 U polynucleotide kinase, and 6 U of T4 DNA polymerase. The reaction was incubated for 30 min at 37°C and heat inactivated at 75°C for 10 min. After heat inactivation the DNA was purified by PC/EtOH and then suspended in TE buffer.

The pUC was prepared for blunt-ended ligation by digestion with SmaI restriction enzyme that created a blunt end. The digestion reaction was incubated with SmaI enzyme and NEBuffer#2 at 25°C. A small aliquot of approximately 10 ng was removed for analysis by gel electrophoresis to confirm that digestion of the plasmid did occur. When fractionated on a 1% LE agarose gel (FMC Bioproducts, Rockland, ME), a band was visible that corresponded to the predicted size of 2,686 bp. The remaining digested pUC plasmid was purified by PC/EtOH and suspended in TE buffer in preparation for the ligation reaction.

A ligation reaction was prepared that contained the blunt-end, phosphorylated 3' β-casein PCR product and pUC-SmaI at equimolar ratio of ends, 1 U T4 DNA ligase, 1X ATP-free ligase buffer (Stratagene, La Jolla, CA), and 50 µM ATP. After incubation for a minimum of 8 hours at 25°C, 5 ng of the reaction was transformed as described above (Experimental Methods, step #1). Ampicillin was the antibiotic used to select for the pUC-containing bacterial colonies. In the previous experiment, kanamycin was the antibiotic used because the plasmid pGFP-C1 was used. By using the multiple cloning site in pUC19, blue-white selection (3) of Lac operon function was used. The media and agar contained 50 µg/ml ampicillin (Sigma Chemical Co., St. Louis, MO) to select for bacterial

colonies containing the pUC vector; and the agar plates contained 20 µg/ml X-gal (Sigma Chemical Co., St. Louis, MO) and 0.1 mM IPTG (Sigma Chemical Co., St. Louis, MO) to select for an insert in the multiple cloning site.

Single white colonies were picked under sterile conditions and inoculated at 37°C with shaking for 12-18 hours in 2 ml of LB media that contained 50 µg/ml of ampicillin. The DNA was isolated by alkaline lysis and suspended in TE buffer, and the uncut plasmids were fractionated on a 1% Seakem LE gel. Those plasmids that migrated differently when compared to a pUC plasmid control, were plasmids that possibly contained an insert. The correct clone was identified by separate restriction enzyme digestions using KpnI and HindIII. The predicted sizes depended on the orientation of the fragment within the plasmid as noted: orientation 1 - KpnI: 2,224 & 2,682 bp, HindIII: 1,174 & 3,732 bp or orientation 2 - KpnI: 4,902 & 4 bp, HindIII: 1,190 & 3,716 bp (the small fragments would not be visible on a 1% gel). The putative correct clones were confirmed by sequencing the complete insert as performed by the USU Biotechnology Laboratory.

Cloning 3'β-casein into p5'β-cas construct

Both p5'β-cas and pUC-3'β-casein were digested with BamHI and KpnI using a double restriction enzyme digestion in which both enzymes were used in one reaction at a maximum amount of 2.5% total volume per enzyme to avoid having star activity. The reaction also contained NEBuffer for BamHI and 1X bovine serum albumin (BSA). This would result in DNA fragments of the following sizes being produced for p5'β-cas: 9,185 & 11 bp, and for pUC-3'β-casein: 2,696 bp (pUC) & 2,210 bp (3'β-casein). The samples were purified by PC/EtOH after incubation with restriction enzymes at 37°C for 1-2 hours. The plasmid and insert-fragments were not separated from each other before the ligation. The DNA samples were suspended in TE and the concentration of DNA measured by depositing 1 µl of each DNA sample onto ethidium bromide plates (0.5 µg/ml ethidium

bromide, 1 % Seakem agarose) and comparing the degree of fluorescence of the samples to the fluorescence of DNA standards of known concentrations. The ligation reaction was designed to have a 3:1 molar ratio of ends of 3'- β -casein DNA to p5'- β -cas. The ligation reaction did not have more than 50 μ g of DNA per μ l of total reaction volume. The appropriate amounts of DNA were incubated along with 1 U of T4 DNA ligase and 1X NEBuffer at 16°C for a minimum of 8 hours. An amount of 5 ng of the ligation reaction was transformed by electroporation using XL1-Blue Epicurian Coli[®] Electroporation-Competent Cells (Stratagene, La Jolla, CA) that had a transformation efficiency of 5×10^9 colony forming units per μ g of DNA (CFU/ μ g) as compared to home-made cells which typically had a transformation efficiency of 1×10^8 CFU/ μ g. This higher efficiency increased the likelihood of one of the transformants containing the desired clone. A volume of 50 μ l of the transformation reaction was plated onto LB agar plates that contained kanamycin at a concentration of 30 μ g/ml. Only a colony that contained the p5'- β -cas plasmid could survive on the kanamycin media; therefore, the possible clones that could be present on the plates were 1) p5'- β -cas that had re-ligated with the 11 bp fragment, 2) p5'- β -cas + pUC, and 3) p5'- β -cas + 3'- β -casein. Any aberrant ligation products involving pUC (not kanamycinresistant) without p5'- β -cas would not result in transformed colonies. The most efficient way to identify the desired clone of p5'- β -cas + 3'- β -casein was by using a radiolabeled 3'- β -casein probe that would bind to the DNA in those clones that contained the 3'- β -casein DNA region. After incubation of the plates overnight at 37°C, the colonies were transferred and bound onto a nitrocellulose membrane (Schleicher & Schuell, Keene, NH) (30) in preparation for analysis by colony hybridization.

The DNA template used for the random labeling reaction to make the probe was generated from the pUC-3'- β -casein clone. It was digested with BamHI and EcoRI using NEBuffer for EcoRI, and 5 X BSA at 37°C for at least 1 hour. This generated two fragments: the 3'- β -casein DNA segment along with 16 bp of the pUC plasmid (2,236 bp) and the remainder of the pUC plasmid (2,670 bp). The DNA bands were fractionated on a

1% Seaplaque GTG agarose gel (FMC Bioproducts, Rockland, ME) and the 2,236 bp fragment was excised from the gel. The DNA was isolated from the gel slice using β -agarase digestion (protocol with product insert, New England BioLabs, Beverly, MA). The β -agarase cleaved the complex carbohydrates into polysaccharides that were no longer capable of forming a gel-matrix. The protocol for accomplishing this was as follows: the gel slice was weighed, placed in a microcentrifuge tube and incubated in one tenth volume β -agarase buffer at 70°C for 10 min. It was placed in a 40°C water bath and the gel slice was allowed to cool to that temperature which took approximately 2 minutes. β -agarase enzyme (1 U) was added for every 200 mg of gel, and the reaction was incubated at 40°C for 1-2 hours with regular vortexing. Once the reaction appeared liquid, 3M sodium acetate solution (pH 5.2) (Mallinckrodt Baker, Inc., Paris, KY) was added to a final concentration of 0.3 mM and placed on ice for 10 min. This caused agarose to precipitate out of solution. Centrifugation at 12,000 rpm for 15 min pelleted the precipitated agarose. The supernatant containing the DNA was transferred to a fresh tube. The DNA was precipitated using ethanol precipitation (EtOH) and suspended in TE buffer. DNA concentration was measured on a DNA fluorometer according to the instructions given with the instrument (Hoefer Scientific Instruments, San Francisco, CA). An amount of 25 ng of template DNA was heated to boiling temperature for 5 min and quenched on ice. The following reaction components were added to make a total volume of 20 μ l: 2 μ l reaction mixture (Boehringer Mannheim, Indianapolis, IN), 0.5 mM of each dNTP (Pharmacia Biotech Inc., Piscataway, NJ), 5 μ l 32 P dCTP 3000 Ci/mmol (ICN Biochemicals, Costa Mesa, CA) and 1 μ l Klenow polymerase (Boehringer Mannheim, Indianapolis, IN). The reaction was incubated at 37°C for 30 min and stopped by adding 1 μ l of 0.5 M EDTA. The efficiency of the reaction in terms of percent incorporation was determined by removing and dispensing 1 μ l of the reaction on a DE81 Whatman filter (Baxter, McGaw Park, IL). The DE81 filter was then subjected to several washing steps (23) which removed unincorporated nucleotides. Using the ratio of the "before" and "after" radioactive counts the

reaction efficiency was calculated. A reaction efficiency of 25% was sufficient for colony hybridization.

Before the filters containing the DNA could be incubated with the probe, they were floated DNA-side up on a solution of 6X SSC for 5 min, completely submerged in the 6X SSC for 5 min, and then sealed in a "seal-a-meal" bag with prehybridization solution (5X Denhardt's, 5X SSC, 1% SDS, and 100 µg/ml denatured salmon sperm DNA, Sigma Chemical Co., St. Louis, MO) at a volume of 1 ml per cm² of membrane. The bag was incubated at 68°C for 30 min, and then the prehybridization solution was drained and fresh hybridization solution added which contained the probe. This was incubated at 68°C for at least 8 hours. Filters were then washed according to medium stringency conditions (30). After wrapping the filters in cellophane, Kodak XAR film (VWR Scientific, Salt Lake City, UT) was exposed to the filters for 1 day. After developing the film, potential positive clones were identified and matched to the original agar plates. A number of putative colonies were inoculated individually in 2 ml LB media with 30 µg/ml kanamycin and incubated overnight at 37°C. The DNA was then isolated by alkaline lysis miniprep, and the correct clone was identified. A clone that had only the 3' β-casein DNA within p5'βcas in the correct orientation was identified by restriction enzyme analysis. The reaction conditions for each enzyme reaction are given in parentheses as follows: HindIII (1X NEBuffer#2, 37°C): 9.9 kb & 1.2 kb, BamHI & KpnI (1X NEBuffer for BamHI, 5X BSA, 37°C): 2.3 kb & 8.8 kb, SalI (1X NEBuffer for SalI, 5X BSA, 37°C): 11.1 kb and XbaI (NEBuffer #2, 37°C): 3.4 kb, 4.5 kb & 3.2 kb.

The clone that contained DNA fragments of 5'β-casein and 3'β-casein in the pGFP vector was referred to throughout the text as p5'3'β-cas.

The overlapping regions were hybridized together and the single-stranded portions were filled in by T4 DNA polymerase to generate the four double-stranded segments used to make the thaumatin gene (figure 8 overview). One μg of each oligonucleotide was suspended in 1X Sequenase buffer (USB, Cleveland, OH) to give a final volume of 19 μl . The reaction mixture was incubated at 70°C for 10 min to denature any oligonucleotides that could have aberrantly hybridized while in solution. The hybridization of the overlapping region then occurred while the reaction was cooled slowly to 37°C. The single-stranded portion of the hybridized oligonucleotides was filled in by the addition of 1 μl (10 U) of Sequenase Version 1.0 enzyme (USB, Cleveland, OH) and 2 μl of a 2.5 mM dNTP mix to give a final reaction volume of 22 μl . The reaction occurred at 37°C for 30 min. The reaction was heat-inactivated by incubation at 70°C for 10 min. The DNA was then purified by PC/EtOH and the pellet suspended in 10 μl of TE buffer. DNA concentration was measured using an agarose ethidium bromide plate as previously described.

Amplification using PCR

It was necessary to amplify each fragment by PCR in order to generate sufficient DNA concentrations for subsequent manipulations (figure 8, overview). The primers for PCR used for each segment were indicated in figure 11. The PCR primers were synthesized by the USU Biotechnology Laboratory. The PCR was performed using the following reaction conditions: 50 ng of DNA template, 10 μM of each primer, 5 U PFU enzyme (Stratagene, LaJolla, CA), 50 $\mu\text{g/ml}$ BSA (Miles Inc., Kankakee, IL), 25 μM each dNTP, dH_2O to a final volume of 50 μl and 1X buffer (Invitrogen Corp., San Diego, CA). For each fragment, five identical reactions were run that differed only by the buffer used. The optimization buffers varied in pH and MgCl_2 concentrations. The thermocycler (Perkin Elmer Cetus, Norwalk, CO) was programmed with the following cycling conditions: 1 cycle for 7 minutes at 95°C and 30 cycles consisting of the following

temperatures: denaturation for 30 sec at 95°C, annealing for 30 sec at 58°C, and extension for 40 sec at 72°C. A volume of 5 µl of the reaction product was run on a 2.5% Seaplaque GTG gel to detect the predicted PCR product sizes that were indicated at the bottoms of figures 11 and 12.

Preparation of the PCR product for cloning into pUC plasmid

The remaining 45 µl of PCR product was treated with proteinase K, which improved the cloning efficiency. Concentrations of 50 µg/ml of proteinase K and 1X proteinase K buffer were used and the reaction was incubated for 30 min at 37°C. It was then heat-inactivated for 20 min at 65°C. The samples were purified by PC/EtOH and the pellet was suspended in TE buffer. The DNA ends were phosphorylated and blunt-ended using the following reaction conditions: 1X T4 DNA polymerase buffer, 1 mM ATP, 0.8 mM dNTP's, BSA to 50 µg/ml, 20 U polynucleotide kinase, and 6 U of T4 DNA polymerase. The reaction was incubated for 30 min at 37°C and heat-inactivated at 75°C for 10 min. The reactions were purified by PC/EtOH and the pellet was suspended in TE buffer.

The pUC plasmid was prepared for blunt-ended ligation as described in a previous section of Experimental Methods.

Ligation and screening of ligation products

Each PCR-generated fragment of thaumatin and the synthetic sequence for brazzein was ligated into pUC-SmaI using a 10:1 molar ratio of insert to plasmid. Ligation reaction components were ATP-free ligation buffer, 50 µM ATP, and 2U T4 DNA ligase. The final DNA concentrations were between 40-50 ng/µl. The reactions were incubated at 25°C overnight.

A volume of 1 µl of each ligation reaction was used to transform bacterial cells that were prepared for heat-shock transformation (30). After transformation, the culture was

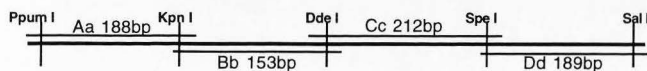
plated onto LB agar plates containing 50 µg/ml ampicillin, IPTG, and X-gal, which allowed for blue-white selection of plasmids containing an insert. The plates were incubated overnight at 37°C. A volume of 2 ml LB broth, containing 50 µg/ml ampicillin, was inoculated with single white colonies and incubated overnight at 37°C in a shaker. DNA was obtained using the alkaline lysis method. The potential clones were then screened by a double-digestion with BamHI and EcoRI restriction enzymes in the presence of NEBuffer for EcoRI and 100 µg/ml BSA. The reactions were incubated at 37°C for 2 hours and then run on a 2.5% Seaplaque GTG gel. Correct putative clones had bands at the following sizes: thaumatin A: 209 bp, thaumatin B: 174 bp thaumatin C: 233 bp and thaumatin D: 210bp.

The identify of the clones was confirmed by DNA sequencing in both directions. Alkaline lysis miniprep DNA was extracted with phenol:chloroform once, chloroform twice, and then ethanol precipitated. They were then sequenced by the USU Biotechnology Laboratory with the primers used previously to amplify the fragments.

Combining thaumatin segments

Joining segments A&B and C&D

The relationship between the four thaumatin fragments is shown in figure 13. The different fragments/segments, that were cloned into the vector pUC, were referred to as A,



Overlapping region between Aa and Bb = 18bp

Overlapping region between Bb and Cc = 20bp

Overlapping region between Cc and Dd = 18bp

FIG. 13. Relationship of thaumatin segments with each other.

B, C & D. Both A and B were digested with KpnI in the presence of NEBuffer for KpnI and 1X BSA for a minimum of 1 hour at 37°C. At this point, the thaumatin fragments were still attached to pUC at one end, but both had a KpnI free overhanging end. Once the KpnI enzyme digestion was completed, both the A and B reactions were purified by PC/EtOH. After suspension in TE buffer and measurement of the DNA concentration, A and B were ligated together using an equal DNA ratio with T4 DNA ligase and 1X T4 DNA ligase buffer at 16°C overnight. The ligation product was then cleaned up once again by PC/EtOH. In preparation for PCR which was done to amplify the AB product, the pUC-AB was linearized using an enzyme, XbaI, that was only present in the pUC plasmid and not in either of the thaumatin segments. The enzyme reaction conditions were as follows: NEBuffer#2, 1X BSA, and XbaI enzyme.

After incubation at 37°C for several hours and purification by PC/EtOH, PCR was performed on the ligation products using PCR primers A+ and B- (figure 11) under the same reaction and cycling conditions that were used previously to amplify the individual fragments as described on page 35.

The generation of C and D ligated together was performed in the same way as the generation of A and B ligated together in the above description, except that SpeI restriction enzyme was used instead of KpnI, and primers C+ and D- were used for the PCR.

The PCR products containing AB and CD fragments were fractionated on a large, 50 cm long, sequencing gel apparatus (Bio-rad, Hercules, CA) that would allow good separation of bands from aberrant products such as AA, BB, CC, or DD. A 10% polyacrylamide gel (PAGE) was prepared (30). The bands at 323 bp for AB and 383 bp for CD were excised from the gel. The gel slice was placed in a microcentrifuge tube, crushed, and subsequently isolated by incubation in PAGE elution buffer (0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA (pH 8.0) and 0.1% SDS) overnight at 37°C while shaking. The reaction was centrifuged at 12,000 rpm for 1 min to

pellet the polyacrylamide. The supernatant was transferred to a separate tube and ethanol precipitated to remove the DNA from the gel. The DNA pellet was dried and suspended in TE buffer, and the concentration was measured using ethidium bromide plates.

Combining thaumatin AB and thaumatin CD

The PCR products AB and CD that were purified from the PAGE were digested with the restriction enzyme DdeI using NEBuffer #4 at 37°C for 1-2 hours and then purified by PC/EtOH. A ligation was then done using the same conditions as described in the previous section when A,B and C,D were being joined. However, in this instance there was no pUC plasmid present because the AB and CD PCR products were not subcloned into pUC19 but were purified from the polyacrylamide gel. PCR was then performed on the ligation mixture with the primers A+ and D-. As indicated above, the same reaction and cycling conditions that were used in the original amplification of the synthetic fragments were utilized. This PCR reaction yielded a PCR product of 686 bp. The PCR product was subcloned into pUC-SmaI as described in B.5 of the methods section. A volume of 1 µl of the ligation was used for transformation by heat-shock treatment and plated on LB/X-gal/IPTG/ampicillin plates to screen the products of the ligation reaction. Putative plasmids that contained an insert were selected using blue/white selection and incubated in LB media that contained 50 µg/ml ampicillin. After the DNA was isolated from the inoculum, putative clones that contained the complete thaumatin synthetic gene were identified by double restriction enzyme digestion with BamHI and EcoRI in the presence of NEBuffer for EcoRI and 1 X BSA for 2 hours at 37°C. Those clones that had a band in the region of 686 bp when fractionated on a 2.5 % agarose gel were sequenced by the USU Biotechnology Laboratory. Sequencing was important to ensure that the entire sequence was exactly as designed and that no random errors had occurred as a result of using PCR to amplify the different ligated segments.

Inserting flavor or sweetener genes into p5'3'βcas construct

Preparation and ligation of plasmid and insert

As outlined in figure 7, the synthetic genes were ligated into p5'3'βcas at the restriction enzyme sites SalI and PpuuI. The p5'3'βcas plasmid and brazzein-pUC or thaumatin-pUC were digested with PpuuI, the reaction was incubated with NEBuffer #4 at 37°C for 2 hours, and then digested with SalI using NEBuffer for SalI and 1X BSA at 37°C for several hours after purification by PC/EtOH. A small amount of each reaction was fractionated by agarose gel electrophoresis to check that the complete digestion had occurred before the reactions were prepared for ligation. The samples of the plasmid (p5'3'βcas) that had been cut with PpuuI and SalI, as well as the thaumatin and brazzein genes that had been cut with PpuuI and SalI, were purified by PC/EtOH and then suspended in TE buffer. The DNA concentrations were measured. A ligation reaction was carried out using a 3:1 equimolar ratio of ends of insert (flavor/sweetener genes) to plasmid (p5'3'βcas), 1X T4 DNA ligase buffer and 4U T4 DNA ligase at 16°C for a minimum of 8 hours.

Screening ligation of flavor genes into p5'3'βcas

To identify a correct clone as seen in figure 7, fourth panel, bacteria containing the ligation reaction were transformed using Stratagene XL1-Blue Epicurian Coli[®] Electroporation Competent cells and the Bio-rad Gene Pulser apparatus. After incubation of the electroporation reaction at 37°C for 1 hour, the bacterial culture was then diluted 1:5 and spread on LB agar plates that contained 30 µg/ml kanamycin and were incubated overnight at 37°C. The reaction was diluted to obtain a concentration of colonies on the surface of the plates in which individual clones could be identified. As previously described in Experimental Methods, the colonies on the agar plates were transferred onto nitrocellulose filters, and the plasmids were screened using a ³²P labeled probe specific for

either brazzein or thaumatin. The template for the probe was obtained by double digestion of thaumatin-pUC or brazzein-pUC with BamHI and EcoRI using NEBuffer for EcoRI and 1X BSA at 37°C for 2-3 hours. The reactions were fractionated on a 2.5 % Seaplaque GTG agarose gel, and the bands at 696 bp and 231 bp for thaumatin and brazzein, respectively, were excised and purified by β -agarase digestion as described previously in Experimental Methods. The radiolabeled probes were prepared using the following reaction conditions: 25 ng DNA template that had been boiled and then quenched on ice, 2 μ l reaction mixture, 0.5 mM of each dNTP, 5 μ l 32 P dCTP 3000 Ci/mmol, 1 μ l Klenow polymerase and dH₂O up to 20 μ l. The reaction was incubated at 37°C for 30 min and stopped by adding 1 μ l of 0.5 M EDTA. After calculating the specific activity of the probe, it was boiled to denature the DNA and added to the filters. The filters along with the probe were incubated overnight at 68°C. The unbound or weakly nonspecific-bound probe was washed from the filter using the conditions described previously in Experimental Methods. Correct putative clones were identified after exposure of the X-ray film to washed filters and subsequent development. These colonies were picked, the DNA isolated by alkaline lysis miniprep, and then screened using a digest with PpumI and then with SalI using the reaction conditions described previously in Experimental Methods. Those clones that contained the correct insert liberated bands corresponding to the size of thaumatin and brazzein in addition to bands of predictable sizes of the plasmid. The orientation of these clones was confirmed with further restriction mapping. Several putative clones were then sequenced by the USU Biotechnology Laboratory. They were sequenced across the junctions between 5'3' β -casein and flavor genes to confirm that they were in the correct translational reading frame.

Objective 2: Generation of transgenic lines to test the expression vector

Preparation of expression vector for microinjection

The expression vector containing the genes of interest had to be removed from the pGFP-C1 plasmid backbone in which it was cloned to prepare it for microinjection into pronuclear-staged embryos. The thaumatin expression vector was digested with AatII using NEBuffer #4 at 37°C for several hours, purified by PC/EtOH, and then digested with MluI enzyme using NEBuffer#3 at 37°C for several hours. The reaction was then fractionated on a 1% Seaplugue GTG gel and the band at 5.8 kb was excised. The DNA was isolated from the gel using β -agarase digestion as previously described in the protocol supplied with the enzyme (New England BioLabs, Beverly, MA). This was followed by a further purification step using an Elutip chromatography column (Schleicher and Schuell, Keene, NH). The DNA obtained from the β -agarase reaction was suspended in a low salt buffer (0.2 M NaCl, 20 mM Tris HCL, 1.0 mM EDTA (pH 7.4)), instead of TE buffer, in preparation for the Elutip procedure. The protocol given in the instructions for the Elutip column was followed. After ethanol precipitation the pellet was suspended in microinjection buffer (5 mM Tris pH 7.2, 1 mM EDTA pH 8.0) at a concentration of 2 ng/ μ l.

When performing restriction enzyme digestions on the brazzein expression vector (Brazzein-5'3' β -casein), an extra MluI site, other than what had been predicted, was identified within the pGFP-C1 portion of the expression vector. It was confirmed by restriction enzyme analysis that there were no significant alterations within the expression vector that would be injected into the embryos. Therefore, to liberate the brazzein expression vector from the plasmid, the DNA was digested first with AatII as described above. Following purification with phenol/chloroform and ethanol precipitation, the DNA was digested with BamHI using NEBuffer for BamHI and 1X BSA at 37°C for 2 hours. The remainder of the Elutip purification procedure in preparing the brazzein expression vector was the same as that described above for thaumatin.

Murine embryo transfer

The following procedure was performed by a team of people in Drs. John Morrey's and Kenneth White's laboratories (Utah State University, Logan, UT). Mice were kept on an automatic 5 am - 5 pm light cycle. With embryo collection taking place on day 4, the treatment regimen for superovulating the donor mice was as follows:

- Day 1: 6U PMSG (CalBioChem, San Diego, CA) (0.85% saline at 1U/0.1 ml) was injected intra-peritoneally (IP) between 10:00 am and 11:00 am.
- Day 3: 6U hCG (Sigma Chemical Co., St. Louis, MO) (0.85% saline at a concentration of 1U/0.1 ml) was given IP between 9 am and 10 am.
- Day 4: embryo collection at 8:00 am: donors were euthanized by CO₂ inhalation.

Embryos were collected from the ampulla and washed in M2 media that contained 350 µg/ml hyaluronidase (Sigma Chemical Co., St. Louis, MO). The embryos were washed with the M2 media, and then hyaluronidase media for 15 sec or until the cumulus cells fell off. They were then washed in M2 media to remove the hyaluronidase. These embryos were held in wash buffer at 37°C until microinjection was performed. Microinjection of the brazzein or thaumatin expression vectors into the embryos was done between 9:30 am to 1:00 pm on the day of embryo collection. After microinjection the embryos were transferred into the infundibulum of recipient mice. The recipients were pseudoimpregnated by putting them with vasectomized males at 1 to 5 pm on day 3, so that the reproductive tract was hormonally prepared to receive the embryos (27).

Screening for potential transgenic animals

Preparation of samples

The pups from the mice that were implanted with the microinjected embryos were weaned at about 3 weeks of age, at which time a tail biopsy was taken. Using a razor blade that had been sterilized by being dipped in 100% ethanol and flamed, 0.5 cm was removed

from the end of the tail and placed in a labeled sterile microcentrifuge tube. The DNA was isolated by digestion with 100 µg/ml proteinase K in 1X proteinase K buffer for 12 hours at 42°C.

The samples were not purified by PC/EtOH because this would increase the likelihood of the samples becoming contaminated.

Screening by PCR

The genomic DNA from putative transgenic mice were screened by PCR using the primers B+ and C- that were used to generate the individual synthetic thaumatin fragments. The PCR reaction conditions were as follows: 0.4 µM each primer, 1X PFU buffer which contains 2 mM MgCl₂ (Stratagene, LaJolla, CA), 25 µM dNTP's, 1 mM MgCl₂ (Gibco BRL, Grand Island, NY), 50 µg/ml BSA, 1.25 U PFU enzyme and 1 µl of DNA template in a total reaction volume of 25 µl. The PCR reaction was performed on all of the animal samples in addition to various positive and negative controls. When the PCR products were fractionated on a 2.5 % Seakem LE agarose gel, a faint band was evident in the region of the predicted size of 345 bp.

Screening by genomic Southern blot hybridization

In addition to screening by PCR, some of the DNA samples were also screened using Southern blot analysis of the genomic DNA. The protocol outlined here, was for genomic DNA but the transfer of the DNA and the probe preparations were the same for the blots that were done on the agarose gels that contained the fractionated PCR products.

The tail samples that had been digested with proteinase K were purified by extraction with phenol:chloroform:isoamyl alcohol (Sigma Chemical Co., St. Louis, MO), two chloroform:isoamyl alcohol (24:1) extractions and ethanol precipitation. The concentration and purity of the samples was determined on a spectrophotometer (Shimadzu Scientific Instruments, Inc., Columbia, MD) at wavelengths 260 nm and 280 nm after

suspending the DNA in TE buffer. An amount of 10 μ g of each DNA sample was digested with the restriction enzyme *Hinf*I, which is a four base pair-cutting enzyme, and is not found in either of the brazzein or thaumatin gene sequences. The digest was done using NEBuffer #2 at an incubation temperature of 37°C. The digested DNA was then run on a 2% Seakem LE agarose gel. The DNA was transferred to a positively charged nylon membrane (GibcoBRL, Grand Island, NY) according to the Stratagene PosiblortTM pressure blotter protocol (Stratagene, La Jolla, CA). After drying overnight the DNA was permanently bound to the membrane by UV-crosslinking in a Strata-link oven (Stratagene, La Jolla, CA). To prepare the membrane for addition of the probe, the membrane was floated, DNA-side up, on top of 6X SSC until the underside of the membrane was completely wet and the membrane was submerged in 6X SSC for 5 min. The membrane was placed in a bag with 1 ml prehybridization solution (PH) per cm² of membrane and the bag was sealed and incubated at 68°C for 15 min.

The DNA template for generating the probe was obtained from the thaumatin DNA sequence used to inject embryos. The region to which it bound was indicated in figure 14. It was cut out of pUC plasmid by a double restriction enzyme digestion with the enzymes *Bam*HI and *Eco*RI, NEBuffer for *Eco*RI, and 1X BSA at 37°C for at least 1 hour. The probe was prepared by random primer labeling as described in objective #1.

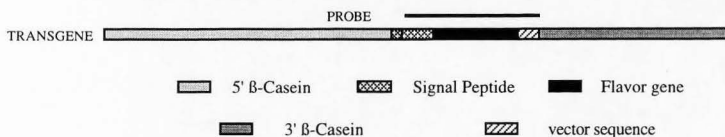


FIG. 14. Location of probe used in detection of transgene. The probe was prepared from the clone that contained the thaumatin gene and signal peptide sequence which was incorporated into the expression vector. The probe was approximately 700 bp in length.

The efficiency of the reaction was determined by measuring the percentage of ^{32}P nucleotides that were incorporated. The volume of the reaction was brought up to 100 μl with TE buffer after 1 μl of EDTA was added to stop the reaction. To measure the “before” radioactive level, 1 μl of the reaction was dotted onto a DE81 filter. The remaining 99 μl of the reaction was run through a Sephadex G50 spin column at 4,000 rpm for 4 min (30) to remove any un-incorporated nucleotides. The elutant was removed and 1 μl dotted onto another DE81 filter to measure the “after” radioactive level. Using the ratio of the “before” and “after” radioactive counts, the efficiency of the reaction was determined in terms of percent incorporation. The percent incorporation was then used to calculate the specific activity. If the probe had a specific activity of 10^8 cmp/ μg , 5 ng of DNA was used per ml prehybridization solution. If the probe had a specific activity of 10^9 cmp/ μg , 2 ng of DNA was used per ml of solution.

To probe the membrane, the PH was drained out of the bag containing the membranes and fresh PH was added along with the probe. This was incubated at 68°C for at least 8 hours. The filters were then washed according to high stringency conditions (3) to remove any unbound probe.

Objective 3: Detection of protein in the milk of the transgenic mice

Generation of antibodies to thaumatin and brazzein

The goat and rabbit antibodies that were used in the assay were generated using the natural thaumatin protein (Sigma Chemical Co., St. Louis, MO) and a multiple antigenic peptide of brazzein was synthesized by USU Biotechnology Laboratory.

A 2 ml volume of 0.25 mg/ml thaumatin was added to one vial of Ribi Adjuvant (RIBI Immunochemicals, Hamilton, MO) that had previously been warmed to 37°C . The mixture was vortexed for a total of four minutes and then kept at 37°C until the injections were given. Two goats were injected intramuscularly (IM) with 1.0 ml. Two rabbits were

injected IM, intradermally and subcutaneously with varying amounts that made a total volume of 0.7 ml. A third rabbit was injected in the same sites but had a total volume of 1 ml. A booster injection was given 16 days later.

A volume of 2 ml of 0.25 mg/ml brazzein multiple antigenic peptide was added to one vial of Ribi adjuvant at 37°C. It was injected into rabbits and goats in the same manner as described above for thaumatin.

The blood was collected from the immunized goats and rabbits 26 days after the first injection was given. The serum was isolated by centrifugation at 1,200 rpm for 15 min. The serum was transferred to a fresh tube and stored at -80°C to preserve the antibodies over a long time period. Aliquots of 100 μ l were prepared and stored at -20°C for use in the ELISA.

ELISA

The ELISA was developed to be able to detect varying levels of either thaumatin or brazzein protein by comparing absorbance levels (O.D., 450nm) of samples that contained unknown concentrations of the protein to the O.D. of known concentrations of protein. Positive and negative controls were incorporated into the assay. Negative controls contained all reagents except for the antigen. Positive controls contained varying known concentrations of antigen.

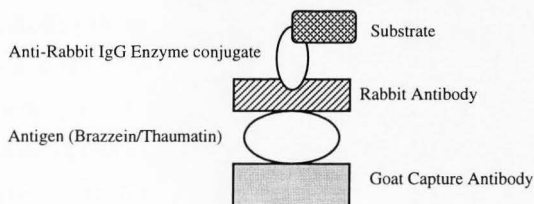


FIG. 15. Design of ELISA. The goat capture antibody was bound to the microtiter plate. After blocking of the plates, thaumatin or brazzein antigen was added. Subsequently the rabbit antibody was incubated followed by the anti-rabbit IgG horseradish peroxidase conjugate and its substrate.

The components of the ELISA were added in the order shown in figure 15. A checker-board design was used for the assay to determine the appropriate levels of antibody to be used. Goat antibody was used at log 10 dilutions between 1×10^{-3} and 1×10^{-6} . Antigen was added at log 10 dilutions from 100 $\mu\text{g}/\text{ml}$ down to 0.01 $\mu\text{g}/\text{ml}$. The dilutions were done both in milk and in phosphate buffered saline (PBS) to see what effect the milk had on the sensitivity of the assay. Rabbit antibody was used at log 10 dilutions from 1×10^{-2} to 1×10^{-6} . The next two steps in the ELISA were added at a constant level in accordance with the manufacturer's instructions.

Dilutions of the goat capture antibody were made in a bicarbonate coating buffer (7.35g NaHCO_3 , 1.36g Na_2CO_3 , 900 ml dH_2O adjust pH to 9.4, dH_2O to 1 L). A volume of 50 μl of the different concentrations of goat antibody were pipetted into the assigned wells of Immulon™ ELISA plates (Intermountain Scientific Corporation, Kaysville, UT). The plates were gently tapped to ensure that the bottom of the well was covered. This was done after each step in the ELISA. The plates were covered and stored at 4°C for a minimum of 12 hours. The wells were emptied by flicking the plates over a sink. The plates were washed by pipeting 200 μl of wash solution (1X PBS) into each well and the wash solution removed by flicking the plates. This was repeated twice. The next step involved blocking the wells so that the antigens added in later steps would not bind to the sides of the wells. A volume of 100 μl of ELISA blocking solution (0.025M PBS, 0.05% Tween 20, 0.5% Gelatin) was put into each well. The plates were covered and incubated for 1 hour at 37°C . The plates were then washed as described above. The antigen was added at the specific concentration in a volume of 50 μl per well. Dilutions of antigen were made in the 1X PBS wash buffer or in bovine milk. The plate was covered and incubated for 1 hour at 37°C . The plates were then washed. A volume of 50 μl of serially diluted rabbit antibody in wash buffer was added to each well. After incubation for 1 hour at 37°C , the rabbit antibody was removed by washing the plates. An anti-rabbit antibody

conjugated to horse radish peroxidase (Sigma Chemical Co., St. Louis, MO) was prepared according to the manufacturer's instructions, stored at 4°C, and kept away from UV light. A 1:100 dilution of the conjugate was made in wash buffer just prior to its addition to the plate. A volume of 50 µl of conjugate was pipetted into each well and the plate was covered and incubated for 1 hour at 37°C. Once again the plates were washed. The substrate for the horse radish peroxidase was ortho-phenylenediamine dihydrochloride (OPD) (Sigma Chemical Co., St. Louis, MO). The OPD solution was prepared according to the package insert that came with the OPD tablets. A volume of 50 µl of the OPD solution was pipetted into each well, and the plates were covered and incubated at room temperature in the dark. The plate was read after 10-15 min on an ELISA plate reader at 450 nm.

Collection of mouse milk

Mouse milk was collected during the lactation period using a method described in Maga et al. (22). The lactating mice were injected with oxytocin (Sigma Chemical Co., St. Louis, MO) and milked by a specifically designed suction device (gift from the original author; James Murray, University of California, Davis). The volume of milk obtained varied from 20 µl to 200 µl. The milk was collected into a microcentrifuge tube and stored at -20°C until used.

Analysis of mouse milk proteins by capillary electrophoresis (CE)

Some of the putative thaumatin transgenic mouse milk samples were analyzed by CE by Marie Strickland, Department of Nutrition and Food Sciences, Utah State University. A positive control was prepared for the CE by adding the native plant thaumatin obtained from Sigma to the milk of non-transgenic mouse milk. The negative control consisted of milk from non-transgenic mice or transgenic mice carrying an unrelated transgene. A range of 20 to 50 µl of milk was mixed with 3 volumes of 9 M urea, 100 mM

phosphoric acid solution and stirred for 30 min. The resulting pH was well below the isoelectric point of all common bovine milk proteins. The 9 M urea disrupted the micelles and dissolved the proteins in milk. The extracts were centrifuged at top speed in an Eppendorf centrifuge. The clear solution below the floating lipid layer was removed, filtered through a 0.2- μ m low protein binding syringe filter, and diluted 5- to 10-fold before CE. The run buffer was 4 M urea, 100 mM phosphate pH 3.35. This pH was below the isoelectric point of most milk proteins so that they are positively charged in solution. The pI of thaumatin was calculated to be greater than 11 using the GCG computer analysis program. Migration towards the negative electrode depended on the charge density of the molecules. The small, highly positive charged species eluted first, followed by larger, less charged species. Ultraviolet light at 210 nm was used for detection of the bands.

MATERIALS

Seaplaque GTG agarose

- from FMC Bioproducts, Rockland, ME
- cat. # 50113

Nuseive GTG agarose

- from FMC Bioproducts, Rockland, ME
- cat. # 50082

Seakem LE agarose

- from FMC Bioproducts, Rockland, ME
- cat # 50004

Ethidium bromide

- from Sigma Chemical Co., St Louis, MO
- cat. # E-8751

Proteinase K

- from Sigma Chemical Co., St Louis, MO.
- cat. # P-2308

Sodium dodecyl sulfate (SDS)

- from Acros Organics, NJ.
- cat. #41953-0010

EDTA

- from Sigma Chemical Co., St Louis, MO.
- cat. #E-5134

Tris

- from Acros Organics, NJ.
- cat. # 42457-5000

Phenol chloroform

- from Sigma Chemical Co., St Louis, MO.
- cat. #P-3803

pGFP-C1

- Clontech, Palo Alto, CA
- cat. #6094-1

Bacto® tryptone

- Difco laboratories (ordered through Fisher Chemical Co.)
- cat. #0123-17-3

Bacto® yeast extract

- Difco laboratories (ordered through Fisher Chemical Co.)
- cat. #0886-17-0

Sodium Chloride

- from Mallinckrodt, Paris, KY
- cat. #7581

Bacto-agar

- from Becton Dickinson and Company, Cockeysville, MD.
- cat. # 4311849

Kanamycin

- from Sigma Chemical Co., St Louis, MO
- cat. # K-4000

ATP

- from Sigma Chemical Co., St Louis, MO.
- cat. # A-2383

dNTP's

- from Pharmacia, Piscataway, NJ.
- cat. #27-2035-01

BSA

- from Pentex, Kankakee, IL.
- cat. #81-001-2

Ampicillin

- from Sigma Chemical Co., St Louis, MO.
- cat # A 9518

IPTG

- from Sigma Chemical Co., St Louis, MO.
- cat # I 5502

X-gal

- from Sigma Chemical Co., St Louis, MO.
- cat # B 4252

1 kb ladder

- Gibco BRL, Grand Island, NY.
- cat. # 5615 SA

Lambda Hind III

- from New England BioLabs, Beverly, MA
- cat. # 301-2S

XL1-Blue Epicurian Coli Electroporation-Competant Cells

- from Stratagene, La Jolla, CA.
- cat. # 200236

Nitrocellulose membrane

- from Schlieher & Schuell, Keene, NH
- cat. #20470

³²P dCTP

- ICN Biochemicals, Costa Mesa, CA
- cat. #39011X

Sodium citrate

- from Sigma Chemical Co., St Louis, MO
- cat. #C-9795

Thaumatococin

- from Sigma Chemical Co., St Louis, MO
- cat. # T-7638

Nylon Membrane

- from Gibco BRL, Grand Island, NY
- cat. # 14856-017

Restriction enzymes:**Aat II**

- from New England BioLabs, Beverly, MA
- cat. # 117S

Age I

- from New England BioLabs, Beverly, MA
- cat. # 552L

Bam HI

- from New England BioLabs, Beverly, MA
- cat. #136L

Dde I

- from New England BioLabs, Beverly, MA
- cat. #175S

Eco RI

- from New England BioLabs, Beverly, MA
- cat. #101L

Hind III

- from New England BioLabs, Beverly, MA
- cat. #104S

Kpn I

- from New England BioLabs, Beverly, MA
- cat. #142L

Pvu II

- from New England BioLabs, Beverly, MA
- cat. #506S

Sal I

- from New England BioLabs, Beverly, MA
- cat. #138S

Sma I

- from New England BioLabs, Beverly, MA
- cat. #141S

Spe I

- from New England BioLabs, Beverly, MA
- cat. #133S

Enzymes:**Taq polymerase**

- Gibco BRL, Grand Island, NY.
- 8038SA

T4 DNA polymerase

- from New England BioLabs, Beverly, MA
- cat. #203S

Polynucleotide kinase

- from New England BioLabs, Beverly, MA
- cat. #210S

T4 DNA ligase

- from New England BioLabs, Beverly, MA
- cat. #202S

Beta-agarase

- from New England BioLabs, Beverly, MA
- cat. #392L

Sequenase

- from USB, Cleveland, OH
- cat. #70722

RESULTS

Objective 1: Construction of expression vector

This results section gives the data obtained from the experiments described in the experimental methods section. The results are given under the same headings and subheadings that were used in the methods section so as to facilitate comparison between the two sections.

Reaction (rxn) and gel numbers, used in the laboratory to trace experimental data, have been noted for some of the important reactions to provide continuity between the thesis and the raw data.

In all of the agarose or polyacrylamide gel electrophoresis shown, one of two standard size markers was used: λ Hind III size marker (New England Biolabs, Beverly, MA) or 1 kb ladder (Gibco BRL, Grand Island, NY). The sizes for the bands seen with these are shown below (figure 16) and should be referenced to as needed.

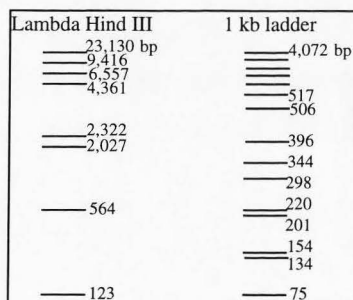


FIG. 16. Marker size map.

Cloning β -casein regions into pGFP-C1

Isolating β -casein fragments

The 4.4 kb 5' β -casein region obtained by PCR was fractionated on a 1% LE agarose gel. The diagnostic 4.4 kb band was demarcated by an arrow on the gel (figure 17). The 2.2 kb 3' β -casein region obtained by PCR was fractionated on a 1% LE agarose gel (figure 18). The diagnostic 4.4 kb band was demarcated by an arrow on the gel. Only a small proportion of each PCR reaction was fractionated by gel electrophoresis for analytical purposes; the remainder of the 5' β -casein PCR product was ligated into pGFP-C1 as described below.

Cloning of 5' β -casein DNA into pGFP-C1

DNA isolated from the colonies resulting from the ligation of 5' β -casein with pGFP-C1 was fractionated on a 1% LE agarose gel without restriction enzyme digestion.

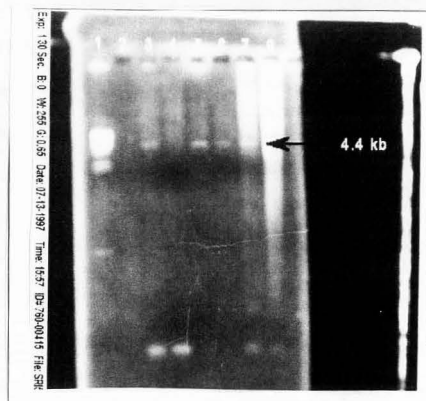


FIG. 17. 5' β -casein DNA region obtained by PCR.

Lane 1: λ Hind III size marker (New England Biolabs, Beverly, MA);

Lane 2: PCR reagent negative control (no DNA template added);

Lanes 3-8: 5' β -casein PCR product.

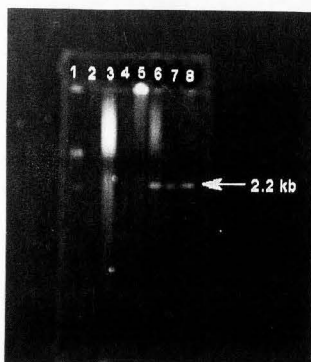


FIG. 18. 3' β -casein DNA region obtained by PCR.

Lane 1: λ Hind III size marker;

Lanes 6-8: 3' β -casein PCR product.

As a comparison, the uncut pGFP-C1 plasmid was also loaded onto the gel so that the control band could be visualized (figure 19). Those clones that had different size bands than the plasmid band were considered possible candidates for the correct clone because the different migration patterns indicated that the size of the plasmids had changed due to ligation of inserts. Because the plasmids were not linearized it was not possible to predict the size of bands seen on the gel. However, the plasmids that had a different electrophoresis pattern than the control were screened by restriction enzyme digestion.

In figure 19 below, and the following figures that show the cloning of 5' and 3' β -casein DNA, the only clone that was ultimately used in the expression vector is indicated by an arrow. Those clones in figure 19 that had an uncut lower band of approximately 6 kb were possible candidates that contained the 5' β -casein.

Examples of such bands in figure 19 are in lanes 4 and 14. These were cut with the restriction enzymes AatII and AgeI in a double digestion reaction which, in a correct clone,



FIG. 19. Screening of 5'β-casein into pGFP-C1 (gel #763).

Lanes 1, 15 and 16: λ Hind III size marker;

Lanes 2 and 17: uncut pGFP-C1;

Lanes 3 - 14 and 18 - 24: uncut putative p5'βcas clones (Rxn # 3531).

would liberate the 5'β-casein from the pGFP resulting in two bands of similar size, a 4.4 kb and 4.7 kb band. A second diagnostic digestion with SpeI was done which would cut the plasmid in two sites generating a 7.3 and 1.8 kb band. These diagnostic enzyme digestions were fractionated on a 1% LE agarose gel (figure 20).

The digest with AatII and AgeI of the clone, seen in lane 3 of figure 20, appeared to contain one thick band at approximately 4.5 kb. This was interpreted to be a result of loading too high of a DNA concentration onto the gel so that it was not possible to differentiate two bands. The digest with SpeI liberated the two predicted bands and a third band which was probably DNA that was not digested (lane 6). The clone shown by the arrows in figures 19 and 20 was then sequenced by the USU Biotechnology Laboratory. It was sequenced across the signal peptide sequence as well as from the 3' end of the 5'β-

casein region. The sequencing results were compared to the published sequence of the bovine β -casein gene using the Wisconsin GCG Sequence Analysis Software Package. The sequence which originated from the 3' end of the 5' β -casein region is shown in figure 21. Part of the restriction enzyme site ACCGGT that is recognized by AgeI was highlighted to orientate the data in relation to how it fits into the expression vector (figure 7). The errors seen at either end of the sequences that are shown throughout this chapter maybe as a result of sequencing artifacts and therefore could not be considered true errors in the sequence. Errors at the beginning of sequencing and at the terminal end when sequencing beyond 300-400 bases are common artifacts of sequencing.

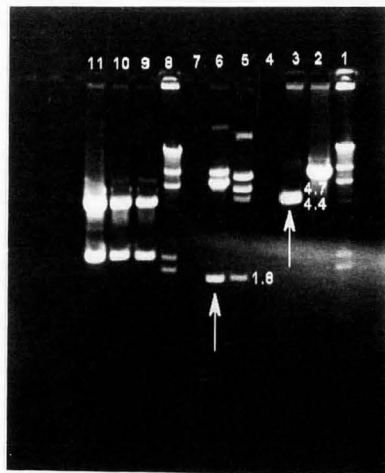


FIG. 20. Screening of putative p5' β cas clones (gel# 765).

Lane 1: λ Hind III;

Lanes 2 & 3: putative clones, from lane 4 & 14 in figure 18, cut with AatII and AgeI;

Lanes 5 & 6: putative clones, from lane 4 & 14 in figure 18, cut with SpeI.

494 AGGACCAAAAGTGAGTNATATGGCATAAAATACAATNGCATGCATATACAA 445
|||:|||||:|||||
4047 AGAACAAAAAGTGAGTAAATATTGCATAAAATACAATTGCATGCATATACAA 4096
444 TCTAGATAAAATATATCTTATTCCAGTGATGAAATATTTGTATCCCTTACT 395
4097 TCTAGATAAAATATATCTTATTCCAGTGATGAAATATTTGTATCCCTTACT 4146
394 GTAGAGTGCTAGGTTTtagctgtgtctattTCAACACAGGATGATACTCCAG 345
4147 GTAGAGTGCTAGGTTTtagctgtgtctattTCAACACAGGATGATACTCCAG 4196
344 AGGATGGTATATCAGACAACAATAAATAATATGTTTCATAATTATAATAAA 295
4197 AGGATGGTATATCAGACAACAATAAATAATATGTTTCATAATTATAATAAA 4246
294 AAGTGTTCAGTAAAAAATTAATAAATCTCTTTTCTGTTACCCATAAAAAAC 245
4247 AAGTGTTCAGTAAAAAATTAATAAATCTCTTTTCTGTTACCCATAAAAAAC 4296
244 TCTTCATTAAAGTAAACAAAAATATACTAATGAAAGTTACTAAATTAA 195
4297 TCTTCATTAAAGTAAACAAAAATATACTAATGAAAGTTACTAAATTAA 4346
194 AAGACTCTCAAAAGACATATAACATTTTTATTTTTCTAGATTGTGAAATA 145
4347 AAGACTCTCAAAAGACATATAACATTTTTATTTTTCTAGATTGTGAAATA 4396
144 GATAGCTCTGAATAAAGCAAGTAAAAATTAGGTAGGAAAAATATTTAATAA 95
4397 GATAGCTCTGAATAAAGCAAGTAAAAATTAGGTAGGAAAAATATTTAATAA 4446
94 TGAGTTGACTGTGGGAACATAAGTGTTTTTTTCTCTTTAGCTGGAAGA 45
4447 TGAGTTGACTGTGGGAACATAAGTGTTTTTTTCTCTTTAGCTGGAAGA 4496

Age I

44 ACTCAATGTACCGG 31
|||||
4497 ACTCAATGTACCTG 4510

FIG. 21. Sequence comparison between the cloned and published sequences of 5' portion of β -casein. The top line starting at position 494 is the sequence obtained from our DNA sequence cloned into pGFP-C1 and the bottom line is the sequence of the bovine β -casein gene (accession #x14711).

Cloning 3' β -casein DNA into pUC-19

A schematic of the cloning of 3' β -casein into pUC-19 is shown in figure 22. It illustrates the possible orientation in which the 3' β -casein insert could have ligated into the pUC plasmid. Putative clones with the DNA insert of 3' β -casein region cloned into pUC-19 in either orientation were identified as shown on the following 1% LE agarose gels (figures 23-24). Figure 23 contained the undigested putative 3' β -casein-pUC clones. As explained in section 2 for figure 19, the uncut clones that had bands with a different migration pattern when compared to the uncut pUC19 were considered candidates for containing the 3' β -casein insert. These putative clones were then definitively diagnosed using restriction enzyme digestion.

Those samples in figure 23 that had a lower band of approximately 3 or 4 kb, as connoted by an arrow, (for example lanes 14 and 16) were subjected to restriction enzyme analysis to confirm that they contained the 3' β -casein. The size of the bands varied

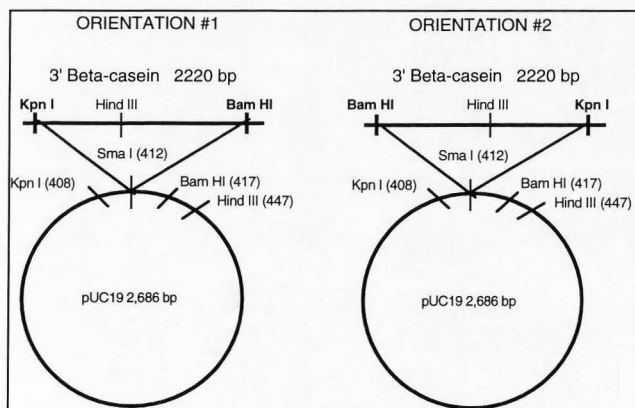


FIG. 22. Two possible orientations of cloning 3' β -casein DNA sequences into pUC-19.

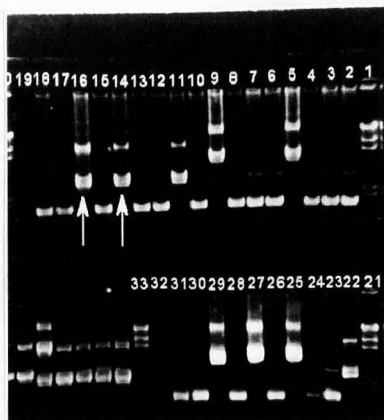


FIG. 23. Screening of putative 3'β-casein-pUC19 (gel #752).
Lanes 1, 20, 21 & 33: λ Hind III marker;
Lanes 2-18 & 22-31: uncut putative 3'β-casein-pUC clones (Rxn# 3505);
Lanes 19 & 32: uncut pUC.

according to the orientation in which the 3'β-casein had ligated into the pUC-SmaI enzyme site (figure 22). In the first enzyme digestion done with KpnI, two bands of 2.7 and 2.2 kb for orientation 1 (figure 24) or 4.8 kb for the second orientation were seen (figure 24). The second enzyme digestion was a HindIII digestion that would liberate similar size bands for both orientations.

Several of the clones, shown by arrows in figure 24, that were likely candidates for the correct insert were sequenced by the USU Biotechnology Laboratory. The cloned sequences were aligned with the published sequences by the Bestfit program in the GCG Sequence analysis package. The sequence of the clone used in the expression vector was matched to the native sequence of the β-casein gene as shown in figure 25. The errors seen at either end of the sequence are probably sequencing artifacts and cannot be

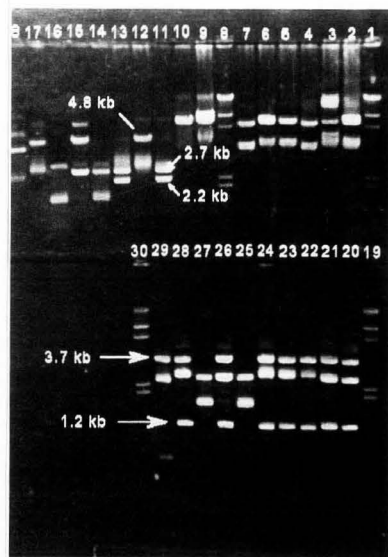


FIG. 24. Screening putative 3'β-casein-pUC clones (gel #754).
Lanes 1,8,19 & 30: λ Hind III size marker;
Lanes 9-18: putative 3'β-casein-pUC clones cut with Kpn I;
Lanes 20-29: putative 3'β-casein-pUC clones cut with Hind III.

considered true errors in the sequence. Errors at the beginning of sequencing and at the terminal end when sequencing beyond 300-400 bases are common artifacts of sequencing.

Subcloning 3'β-casein DNA into p5'β-cas construct

After ligating the 3'β-casein DNA into p5'β-cas, the initial screening was done by colony hybridization using a radiolabeled probe specific for the 2.2 kb 3'β-casein DNA segment. Colonies of interest that showed up as dense areas on the autoradiograph were

```

41  CCAGGATGGGTACCAGGATAAAATCCACCCCTTTGCCCAGACGCAGTCTC 90
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
8024 CCAGGATGAATCCAGGATAAAATCCACCCCTTTGCCCAGACGCAGTCTC 8073

91  TAGTCTATCCCTTCCCTGGGCCCATCCATAACAGCCTCCACAAAACATC 140
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
8074 TAGTCTATCCCTTCCCTGGGCCCATCCATAACAGCCTCCACAAAACATC 8123

141 CCTCCTCTTACTCAAACCCCTGTGGTGGTGCCGCCTTTCCTTCAGCCTGA 190
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
8124 CCTCCTCTTACTCAAACCCCTGTGGTGGTGCCGCCTTTCCTTCAGCCTGA 8173

191 AGTAATGGGAGTCTCCAAAGTGAAGGAGGCTATGGCTCCTAAGCACAAAG 240
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
8174 AGTAATGGGAGTCTCCAAAGTGAAGGAGGCTATGGCTCCTAAGCACAAAG 8223

241 AAATGCCCTTCCCTAAATATCCAGTTGAGCCCTTTACTGAAAGCCAGAGC 290
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
8224 AAATGCCCTTCCCTAAATATCCAGTTGAGCCCTTTACTGAAAGCCAGAGC 8273

291 CTGACTCTCACTGATGTTGAAAACTGCACCTTCCTCTGCCTCTGCTCCA 340
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
8274 CTGACTCTCACTGATGTTGAAAACTGCACCTTCCTCTGCCTCTGCTCCA 8323

341 GTCTTGGATGCACCAGCCTCACCAGCCTCTTCCTCCAACGTGCATGTTTC 390
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
8324 GTCTTGGATGCACCAGCCTCACCAGCCTCTTCCTCCAACGTGCATGTTTC 8373

391 CTCCTCAGTCCGTGCAGTCCCTTTCTCAATCCAAAGTCTGCCTGTTCCC 440
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
8374 CTCCTCAGTCCGTGCTGTCCCTTTCTCAGTCCAAAGTCTGCCTGTTCCC 8423

441 CAAAAGCAATGGCCTATCCCAGAAAAATATG. CCATTGAGGCTTTCT 489
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
8424 CAGAAAGCAGTGCCCTATCCCAGAGAGATGCCCATTCAGGCTTTCT 8473

490 GCTGTA. CAAGAACCTGTA. TCNGTCCTGTCCGGGAAC 525
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
8474 GCTGTACCAGGAGCCTGTACTCGGTCTGTCCGGGAC 8511

```

FIG. 25. Sequence analysis of 3' portion of β -casein cloned into pUC. The top line is the sequence obtained from DNA cloned into pUC19. The bottom line is the sequence of the bovine β -casein gene (accession #X14711).

Subcloning 3' β -casein DNA into p5' β -cas construct

After ligating the 3' β -casein DNA into p5' β -cas, the initial screening was done by colony hybridization using a radiolabeled probe specific for the 2.2 kb 3' β -casein DNA segment. Colonies of interest that showed up as dense areas on the autoradiograph were matched to the original agar plate that contained the bacterial colonies. The plasmid DNA was isolated and further screened by restriction enzyme analysis to identify p5'3' β -cas clones.

The samples were screened by several different diagnostic enzyme digestion reactions. The reactions were HindIII, a double digest using BamHI and KpnI, and Sall and XbaI.



FIG. 26. Autoradiograph of colony hybridization to detect p5'3' β cas.

Figure 27 shows the HindIII digested-plasmid DNAs of all the colonies that were isolated from the plates. For each clone, the uncut samples were run next to the corresponding samples that were cut with HindIII (see legend for figure 27 for identification of the samples). This aided in accounting for different size bands that were visible on the agarose gel due to incomplete restriction enzyme digestion.

The DNA samples tentatively identified in figure 27 were cut by several enzymes as indicated above (figure 28). All of the predicted sized bands correlated with the actual sizes, except for the plasmids cut with XbaI. Lanes 2-5 all exhibit the predicted bands for a correct clone that was digested with BamHI and KpnI. In lanes 7-10, the upper band

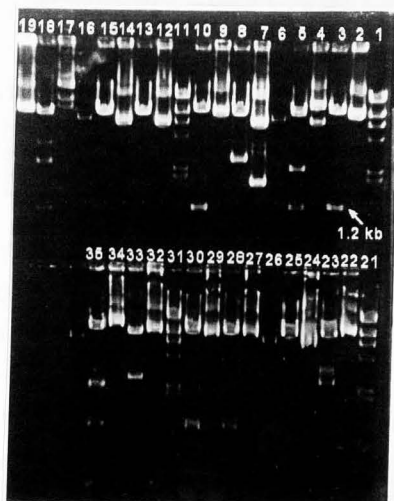


FIG. 27. Screening of p3'5'β-cas #1 (gel# 945).
Lanes 1 & 21: λ HindIII size marker;
Lanes 6, 16, 26 & 36: pGFP5'β-casein cut with KpnI;
Lanes 2-5, 7-10, 12-15, 17-20, 22-25, 27-30 & 12-15: alternating uncut sample next to sample cut with HindIII.

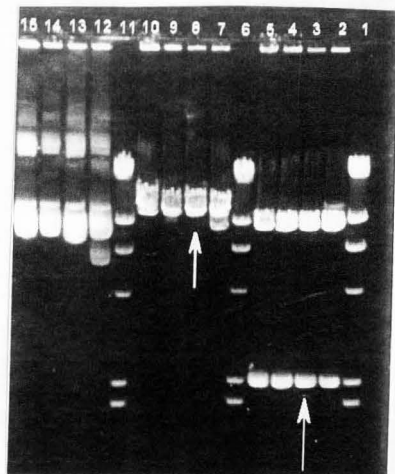


FIG. 28. Screening of putative p3'5'β-cas #2 (gel #946).

Lanes 1,6 & 11: λ Hind III size marker;

Lanes 2-5: putative 3'5'β-casein clones cut with BamHI and KpnI;

Lanes 7-10: putative 3'5'β-casein clones cut with Sall;

Lanes 12-15: putative 3'5'β-casein clones cut with Xba I.

has shifted upwards slightly when compared to the upper band in lanes 2-5. This upper band at approximately 11 kb was diagnostic of a correct clone. It was determined on the basis of this and other restriction enzyme digestions as well as from sequencing data (not shown) that our 5'β-casein region did not have all the XbaI restriction enzyme sites that were expected based on the published sequence. Lanes 3 and 8 contain the same plasmid that was used for the expression vector.

Generating separate segments of the synthetic genes

Amplification using PCR

As outlined in figures 8 and 9, the thaumatin and brazzein genes respectively were amplified by PCR. In the initial attempt to amplify thaumatin and brazzein DNA using

PCR, the reactions used the buffer supplied with the polymerase. The MgCl_2 concentrations were varied. This resulted in multiple bands as could be seen on the ethidium bromide stained 2.5 % Seaplugue GTG agarose gel in figure 29.

Optimization buffers were used to optimize the PCR reaction so that some of the extraneous bands would be eliminated. The buffers differed in pH and MgCl_2 concentration. The PCR cycling conditions and other reaction conditions such as primer concentrations were not altered from the previous reaction that was run in figure 29. The results from this PCR are shown in figure 30.

Figure 30 shows the results for brazzein and three of the thaumatin fragments. The remaining thaumatin D fragment was analyzed on separate gel (figure 31). The PCR



FIG. 29. Synthetic gene amplification (gel # 843).
Lanes 1, 18 & 19: 1 kb ladder size marker (Gibco BRL, Grand Island, NY);
Lanes 3-5 & 8-11: brazzein PCR with 1 mM, 2 mM or 3 mM MgCl_2 added;
Lanes 13-15: thaumatin A PCR with 1 mM, 2 mM or 3 mM MgCl_2 added;
Lanes 21-23: thaumatin C PCR with 1 mM, 2 mM or 3 mM MgCl_2 added;
Lanes 26-28: thaumatin B PCR with 1 mM, 2 mM or 3 mM MgCl_2 added;
Lanes 6, 11, 16, 24 & 29: negative control for each reaction.

product, using thaumatin D as the template, was fractionated on a 2.5% Seaplugue agarose gel (figure 31).

In figures 30 and 31, the predominant PCR bands of the brazzein, thaumatin A, B, and D reactions, were the bands of the predicted size. In some of the lanes, additional bands could be seen, but they were of low abundance. This was adequate to use for cloning the PCR products into pUC. The PCR product from this reaction for thaumatin C, however, could not be used due to the almost nonexistent band. The thaumatin C PCR product seen in figure 29, lane 23 was used for cloning the DNA into pUC because of its abundance.

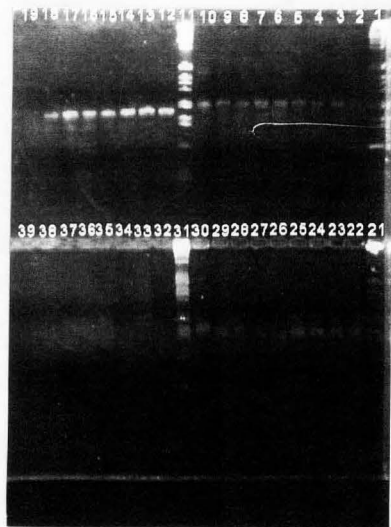


FIG. 30. Synthetic gene amplification using optimization buffers #1 (gel #846).
Lanes 1, 11, 21 & 31: 1 kb size marker;
Lanes 2-10: brazzein PCR using optimization buffers A, C, G, H, I, K, L & P;
Lanes 12-20: thaumatin A PCR using optimization buffers A, C, G, H, I, K, L & P;
Lanes 22-30: thaumatin B PCR using optimization buffers A, C, G, H, I, K, L & P;
Lanes 32-40: thaumatin C PCR using optimization buffers A, C, G, H, I, K, L & P.

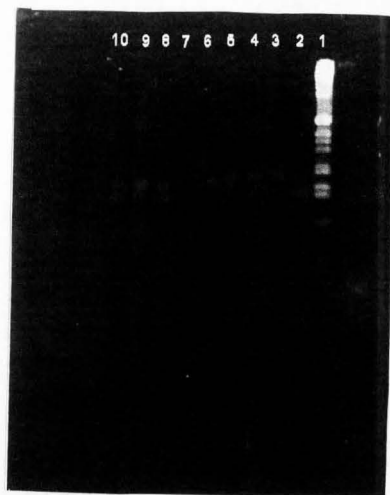


FIG. 31. Synthetic gene amplification using optimization buffers #2 (gel #847).

Lane 1: 1 kb size marker;

Lane 2-10: thaumatin D PCR using optimization buffers A, C, G, H, I, K, L & P.

Ligation and screening of ligation products

White colonies carrying clones with DNA inserts identified on the basis of lac Z (blue-white) selection were picked from the plates, and the DNA was isolated and subsequently diagnosed by digestion with BamHI and EcoRI. Each of these enzyme sites (BamHI and EcoRI) was located either to side of the SmaI site in pUC into which the fragments were cloned and not in the fragment itself.

The lanes that contained diagnostic lower bands ranging from 160 bp to 230 bp, depending on which thaumatin fragment was used, were considered putative positive clones: thaumatin A: lane 4, thaumatin B: lanes 8 & 10, thaumatin C: lanes 16 & 17,

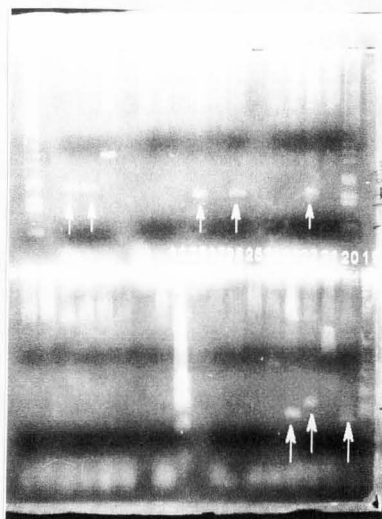


FIG. 32. Screening of thaumatin-pUC (gel # 841).
Lanes 1,8,19 & 29: 1 kb size marker;
Lanes 2-5: putative thaumatin A cloned into pUC cut with BamHI and EcoRI;
Lanes 7-12: putative thaumatin B cloned into pUC cut with BamHI and EcoRI;
Lanes 14-16: putative thaumatin C cloned into pUC cut with BamHI and EcoRI;
Lanes 20-27: putative thaumatin D cloned into pUC cut with BamHI and EcoRI.

thaumatin D: lanes 24 & 25 (figure 32). Figure 33 showed brazzein DNA insert cut out of the pUC plasmid with PpuMI and SalI, which were the enzymes sites on either side of the brazzein synthetic gene. The gel was an 8% polyacrylamide gel, which provided more precision for estimating the relative sizes of the bands. A brazzein diagnostic band at approximately 200 bp was apparent in lanes 6 through 8 of figure 33.

The modified protocol for cloning these oligonucleotides was important, because a high number of random errors were seen when using the unmodified protocol. An example of a sequence containing multiple errors is shown below in figure 34. A high frequency of random errors were seen in DNAs generated from the original, unmodified

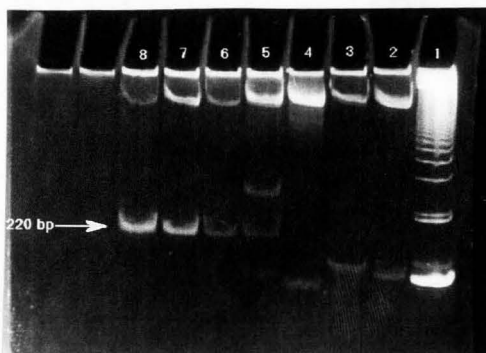


FIG. 33. Brazzein digested out of pUC (gel # 805).
Lane 1: 123 bp ladder (Gibco BRL, Grand Island, NY);
Lanes 5-10: putative brazzein-pUC clones.

nucleic acid synthesis procedure. The errors are spuriously located throughout the sequence in figure 34, without any predictable pattern. Therefore, some of the brazzein and thaumatin fragments had to be remade using a modified procedure with a longer deprotection time (overnight instead of several hours) and lower temperature (60° C instead of 80° C). This modified procedure yielded far fewer errors in the oligonucleotide sequences. The number of errors for each oligonucleotide varied. Therefore, instead of remaking the oligonucleotides for the thaumatin fragments A and D, a high number of putative clones were sequenced until a complete error-free sequence was found for both A and D. The following series of figures (34-38) shows the sequences of the synthetic DNA with the predicted sequences. The correct error-free sequences showing the correct brazzein and thaumatin clones were given after the sequence that contained random errors (figure 33). In each figure the sequence obtained from the synthetic clone is on the upper row. The synthetic sequence (top row) was compared to the predicted sequence (lower row) that was used in the design of the expression vector (figures 5 and 6).

Figure 34 shows a number of errors that occurred in the PCR-cloned thaumatin C fragment. In other clones from the PCR reaction, different errors were seen. With the modifications in the deprotection protocol given above, the number of errors was significantly reduced. A varying number of clones for brazzein and each thaumatin fragment were sequenced until one was found which had a perfect sequence when compared to the predicted sequence.

The sequence of the brazzein cloned into pUC was shown on the top line of each row in figure 35. Each nucleotide of the clone matched with the predicted sequence for this

```

3  ANATCTGGGNCAGAAC..ACTGCTATTTTGACGATAGAGAAAGAGGCATC 50
   |:|||||:||||| ||||||||||||||||| | |||||||||
193 agatctgggccagaaccgactgctattttgacgatagcggaagaggcatc 242
   .
51  AGCCGGACCGCGGACTGTGGCGGACTGTTGCAGTGGAAAAGATTGGCA 100
   ||||||||| ||||||||| ||||||| | |||||||||
243 tgccggaccggc.gactgtggcgactgctgcagtgcaagagatttgca 291
   .
101 GACCTCCTACGACTCTGGCTGAATT 125
     ||||||||| ||||||||| ||
292 gacctcctaccactctggctgagtt 316

```

FIG. 34. Sequence analysis of a thaumatin C clone that contained multiple errors. This clone is an example of one that was prepared using the unmodified protocol.

```

3  CTGCCTGGTGGCCGTGGCCCTGGCCGACAAGTGCAAGAAAGGTGTATGAGA 52
   |||||||||||||||||||||||||||||||||||
23  ctgcctggtggcctgtggccctggccgacaagtgcaagaaggtgtatgaga 72
   .
53  ACTATCCTGTGAGCAAGTGCCAGCTGGCCAACCACTGCAACTATGACTGC 102
   |||||||||||||||||||||||||||||||||||
73  actatcctgtgagcaagtgccagctggccaaccagtgcaactatgACTGC 122
   .
103 AAGCTGGACAAGCAGCCAGAAGCGGCGAGTGCTTTTATGACGAGAAGAG 152
   |||||||||||||||||||||||||||||||||||
123 AAGCTGGACAAGCAGCCAGAAGCGGCGAGTGCTTTTATGACGAGAAGAG 172
   .
153 AAACCTGCAGTGCATCGGCGACTATTGCGGCTAGTAAGTCGACGCGCGC 201
   |||||||||||||||||||||||||||||||||||
173 AAACCTGCAGTGCATCGGCGACTATTGCGGCTAGTAAGTCGACGCGCGC 221

```

FIG. 35. DNA sequence of synthetic brazzein (rxn: 3786#4, upper row) as compared to the predicted sequence (lower row).

for this particular clone, which was used to construct the vector.

The sequence of the thaumatin segment that was cloned into pUC is shown on the top line of each row for each of the thaumatin fragments (figures 36-39). All of the nucleotides matched with thaumatin A DNA (figure 35), except for 167 position. Since

```

3  CTGCCTGGTGGCCGTCGCCCTGGCCGCCACCTTTGAGATCGTGAACAGAT 52
   ||||||||||||||||||||||||||||||||||||||||||||||||
23 ctgcctggtggccgtcgccctggccgccacctttgagatcgtgaacagat 72
   .
53 GCAGCTATACCGTGTGGGCTGCCGCTAGCAAAGGCGATGCCGCCCTGGAT 102
   ||||||||||||||||||||||||||||||||||||||||||||||||
73 gcagctataccgtgtgggctgccgctagcaaaggcgatgccgccctggat 122
   .
103 GCCGGAGGCAGACAGCTGAACTCCGGCGAGAGCTGGACCATCAACGTGGA 152
   ||||||||||||||||||||||||||||||||||||||||||||||||
123 gccggaggcagacagctgaactccggcgagagctggaccatcaacgtgga 172
   .
153 GCCCGGTACCAAGGCCG 169
   ||||||||||||||||
173 gcccggtaccaagggcg 189

```

FIG. 36. DNA sequence of correct thaumatin A (rxn: 3787 # 12, upper row) as compared to the predicted sequence (lower row).

```

41 GAGCCCGGTACCAAGGGCGGCAAGATCTGGGCCAGAACCGACTGCTATTT 90
   ||||||||||||||||||||||||||||||||||||||||||||||||
171 gagcccggtaccaagggcggcaagatctgggccagaaccgactgctattt 220
   .
91 TGACGATAGCGGAAGAGGCATCTGCCGGACCGGCGACTGTGGCGGACTGC 140
   ||||||||||||||||||||||||||||||||||||||||||||||||
221 tgacgatagcggaagaggcatctgccggaccggcgactgtggcggactgc 270
   .
141 TGCAGTGCAAGAGATTGGCAGACCTCCTACCACTCTGGCTGAGTTTAGC 190
   ||||||||||||||||||||||||||||||||||||||||||||||||
271 tgcagtgcagagatttggcagacctcctaccactctggctgagttagc 320
   .
191 CTG 193
   |||
321 ctg 323

```

FIG 37. DNA sequence of correct thaumatin B (rxn: 4059 #8 upper row) as compared to the predicted sequence (lower row).

```

37  TCCCCCTCTGGCTGAGTTTAGCCTGAACCAAGTATGGCAAGGACTATATCG 86
   |||
299  taccactctggctgagtttagcctgaaccagtatggcaaggactatatcg 348
   |||
87  ATATCAGCAACATCAAGGGCTTTAATGTGCCTATGGACTTTAGCCCTACT 136
   |||
349  atatcagcaacatcaagggtttaatgtgcctatggactttagccctact 398
   |||
137  ACCAGAGGCTGCAGAGGCGTGAGATGCGCTGCCGACATCGTCGCCCAGTG 186
   |||
399  accagaggtgcagagggcgtgagatgcgctgccgacatcgtcgcccagtg 448
   |||
187  CCCTGCCAAGCTGAAGGCCCTGGAGGCGGCTGCAATGACGCCTGCACCG 236
   |||
449  ccctgccaaagctgaaggccctggaggcggtgcaatgacgcctgcaccg 498
   |||
237  TGTTCAGACTAGTGAG 253
   |||
499  tgtttcagactagttag 515

```

FIG. 38. DNA sequence of correct thaumatin C (rxn: 4060 #19 upper row) as compared to the predicted sequence (lower row).

this nucleotide was not used in the final thaumatin sequence, this error was allowable and the DNA was used for DNA construction. The sequences in figures 37 and 38 for thaumatin B and C sequences, respectively, matched perfectly with the predicted sequences. In order to verify the identity of a correct clone of thaumatin D, it was necessary to sequence in both directions, to confirm that the errors seen at the ends of the sequence were actually artifacts from the sequencing. By comparing the results of sequence a) and b) in figure 39, it was possible to confirm that reaction 3989 #26 had the correct sequence with no true errors.

Combining thaumatin segments

Joining segments A&B and C&D

Due to the presence of both a KpnI and SpeI enzyme recognition sequences in the multiple cloning site (MCS) of pUC19, each pUC vector that contained an insert was cut in two places when it was digested with either KpnI or SpeI. As a result there were a number of possible ligation products from the ligation of A & B and C & D. Many of them

a) sequencing with M13 forward primer

```

173 CCCCCTGTTTCAGACTAGTGAGTATTGCTGCACCACTGGCAAGTGC GGCC 124
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
494 caccgtgtttcagactagtgtagttattgctgcaccactggcaagtgcggcc 543

123 CTACCGAGTATAGCAGATTTTTTAAGAGACTGTGCCCTGACGCCTTTAGC 74
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
544 ctaccgagtatagcagattttttaagagactgtgccctgacgccttttagc 593

73 TATGTGCTGGACAAGCCTACCACCGTGACCTGCCCTGGCAGCAGCAACTA 24
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
594 tatgtgctggacaagcctaccaccgtgacctgacctggcctggcagcagcaacta 643

23 TAGAGTGACCTTTTGCCGTACC 2
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
644 tagagtgtaccttttgcctacc 665

```

b) sequencing with the M13 reverse primer

```

1 TTTGTGCACCACTGGCAAGTGC GGCCCTACCGAGTATAGCAGATTTTTTA 50
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
518 ttgtctgcaccactggcaagtgcggccctaccgagtatagcagatttttta 567

51 AGAGACTGTGCCCTGACGCCTTTAGCTATGTGCTGGACAAGCCTACCACC 100
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
568 agagactgtgccctgacgcctttagctatgtgctggacaagcctaccacc 617

101 GTGACCTGCCCTGGCAGCAGCAACTATAGAGTGACCTTTTGCCCTACCGC 150
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
618 gtgacctgccctggcagcagcaactatagagtaccttttgcctaccgc 667

151 CTAGTAGGTCGACGCTGCG 169
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
668 ctagttaggtcgacgctgcg 686

```

FIG. 39. Sequencing of the correct thaumatin D (rxn: 3989#26).

involved the re-ligation of pUC plasmid with itself as well as with the small regions cut away from pUC multiple cloning site and the insert which would have resulted in ligation products that did not contain any of the thaumatin fragments. The three possible ligation events involving both thaumatin fragments are depicted in a linear fashion in figure 40, although they could have been linear or circularized depending on the number of ligation events that had occurred. The various pUC plasmids that contained the four different

thaumatin fragments were referred to as pUC 1 - 4 throughout figure 40 to distinguish them from each another.

Because of the many possibilities for ligation, multiple PCR bands were predicted. Therefore, the PCR product was fractionated on a non-denaturing 10% polyacrylamide gel (PAGE) using a large sequencing gel apparatus. The apparatus was used to permit good separation of the similar-sized bands. A large-volume PCR reaction was performed so that enough DNA could be eluted from the PAGE to perform the next ligation. The ethidium bromide-stained gel of the PCR reaction was photographed and is shown in figure 41.

Predominant bands appeared approximately at the predicted size of 323 bp for thaumatin AB (lanes 3 through 7). In lanes 10 through 15 a strong band at 383 was seen for thaumatin CD.

Combining thaumatin AB and thaumatin CD

The ligation products of AB with CD were amplified by PCR, ligated into pUC, and screened by double digestion with BamHI and EcoRI which, in a correct clone, would liberate a band of approximately 680 bp. Those clones that had a diagnostic 680 bp band were identified by arrows in figure 42. Several of these clones were sequenced by the USU Biotechnology Laboratory.

Inserting flavor/sweetener genes into p5'3'βcas construct

Screening ligation of the thaumatin gene into p5'3'βcas

This section describes the ligation of thaumatin DNA sequence into the expression vector, p5'3'βcas. The brazzein DNA sequence was ligated into p5'βcas and then the 3'βcasein was ligated into the brazzein p5'βcas. For the sake of brevity, the procedure for doing this was not described because there was no significant difference in the protocol other than the order in which DNA fragments were ligated into the plasmid backbone.

Putative thaumatin-p5'3'βcas clones were selected on the basis of the colony hybridization autoradiograph (not shown). These clones were diagnosed by a double digestion with PpumI and SalI, which liberated the thaumatin sequence (680 bp) and approximately 1.5 kb of the 3'β-casein sequence, thereby leaving behind the remaining 8.4 kb of the expression vector.

Those clones that were diagnosed by having the three predicted bands, as indicated by arrows, on the gel seen in figure 43 were sequenced by the USU Biotechnology Laboratory (not shown). The sequencing was done to confirm three things: 1) that the thaumatin sequence was still correct and no errors had occurred as a result of the PCR

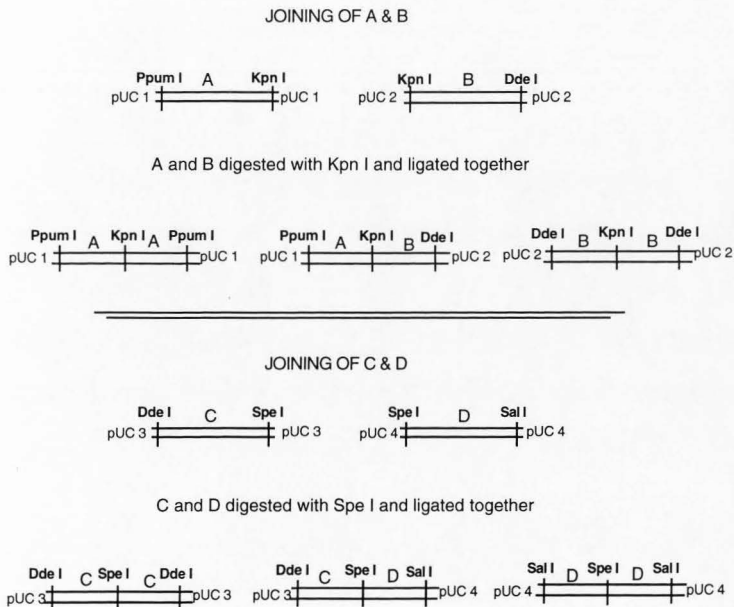


FIG. 40. Possible ligation events when joining thaumatin A & B and C & D.

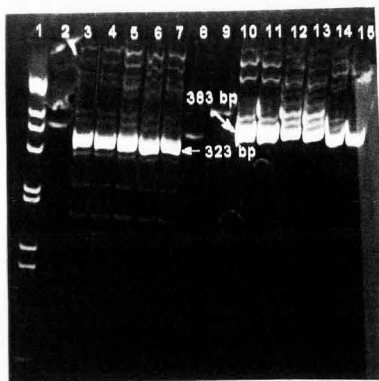


FIG. 41. 10% PAGE with PCR product of thaumatin AB and CD (gel #1012).
Lane 1: 1 kb ladder size marker;
Lanes 3-7: AB PCR product;
Lanes 10-15: CD PCR product.

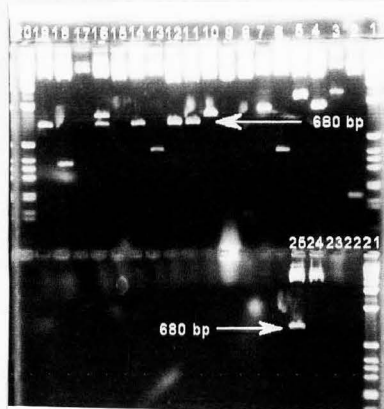


FIG. 42. Screening of thaumatin-pUC (gel # 1019).
Lanes 1, 20 & 21: 1 kb size marker;
Lanes 2-19, 24: putative thaumatin-pUC clones.

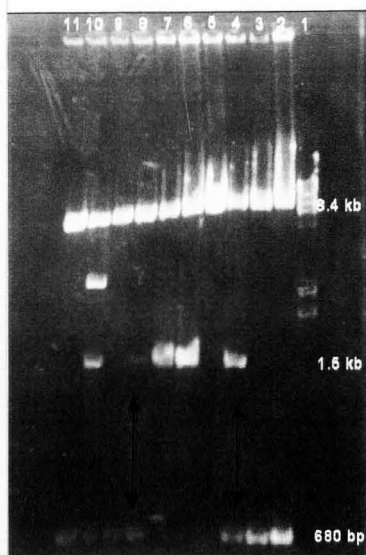


FIG. 43. Screening of thaumatin-p5'3'β-cas (gel #1049).

Lane 1: λ Hind III size marker;

Lanes 2-11: putative thaumatin-p5'3'βcas clones.

amplifications, 2) that the junction between thaumatin and the 5' region was in the correct translational reading frame, and 3) that the junction between thaumatin and the 3' region was in the correct translational reading frame. The sequencing data was analyzed to confirm that the thaumatin expression vector was correct. The brazzein expression vector was also sequenced and analysed prior to it being used for microinjection.

Preparation of expression vector for micro-injection

The DNA that comprised the expression vector (5'β-casein region, synthetic gene and 3'β-casein region) was digested away from the plasmid backbone to prepare for microinjection. This was accomplished by digesting the brazzein or thaumatin expression vectors with the enzymes AatII and MluI. It was predicted that the following bands would result from the reaction: 7.4 or 7.0 kb (desired band for thaumatin or brazzein expression vector, respectively) and 3.1 kb (plasmid). The thaumatin expression vector exhibited the predicted bands (figure 44) while the brazzein expression vector exhibited multiple bands of various sizes. An example of the preparative 1% Seaplaque GTG agarose gel used to isolate the thaumatin expression vector was shown below in figure 44.

The 7.4 kb band that was to be excised from the gel was in close proximity to uncut and partially cut DNA. The gel (figure 44) was run for a longer time (5-6 hours at

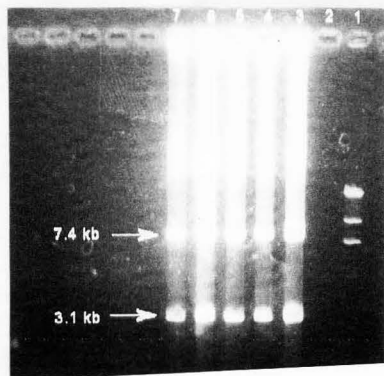


FIG. 44. Thaumatin expression vector cut for microinjection (gel #1070).

Lane 1: λ Hind III size marker;

Lanes 3-7: thaumatin 3'5'βcas cut with AatII and MluI.

75 V) than would normally be used, so that the bands would be well separated. In lane 5 of figure 44, a faint band above the 7.4 kb band can be seen. As the 7.4 kb band was excised from the gel, care was taken not to inadvertently include any uncut or partially cut DNA.

When the brazzein expression vector was digested with the above-mentioned enzymes, AatII and AgeI, additional bands were observed. To determine why these bands were present, several diagnostic restriction enzyme reactions were performed. The brazzein expression vector was cut separately with AatII and MluI. Both enzymes only have one known enzyme site; therefore, the reaction should linearize the plasmid resulting in one band at 10.1 kb. The AatII digested DNA (figure 45, lanes 3-6) yielded one band but, the MluI digested DNA (lanes 8-11) resulted in two strong bands.

It was evident from the gel in figure 45 that there were two MluI sites instead of one. The expected enzyme sites of the brazzein expression vector were shown in figure 46. This schematic can be referenced to while examining the diagnostic gels to determine that the newly identified MluI site was not located in the 3'β-casein region. It was determined that the additional enzyme recognition site was not in the 3'β-casein DNA sequence by digesting the 3' region with MluI. In figure 47, only one 2 kb-sized band was seen in lanes 7 and 8, indicating that no additional MluI site was present in the 3'β-casein DNA.

To avoid working with the MluI restriction digestion, another restriction enzyme site, BamHI, at the 3' end of the 3'β-casein gene (figure 46) was selected to excise the expression vector away from the plasmid sequences, BamHI only cut once and the reaction generated the predicted band of approximately 10.1 kb (lanes 2-5, figure 47). Additional diagnostic restriction enzyme digestions were done to ensure the extra MluI site was not within the expression vector, and that all the other pertinent restriction enzyme sites were as they should be, and would generate the correct size bands (data not shown).

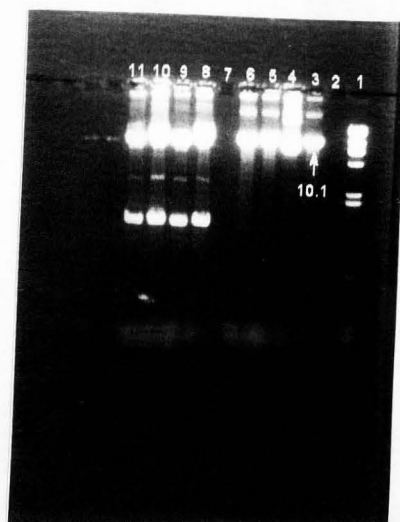


FIG. 45. Diagnostic agarose gel for brazzein expression vector using restriction enzymes AatII and AgeI (gel # 974).

Lane 1: λ HindIII size marker;

Lanes 3-6: brazzein expression vector clones cut with AatII;

Lanes 8-11: brazzein expression vector clones cut with MluI.

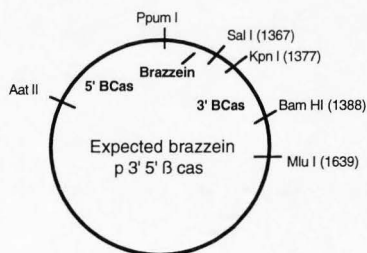


FIG. 46. Pertinent predicted restriction enzyme sites of brazzein expression vector.

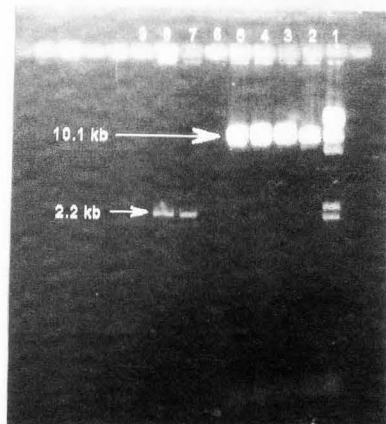


FIG. 47. Diagnostic agarose gel for brazzein expression vector using restriction enzymes BamHI and MluI (gel # 977).

Lane 1: λ HindIII size marker;

Lanes 2-5: Brazzein 3'5'βcas cut with BamHI;

Lanes 7-9: 3'β-casein region cut with MluI.

The brazzein expression vector was digested first with AatII and then with BamHI to prepare the vector for microinjection. This reaction yielded a band of approximately 6.75 kb. An example of such a preparative gel (Seaplaque GTG agarose) to isolate the brazzein expression vector for microinjection is shown in figure 48. The 6.7 kb fragment that was excised for microinjection was indicated by an arrow. The upper band was uncut DNA that was not excised from the gel.

Screening of potential transgenic animals

The ethidium bromide-stained agarose gel used for the Southern blot analysis had the following DNAs loaded in the indicated lanes (figures 49 & 50).

The positive control carrying an equivalent of about 10 copies/cell had one band on

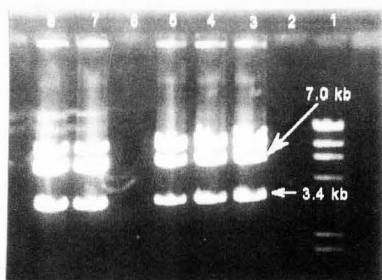


FIG. 48. Brazzein expression vector cut for microinjection (gel # 984).
Lane 1: λ HindIII size marker;
Lanes 3-5: brazzein 3'5'βcas cut with AatII and BamHI.

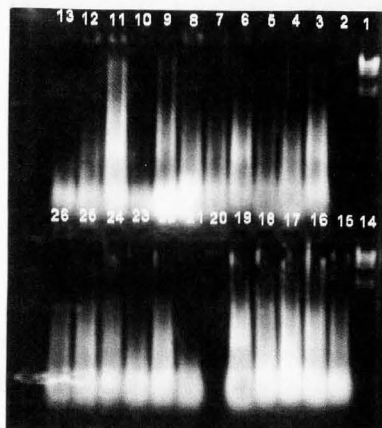


FIG. 49. Ethidium bromide stained agarose gel prepared for murine genomic Southern blot (gel # 1164).
Lanes 1 & 14: λ HindIII size marker;
Lanes 2 & 20: linearized thaumatin plasmid positive control: 10 copies and 1 copy respectively;
Lanes 3-13, 15-19 & 21-26: approximately 10 μ g of putative transgenic murine DNA per lane cut with HinfI.

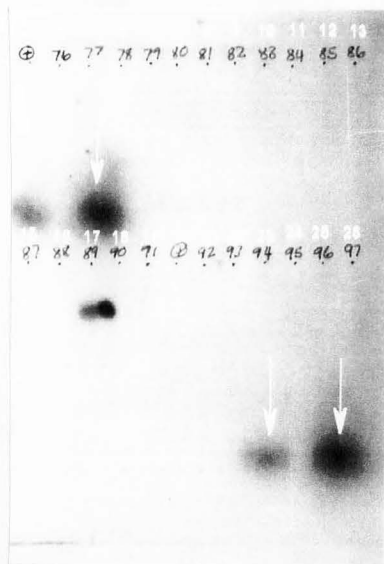


FIG. 50. Autoradiograph of murine genomic Southern blot (gel # 1164). Positive bands were identified using a probe of ^{32}P labeled thauMATIN gene. Positive bands are seen in:
Lane 2: positive control;
Lane 4: mouse #77;
Lane 23: mouse #94;
Lane 26: mouse #97.

the Southern blot hybridized filter (lane 2, figure 50). Some of the lanes containing putative transgenic murine genomic DNA also revealed signal bands, which indicated that these mice possessed DNA homologous to the thauMATIN transgene.

To verify these results, another hybridization blot was done and the results obtained (data not shown). The second blot was done using alkaline transfer (3) instead of Southern transfer, to try and improve the quality of the hybridization. The different transfer protocol showed no improvement. As in the autoradiograph shown in figure 50,

the bands were not discrete and accurate lane identification was difficult. In future reactions, fresh Denhardt's solution was made. Freshly made Denhardt's removed the problem of diffuse bands as is evident in figure 54 using goat genomic DNA.

Tail biopsies of founder (F0) and first-generation progeny (F1) mice were screened for the transgene (figures 51 & 52) to determine if the transgene was inheritable. The only positive control that was visible on the autoradiograph after exposure for 6 weeks at 80°C was the 300 copy equivalent (concentration of positive controls was calculated so as to approximate copy numbers to determine what levels of copy number could be detected in murine samples). This indicated that the sensitivity of the probe was insufficient to detect all transgenic mice with this particular gel. However, two putative

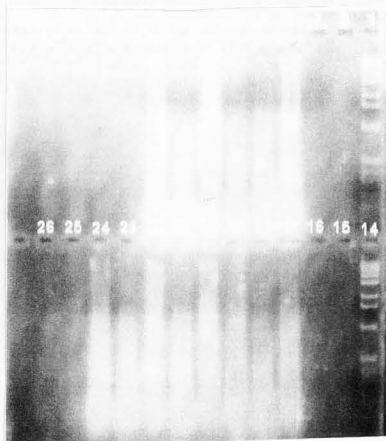


FIG. 51. Agarose gel prepared for F0 and F1 murine genomic Southern blot (gel # 1126).

Lane 1: 1 kb ladder size marker;

Lane 2: 5.5 pg positive control (thaumatin plasmid cut with *Hinf*I) - 3 copy equivalent;

Lanes 4-9: 10 μ g of putative transgenic murine DNA per lane cut with *Hinf*I;

Lane 11: 55 pg positive control - 30 copy equivalent;

Lane 13: 550 pg positive control - 300 copy equivalent;

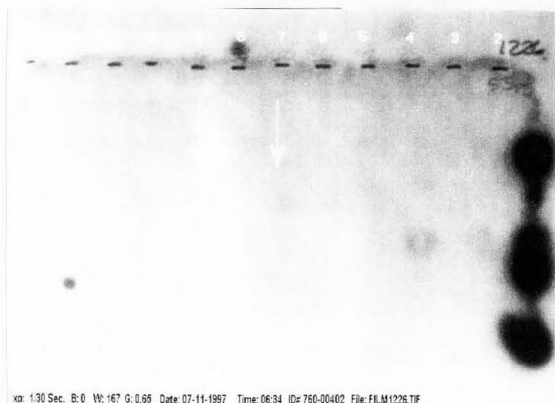


FIG. 52. Autoradiograph of F0 and F1 murine genomic Southern blot (gel # 1126).

F1 transgenic mouse samples had bands. They were mouse number 165, which was the offspring of mouse number 77 (figures 51 & 52), and 140, a founder mouse.

Although it was not part of this candidate's objectives to use goats to test the expression vector, caprine embryos were injected with the thaumatin expression vector. Fetal tissue was obtained from the umbilicus (25) at birth and screened for the transgene.

In order to have a replica of the data, the lower section of the gel (figure 53) contained the same samples as were loaded on the upper portion of gel. The only difference between the lower and upper portions of the gel being the order in which the samples were loaded onto the gel.

The positive controls were evident in lanes 2 and 4 of figure 54 on a film that was developed after 3 days. Goat #28 in lane 13 had a strong band that lined up with the lower band of the thaumatin plasmid control in lane 2.



FIG. 53. Agarose gel prepared for genomic caprine Southern blot (gel # 1218).

Lane 1: 1 kb ladder size marker;

Lane 2: 1000 copy equivalent thaumatin p3'5'βcas cut with HinfI;

Lane 4: 330 copy equivalent thaumatin-pUC cut with HinfI;

Lane 6: negative control: 10 μg non-transgenic goat DNA cut with HinfI;

Lanes 7,8 & 10-28: 10 μg putative transgenic goat DNA cut with HinfI.

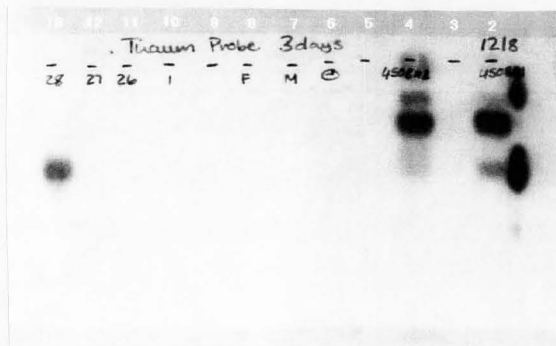


FIG. 54. Autoradiograph of genomic caprine Southern blot after 3-day exposure (gel # 1218). Positive bands were identified using a probe of ^{32}P labeled thaumatin gene.

Lane 1: 1 kb ladder size marker;

Lane 2: 1000 copy equivalent thaumatin p3'5'βcas cut with HinfI;

Lane 4: 330 copy equivalent thaumatin-pUC cut with HinfI;

Lane 6: negative control: 10 μg non-transgenic goat DNA cut with HinfI;

Lanes 7,8 & 10-28: 10 μg putative transgenic goat DNA cut with HinfI.



FIG. 55. Autoradiograph of genomic caprine Southern blot after 6-week exposure (gel # 1218). Positive bands were identified using a probe of ^{32}P labeled thaumatin gene.

Lane 14: 1 kb ladder size marker;

Lane 15: 1000 copy equivalent thaumatin p3'5'βcas cut with *Hinf*I;

Lane 16: 330 copy equivalent thaumatin-pUC cut with *Hinf*I;

Lane 17: negative control: 10 μg non-transgenic goat DNA cut with *Hinf*I;

Lanes 18,19,21,23-24 & 26: 10 μg putative transgenic goat DNA cut with *Hinf*I.

The overexposed autoradiograph at 6 weeks demonstrated that the positive band of number 28 lined up with the lower band seen on the positive controls (demarcated by an arrow, figure 55). In addition, there was no background seen in the lanes that contained the other putative transgenic goat samples (lanes 18,19,21,23-24) or the negative control (lane 17). The extraneous bands to the left of lane 15 were residual radiographic signals from the lane containing the 1 kb ladder DNA size marker that was cut away from the filter.

Objective 3: Detection of protein in the milk of the transgenic mice

ELISA assays were developed to detect the thaumatin or brazzein proteins (figure 15). In order to optimize the reaction for the various components, multiple checker-board assays were done. The optimal amount of goat and rabbit antibody was determined by using \log_{10} serial dilutions of the antibody-containing serum and comparing the results at different concentrations. The first antibody (goat) was used at three dilutions between 1×10^{-3} to 1×10^{-5} in the following two figures (figures 56 and 57), which shows the results of the brazzein assay. The second antibody (rabbit), directed against brazzein, was used at 1×10^{-2} dilution throughout the assay. The difference between the two figures was that in figure 56 the antigen was diluted in phosphate-buffered saline (PBS) (3), while in figure 57 the antigen was diluted in bovine milk.

In figure 56, the absorbance reading at 450 nm decreased in a linear fashion corresponding to the decreasing antigen concentration. When the milk was used as a diluent instead of PBS, however, no linearity was observed (figure 57). The reason for inhibition was not known.

The thaumatin assay was done using the same parameters as with the brazzein, except that antisera specific for the crude thaumatin plant protein were used for the standard curve. The same interference with linearity observed with the brazzein assay was also observed with the thaumatin assay when milk was used as the diluent instead of PBS (figures 59 & 58, respectively).

It was evident that transgenic proteins secreted into the milk could not be identified using antibodies without further development of the ELISA. As another method for identifying the brazzein or thaumatin proteins in mouse milk, a milk sample from a transgenic female mouse (#77) was analyzed by capillary electrophoresis. Controls of crude thaumatin plant protein were electrophoresed to determine the electrophoretic location of thaumatin when compared to the normal mouse milk proteins. The main peak of crude thaumatin plant protein (figure 60, top panel) was compared with the proteins from the nontransgenic mouse milk control (figure 60, middle panel) and a sample containing an

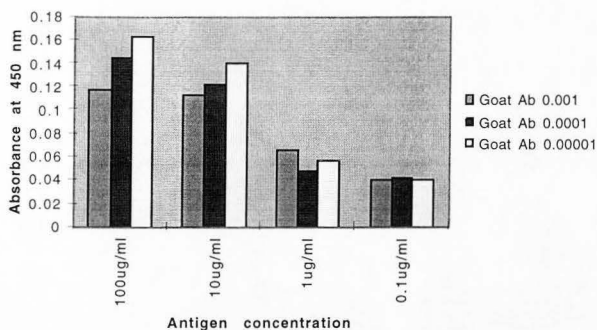


FIG. 56. ELISA with brazzein multiple antigenic peptide diluted in PBS. The optical density (OD) is shown on the y-axis against the concentrations of antigen added to the wells on the x-axis. Three different concentrations of goat capture antibody were used.

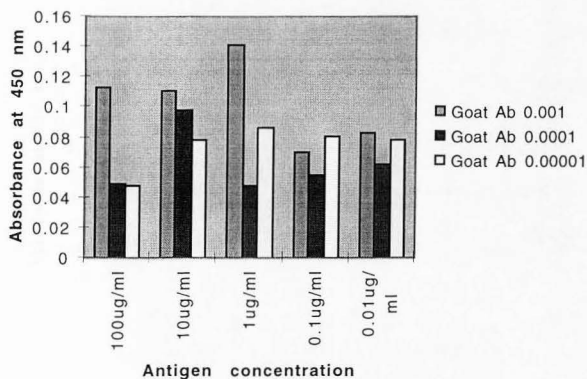


FIG. 57. ELISA with brazzein multiple antigenic peptide diluted in bovine milk. The optical density (OD) is shown on the y-axis against the concentrations of antigen added to the wells on the x-axis. Three different concentrations of goat capture antibody were used.

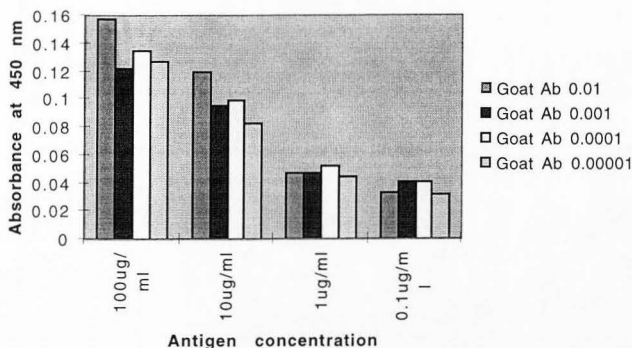


FIG. 58. ELISA with thauMATIN plant protein diluted in PBS. The optical density (OD) is shown on the y-axis against the concentrations of antigen added to the wells on the x-axis. Four different concentrations of goat capture antibody were used.

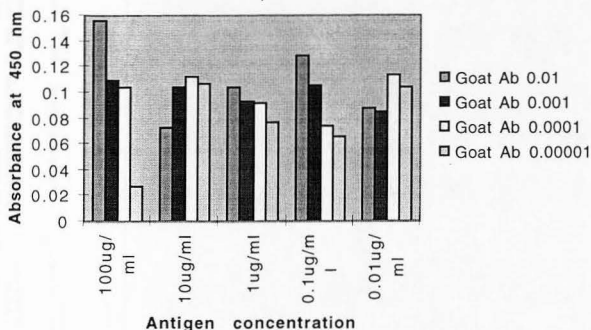


FIG. 59. ELISA with thauMATIN plant protein diluted in bovine milk. The optical density (OD) is shown on the y-axis against the concentrations of antigen added to the wells on the x-axis. Four different concentrations of goat capture antibody were used.

equal (v/v) mixture of thaumatin and mouse milk (figure 60, bottom panel). The thaumatin peak was to the right of two mouse milk protein peaks. This indicated the region in which one could expect transgenic thaumatin to appear in transgenic mouse milk, although there could be slight variation due to differences between glycosylation of the protein in the mammary gland and glycosylation in the wild-type plant protein.

In the lower panel of figure 60 that contained mouse milk with added thaumatin plant protein, the profile of the mouse milk proteins was altered. The intensity and location of the various endogenous milk proteins were altered as a result of the added crude plant thaumatin protein. This could be due to the possible existence of contaminating proteinase activity in the crude thaumatin preparation from plants that would alter the normal mouse milk profile. Contaminating proteases would not be present in the transgenic protein, however.

Based on the results from figure 60, the transgenic thaumatin protein was expected to be located to the right of a milk-protein doublet. Enlargements were made of two of the graphs seen in figure 60 for the purpose of comparison with the transgenic mouse samples (figures 61 & 62). Milk samples from a previously identified transgenic thaumatin mouse #77 (figure 63), and samples from three non-thaumatin transgenic mice, were assayed. The data have been represented in figures 63 through 66 as computer enlargements of the doublet region where exogenous thaumatin was normally located. An additional peak (arrow) at this location was identified in the transgenic mouse #77 (figure 63), but it was not seen in the three other non-thaumatin transgenic mice (figures 64-66). The absence of a peak in the same region in the non-thaumatin transgenic mouse samples was indicated by an arrow corresponding to the position of the extra peak seen in the transgenic mouse (figures 64-66).

Goat milk cannot be analyzed at this stage because the kids need to mature and give birth before being able to obtain milk from them.

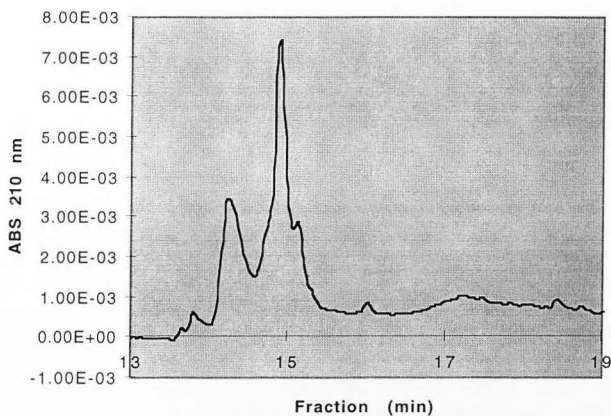


FIG. 61. Capillary electrophoresis of mouse milk control. The fraction area identified as containing the thaumatin spike has been enlarged. Note the absence of peak (arrow) at 16-16.5 minutes as compared to the large peak seen in this region in figure 62.

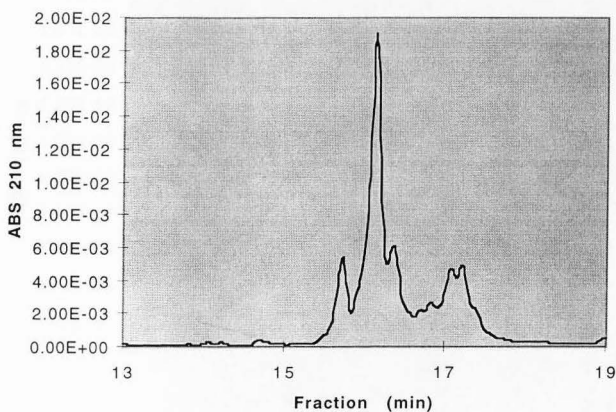


FIG. 62. Capillary electrophoresis of crude thaumatin protein. The peak seen with thaumatin protein appears in the fraction region of 16-16.5 min.

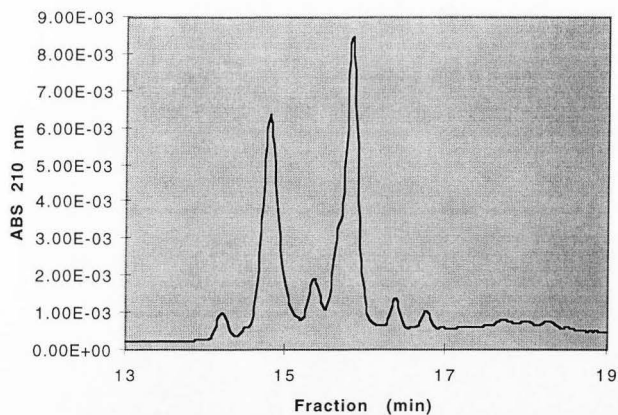


FIG. 63. Capillary electrophoresis of milk from mouse #77. An additional peak is seen in the region of 16-16.5 min (arrow).

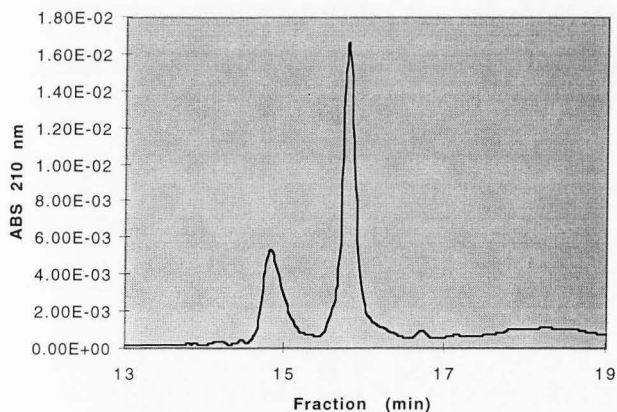


FIG. 64. Capillary electrophoresis of nontransgenic mouse #8 litter mate milk. No additional peak is seen in the region of 16-16.5 min.

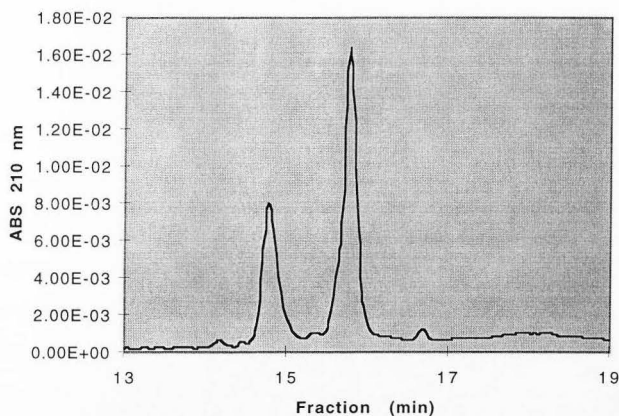


FIG. 65. Capillary electrophoresis of nontransgenic mouse #19 litter mate milk. No additional peak is seen in the region of 16-16.5 min.

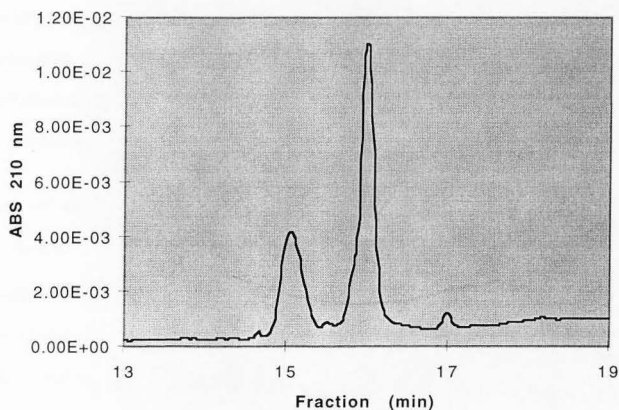


FIG. 66. Capillary electrophoresis of milk from a nontransgenic mouse.

DISCUSSION

Objective 1: Construction of expression vector

β -casein gene

The regulatory elements of the protein expression vectors used in this research consisted of 3.8 kb of the 5' bovine beta-casein and 2.2 kb of the 3' bovine beta-casein gene regulatory sequence. This was predicted to be an effective promoter based on the beta-casein gene being one of the most predominant milk proteins at an amount of 9 g/L (28% of total milk protein). The 5' and 3' sequences included three introns. The inclusion of introns in transgenic expression vectors was shown to markedly increase the transcription efficiency (22). Varying sizes of regulatory sequences from the bovine beta-casein gene have been successfully used in expression vectors for a number of different proteins that have differing post-translational modifications (37). We hypothesize that using a larger regulatory region than has been previously used containing multiple introns and non-coding exons increases the probability of a functional protein being secreted at a concentration suitable for commercial production. Consequently, the use of bovine beta-casein regulatory sequences for this study was justified.

Cloning of the expression vector

Several novel cloning techniques were used in the making of the expression vector. These approaches worked well and should have applications for cloning a variety of genes into plasmids quickly and efficiently. Firstly, PCR products were treated with proteinase K to digest interfering substances, presumably so that the PCR products could be cloned more effectively. Secondly, T4 polymerase and polynucleotide kinase were used together in the blunt end reaction of the PCR products in preparation for ligation into pUC-SmaI. Additionally, our plasmid was found to have low transformation efficiency when subjected to agarose gel purification using beta-agarase digestion or dialysis. This was possibly due

to traces of agarose that were remained with the DNA. To bypass this problem, no gel purification was done. The ligation was done directly after restriction enzyme digestion and the transformants were screened using colony hybridization (figure 26).

Synthesis of long oligonucleotides

In order to have the gene containing common codons that would facilitate mammalian protein expression, synthetic genes had to be made. Synthesis of long oligonucleotides is not frequently done due to high error rates that can occur when making long oligonucleotides of more than 60 bases. Therefore, thaumatin and brazzein were synthesized by making several oligonucleotides, ranging in size from 50 to 100 bases, that could be assembled together into the complete gene. This permitted the generation of DNA strands 220 and 660 bp in length. Effective cloning of the long synthetic gene for thaumatin (660 bp) was accomplished by designing overlapping oligonucleotides that could be hybridized together to make separate fragments of the gene. These fragments were then ligated together using restriction enzyme sites that were designed into the ends of the fragments. The restriction enzyme sites were designed so that when oligonucleotides were ligated together, the complete gene sequence would be in the correct translational reading frame (figure 8). A non-palindromic enzyme site was selected for the middle of the gene to minimize the formation of self-ligation between oligonucleotides. Ligation was done with either PCR product that had been isolated from a PAGE or directly on PCR product. No transformation problems were seen with the PAGE purified sequences that had been ligated into pUC.

Objective 2: Generation of transgenic lines to test the expression vector

Identification of transgenic animals

Probes from the vector DNAs were used to identify transgenic mice and a goat. The thaumatin transgene was identified in founder mice #77 (female) and #140 (male) and

has been shown to be inheritable by identification of offspring from mouse #77. The studies on brazzein mice were stopped because an error in the original brazzein sequence published in the patent was synthesized and would have produced changes in the amino acid sequence. These errors would have probably yielded a nonfunctional protein. Therefore, for the sake of time and money, further investigation was suspended on the brazzein work. However, the brazzein gene sequence error is currently being corrected.

Objective 3: Detection of protein in the milk of the transgenic mice

Identification of thaumatin protein in mouse milk

The thaumatin protein has tentatively been identified in mouse milk from transgenic mouse #77. The protein was detected as an additional peak when compared to nontransgenic mouse milk samples using capillary electrophoresis. This could be due to the possible existence of contaminating proteinase activity in the crude thaumatin preparation from plants that would alter the normal mouse milk profile. Contaminating proteases would not be present in the transgenic protein, however.

It is anticipated that the protein will be produced at a yield of approximately 2-3g/L. As stated in the Literature Review, all indications are that the thaumatin protein will be expressed as a functionally sweet protein. However, this is not known at this point. This will be tested using non-human primate taste tests of thaumatin milk compared to negative controls of non-transgenic milk from litter mates and milk from the transgenic animal that has been deactivated by heat treatment. Thaumatin has an extremely sweet taste and even at low levels its presence should be evident. There are several reasons why functional protein would not be secreted. The signal peptide could be incorrectly or poorly recognized, resulting in little or no protein present in the milk. The protein could also be incorrectly folded in a mammalian system, such that it is not functionally sweet. Additionally, the

physiological environment provided by the milk could result in the protein not being functional, although I believe that this is unlikely.

From an economic standpoint, if the protein is not functionally sweet, because of the relatively low costs of mammary gland production as compared to the natural supply and some of the other methodologies used, it could still be cost effective to produce the protein which is not functional and then chemically modify it.

Immunoassay to detect protein

An ELISA has been designed to detect and quantify the thaumatin and brazzein proteins. It was determined during development of the assay that the presence of milk interfered with the specificity or linearity of the assay. It is hypothesized that the milk contains inhibitory substances that block the epitope on the goat antibody, thereby preventing antigen recognition. Alternatively, there could be other antibodies present in the serum that are binding the milk proteins. This would result in a high background level, thereby preventing the detection of any significant color change resulting from the varying thaumatin protein concentrations. As evidence of this, when the assay was performed with antigen diluted in PBS instead of milk, a linear relationship between absorbance and antigen concentration was seen. While there is a linear relationship, diluting in PBS is only effective for the controls. Because the proteins are produced in milk, one of two things needs to be done. The antigen (thaumatin protein) needs to be isolated from the milk prior to performing the assay. Alternatively, the antibody could be purified by affinity chromatography. By binding purified antigen to chromatography beads and passing the serum over the matrix, the antibody specific to thaumatin and brazzein would be retained in the column and could then be eluted.

CONCLUSION

This project has resulted in:

1. Successful cloning of an expression vector that has tentatively been shown to produce the thaumatin protein in the milk of genetically engineered animals.
2. Development of a novel technique to produce long error-free synthetic genes.
3. Development of an ELISA that can detect the thaumatin and brazzein proteins in a concentration-dependent manner.

In order to achieve the long-term goal of using these expression vectors to produce and detect the proteins in the milk of dairy animals, the following work remains to be done:

1. Characterize the stability of the transgene from one generation to the next. This will be done by continuing to screen transgenic mice from the founder animals that have been shown to carry the transgenes. The relative copy numbers of the transgenes over several generations will then be compared as an indicator of stability of the gene at the integration site peculiar to each transgenic line.
2. Mutate the incorrect brazzein expression vector into the correct sequence to yield functional protein expression. Insert the correct gene sequence into the expression vector using the same protocol described in the Methods section. Generate transgenic mice to test the brazzein expression vector and characterize the transgene and protein as described for thaumatin.
3. Optimize the protocol for obtaining and storing mouse milk so that sufficient quantities can be obtained from each mouse to screen for the protein. There are several stages at which this process could potentially be enhanced. The oxytocin levels and time between injection and initiation of the milking should be optimized. In addition, the homemade mouse milking pump could possibly be modified to get better suction. Lastly, when frozen for long periods of time, the mouse milk seems to become very thick and difficult to work with for the ELISA. Perhaps the milk should either be

assayed soon after it is obtained with only refrigeration at 4 C or a medium could be obtained to store the milk in. Because mouse milk is initially thicker and more concentrated than milk obtained from dairy animals, I do not anticipate this being a problem encountered with large-scale protein production.

4. Develop a protocol for purifying transgenic protein away from milk so that the protein can be characterized biochemically.
5. Obtain milk from goat # 28 and screen for the protein by ELISA, capillary electrophoresis, and Western blot to validate the presence of the thaumatin.
6. Once the protein presence has been confirmed, perform physiological taste studies and biochemical studies to show that the protein is secreted and folded such that it is functional.

LITERATURE CITED

1. **Adesina, S. K.** 1994. Thaumato-coccous danielli, source of sweet proteins, p. 20-34. *In* M. Witty and J. D. Higginbotham (eds.), Thaumatin. CRC Press, Boca Racaon, Fla.
2. **Ashktorab, H., C. Thonabulsombat, W. A. Reed, and K. L. White.** 1995. Bovine Beta-casein (BBC) gene promoter activity and hormonal induction (HI) of its expressionin Mammary epithelial cell (MEC) line. FASEB 9:83.
3. **Ausubel, F.M., R. Brent, F.E. Kingston, D.D. Moore, J. G. Seidman, J.A. Smith, and K. Struhl.** 1994. Current Protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience. New York, N.Y.
4. **Bischoff, R., E. Degryse, F. Perraud, W. Dalemans, D. Ali-Hadjji, D. Thejpot, E. Devinoy, L. Houdebine, and A. Pavirani.** 1992. A 17.6 kbp region located upstream of the rabbit WAP gene directs high level expression of a functional human protein variant in transgenic mouse milk. FEBS 305:265-268.
5. **Boy, C.** 1994. The natural supply of thaumatin, p. 37-45. *In* M. Witty and J. D. Higginbotham (eds.), Thaumatin. CRC Press, Boca Raton, Fla.
6. **Brinster, R. L., J. M. Allen, R. R. Behringer, R. E. Gelinas, and R. D. Palmiter.** 1988. Introns increase transcriptional efficiency in transgenic mice. Proc. Nat. Acad. Sci. USA 85:836-840.
7. **Brown, T. A.** 1995. The polymerase chain reaction, p. 228-242. *In* T. A. Brown (ed.), Gene cloning, an introduction. Chapman & Hall, London.
8. **Crowe, J. S.** 1991. Improved cloning efficiency of polymerase chain reaction (PCR) products after proteinase K digestion. Nucleic Acids Research 19:184.
9. **Edens, L., L. Heslinga, R. Klok, A. M. Ledeboer, J. Maat, M. Toonen, and C. Visser.** 1982. Cloning of cDNA endoding the sweet-tasting plant protein thaumatin and its expression in Escherichia coli. Gene 18:(1)1-12.
10. **Edens, L., and H. Van Der Wel.** 1985. Microbial synthesis of the sweet-tasting plant thaumatin. Trends in Biotechnology. 3:61-63.
11. **Etheridge, K.** 1994. The sales and marketing of talin, p. 47-59. *In* M. Witty and J. D. Higginbotham (eds.) Thaumatin. CRC Press, Boca Raton, Fla.
12. **Gordon, J. W., G. A. Scangos, D. J. Plotkin, J. A. Barbosa, and F. H. Ruddle.** 1980. Genetic transformation of mouse embryos by microinjection of purified DNA. Proc. Natl. Acad. Sci. USA 77:7380-7384.

13. **Hellekant, G., and D. Ming**, inventors. 1994. Brazzein sweetener. USA. (Patent # 5,326,580)
14. **Higginbotham, J. D.** 1994. The application of thaumatin in animal feed and in pet food, p. 83-97. *In* M. Witty and J. D. Higginbotham (eds.), Thaumatin. CRC Press, Boca Raton, Fla.
15. **Illingworth, C., G. Larson, and G. Hellekant**. 1988. Secretion of the sweet-tasting plant protein thaumatin by *Bacillus subtilis*. *Biotechnology Letters* **10**:587-592.
16. **Janne, J., J. Hyttinen, T. Peura, M. Tolvanen, L. Alhonen, R. Sinervirta, and M. Halmekyto**. 1994. Transgenic bioreactors. *Int. J. Biochem.* **26**:859-870.
17. **Koozan, D., G. Hobom, and H. M. Seyfer**. 1991. Genomic organization of the bovine alpha-S1 casein gene. *Nucleic Acids Research* **19**(20):5591-5596.
18. **Krimpenfort, P., A. Rademakers, W. Eyestone, A. Van der Schans, S. Van den Broek, P. Kooiman, E. Kootwijk, G. Platenburg, F. Pieper, R. Strijker, and H. De Boer**. 1991. Generation of transgenic dairy cattle using invitro embryo production. *Biotechnology* **9**:844-847.
19. **Lee, J., J. L. Welckmann, R. K. Koduri, P. Ghosh-Dastidar, K. Saito, L. C. Blair, T. Date, J. S. Lai, S. M. Hollenberg, and R. L. Kendall**. 1988. Expression of synthetic thaumatin genes in yeast. *Biochemistry* **27**:5101-5107.
20. **Lewin, B.** 1994. *Genes*. Vth ed. Oxford University Press Inc., Oxford, England.
21. **Lewis, S. M., and J. H. Carraway**. 1992. Large animal models of human disease. *Lab Animal* **Jan**:22-29.
22. **Maga, E. A., G. B. Anderson, and J. D. Murray**. 1995. The effect of mammary gland expression of human lysozyme on the properties of milk from transgenic mice. *J Dairy Sci* **78**:2645.
23. **Maniatis, T., E. F. Fritsh, and J. Sambrook**. 1982. DNA manipulation, p. 1.21-1.30 *In* N. Ford (ed.), *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
24. **Ming, D. and G. Hellekant**. 1994. Brazzein, a new high-potency thermostable sweet protein from *Pentadiplandra brazzeana* B. *FEBS letters* **355**:106-108.
25. **Moore, K. L.** 1977. *The developing human*. Saunders Press, Philadelphia, Pa.
26. **Mullis, K. B., and F. A. Faloona**. 1986. Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harbor Symp. Quant. Biol.* **51**:263-273.

27. **Palmiter, R. D., R. L. Brinster, R. E. Hammer, M. E. Trumbauer, M. G. Rosenfeld, and N. C. Birnberg.** 1982. Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. *Nature* **300**:611-615.
28. **Pfeffer, S. R., and J. E. Rothman.** 1987. Biosynthetic protein transport and sorting by the endoplasmic reticulum and golgi. *Ann. Rev. Biochem* **56**:829-852.
29. **Riego, E., J. Limonta, A. Aguilar, A. Perez, R. De Armas, R. Solanto, B. Ramos, F. O. Castro, and J. De la Fuente.** 1993. Production of transgenic mice and rabbits that carry and express human tissue plasminogen activator cDNA under the control of a bovine alpha S1 casein promoter. *Theriogenology* **39**:1173-1185.
30. **Sambrook, J., E. F. Fritsh, and T. Maniatis.** 1989. *Molecular Cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
31. **Swaigood, H. E.** 1992. *Advanced dairy chemistry*, Elsevier Science Publishers Ltd., Cambridge, England.
32. **Velander, W. H., J. L. Johnson, R. L. Page, C. G. Russell, A. Subramanian, T. D. Wilkins, F. C. Gwazdauskas, C. Pittius, and W. N. Drohan.** 1992. High-level expression of a heterologous protein in the milk of transgenic swine using the cDNA encoding human protein C. *Proc. Natl. Acad. Sci. USA* **89**:12003-12007.
33. **Villote, J., S. Soulier, M. Stinnakre, M. Massoud, and J. Mercier.** 1989. Efficient tissue-specific expression of bovine alpha-lactalbumin in transgenic mice. *Eur. J. Biochem.* **186**:43-48.
34. **Wall, R. J., V. G. Pursel, A. Shamay, R. A. McKnight, C. W. Pittius, and L. Henninghausen.** 1991. High level synthesis of a heterologous milk protein in the mammary glands of transgenic swine. *Proc. Nat. Acad. Sci. USA* **88**:1696-1699.
35. **Weickmann, J. L., L. C. Blair, and G. L. Wilcox.** 1994. High level expression of thaumatin in *Saccharomyces cerevisiae*, p. 151-159. *In* M. Witty and J. D. Higginbotham (eds.), *Thaumatococcus*. CRC Press, Boca Raton, Fla.
36. **Witty, M.** 1990. Thaumatococcus II--a palatability protein. *Trends in Biotechnology* **8**(May) 113-116.
37. **Young, M. W., W. B. Okita, E. M. Brown, and J. M. Curling.** 1997. Production of biopharmaceutical proteins in the milk of transgenic dairy animals. *BioPharm* **10**(6):34-38.