2004

BIOINFORMATIC ANALYSIS OF A MAMMALIAN BIP GENE FOR INSERTION INTO GREEN ALGAE AND COMPARISON OF ITS POSSIBLE EFFECTS ON THE SYNTHESIS OF A MAMMALIAN ANTIBODY

Katrina Ghazanfar
Virginia Commonwealth University

School of Medicine
Virginia Commonwealth University

This is to certify that the dissertation prepared by KATRINA PATRICIA GHAZANFAR entitled BIOINFORMATIC ANALYSIS OF A MAMMALIAN BIP GENE FOR INSERTION INTO GREEN ALGAE AND COMPARISON OF ITS POSSIBLE EFFECTS ON THE SYNTHESIS OF A MAMMALIAN ANTIBODY has been approved by the student advisory committee as satisfactory for completion of the dissertation requirement for the degree of Doctor of Philosophy.

Fang-Sheng Wu, PhD., School of Medicine and College of Humanities and Sciences, Director of Dissertation

Guy A. Cabral, PhD., School of Medicine

Kathleen McCoy, PhD., School of Medicine

Jennifer K. Stewart, PhD., College of Humanities and Sciences

L. Shozo Ozaki, PhD., School of Medicine

Jerome F. Strauss, III, M.D., PhD., Dean, School of Medicine

F. Douglas Boudinot, PhD., Dean, Graduate School

Date

BIOINFORMATIC ANALYSIS OF A MAMMALIAN BIP GENE FOR INSERTION INTO GREEN ALGAE AND COMPARISON OF ITS POSSIBLE EFFECTS ON THE SYNTHESIS OF A MAMMALIAN ANTIBODY

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, Microbiology and Immunology with concentration in Molecular Biology and Genetics at Virginia Commonwealth University.

by

KATRINA PATRICIA GHAZANFAR
M.S., Health Services Administration, Central Michigan University, 2000
B.S., Medical Technology, Fayetteville State University (UNC), 1994

Director: Dr. Fang-Sheng Wu,
Associate Professor, Department of Microbiology and Immunology

Virginia Commonwealth University
Richmond, Virginia
December 2009

Acknowledgement

Taking a moment to reflect on these years that have passed in such a blur, I must acknowledge those special individuals who have enabled me to achieve this educational goal. First, I would like to give my sincere appreciation to Dr. Fang-Sheng Wu, who guided me through this dissertation project with never-ending patience and much needed support. Next, I would like to express my heartfelt gratitude to the members of my Graduate Advisory Committee, Dr. Guy Cabral, Dr. Kathleen McCoy, Dr. Jennifer Stewart, and Dr. Shozo Ozaki.

Without hesitation, I must acknowledge three of my closest Sisters, who have stood beside me during these arduous times, and who have picked me up and carried me when I stumbled and fell and could not find my bearings again. Sis. Rosalynn Intisar Hasaan, US Army CW2 (retired), Sis. Joy Renee' Fitchett, MT (ASCP), and Sis. Aneeqah Siddeeq Ferguson, M.S., Biostatistics, may Allah (SWT) always reward you for the unwavering support you have given so abundantly to me and my daughters, Ameenah Nuri and Aaliyah Taaj Saalih. Jazakallah for your guidance, your love, your nurturing, innumerable times led me back to the Straight Path and encouraged me to remember who is Most Merciful!

Time would surely stand still without my acknowledgement of my dearest Brother, Raja Ghazanfar Hussain. Thank you for loving me and shining that bright light so that I could find my way out of that dark tunnel. Thanks for enriching my life and enhancing it with two small joys: Ambar-Faith Bukola and Raja Dare'-Sulaimon Ghazanfar. For all of the good deeds that you have done, may Allah (SWT) bestow His greatest blessings upon

you and grant you everlasting happiness. Each Eid-ul-Adha will always be especially remembered with immeasurable fondness.

Lastly, I would like to dedicate this dissertation to the memory of my Blue Point Siamese, Pumpkin, who at the tender age of 12 years and 11 months lost her life to hepatic neoplasia. I miss you more each day. Thanks for loving me unconditionally for so many years. Thanks for staying up with me to watch the many sunrises in North Carolina, Florida, New York, Maryland, New Jersey and finally, Virginia. Thanks for "singing" with me in the shower. Thanks for being my truest friend who intently listened without judgment. You can never be replaced in my heart and you are sorely missed by all of us, especially Brooklyn. May you rest in peace and insha Allah, we will be reunited in Paradise.

Table of Contents

Page

Acknowledgements……..……..…….iii

List of Tables.....viii

List of Figures ix

List of Abbreviations..... xi

Abstract xv

Chapter

1. BACKGROUND 1

A. Bioinformatics and Molecular Biology 1

B. Genetically Modified Organisms - Bacteria 3

C. Genetically Modified Organisms – Plants 5

D. Genetically Modified Organisms - Animals..... 6

E. Molecular “Pharming” of Mammalian Proteins 7

F. Introducing Mammalian Proteins into Plants..... 12

G. Increasing Yield of Mammalian Proteins in Plants 13

H. Advantages of Using Green Algae…. ……..…….16

I. Plasmolysis and Electroporation 18

2. PROTEINS ANALYZED IN THIS INVESTIGATION..... 23

A. Glucose Regulated Protein 78 23

B. Humanized anti-carcinoma antibody 49 25

C. Green Fluorescent Protein..... 27

D. Ble^{R} gene 29

3. ORIGINS of the PROTEINS ANALYZED in THIS STUDY 31

A. p35-49-GFP and 35s promoter 31

B. pSP124S and RBCS2 promoter 31

4. METHODS and MATERIALS..... 34

A. National Center for Biotechnology Information (NCBI) Database…..34

B. Assessing Homology Using a Phylogeny Tree……..….35

C. CLC Main Workbench 5.1…….35

D. Media and Growth Requirements for Algae……..37

E. Electroporation Techniques Used for Algae……..38

F. Starch-Embedding Method……40

5. RESULTS……42

A. National Center for Biotechnology Information (NCBI) Database…...42

B. Assessing Homology Using a Phylogeny Tree……..….60

C. CLC Main Workbench 5.1…….62

D. Electroporation Techniques Used for Algae……..65

6. DISCUSSION 70

References 75

Appendices 87

A Plasmid Map of pA78H 88

B Plasmid Map of p35S-49-GFP 89

C Plasmid Map of pSP124S 90

Vita……91

List of Tables

Page

Table 1: Sources of Proteins Analyzed…….……...…..33

Table 2: Stock Solutions for TAP Media. ..... .37

Table 3: Electroporation Parameters Used for Green Algae Experiments. 39

Table 4: List of Eukaryotes Used in Comparison Study of GRP 78 Proteins..... 43

Table 5: Sequence Alignment Order and Scores from NCBI Database…..……..…….45

Table 6: Comparison of Known Homologue and Non-Homologue to HuGRP 78……...63

Table 7: Comparison and Contrast of Atomic Composition and Count Frequency……..64

List of Figures

Page

Figure 1: Single Chain IgG molecule……..26

Figure 2: GFP Structure Computerized Depiction.....……28

Figure 3: pSP124S RubisCO Cassette……...……33

Figure 4: NCBI Data from GRP78 Comparison Output…….……44

Figure 5: Human GRP 78 and GRP 78 from *Mus musculus* (mouse) Data…….……...46

Figure 6: Human GRP 78 and Heat Shock Protein 5 from *Rattus norvegicus* (rat)……...47

Figure 7: Human GRP 78 and GRP 78 from *Plasmodium falciparum* (malaria) 48

Figure 8: Human GRP 78 and HSP 70 chaperone from *Aspergillus fumigatus* (fungi).....49

Figure 9: Human GRP 78 and GRP 78 from *Gallus domesticus* (chicken)……..….50

Figure 10: Human GRP 78 and HSP 70 from *Entamoeba histolytica* (amoeba)……....51

Figure 11: Human GRP 78 and HSP 7C from *Drosophilia melanogaster* (fruit fly)…..…52

Figure 12: Human GRP 78 and a GRP 78 from *Saccharomyces cerevisiae* (yeast)……...53

Figure 13: Human GRP 78 and a GRP 78 from *Mesocricetus auratus* (hamster)……..54

Figure 14: Human GRP 78 and a GRP 78 from *Bos taurus* (cow)…….55

Figure 15: Human GRP 78 and a GRP 78 from *Xenopus laevis* (frog)…….…….56

Figure 16: Human GRP 78 and a GRP 78 from *Spinacia oleracea* (spinach)……..….57

Figure 17: Phylogeny Tree from Organisms BLASTed Against Human GRP 78…….61

Figure 18: Effect of Field Strength on Transformation Efficiency in 0.4 M Sucrose……66

Figure 19: Effect of Field Strength on Transformation Efficiency in 0.4 M Sorbitol…….67

Figure 20: Effect of Field Strength on Transformation Efficiency in Ultra pure Water…..69

Figure 21: Effect of Field Strength on Transformation Efficiency in 0.4 M Mannitol……69

List of Abbreviations

ASCP	American Society of Clinical Pathologists
BiP	Immunoglobulin Binding Protein
BLAST	Basic Local Alignment Search Tool
Ble^R	Bleomycin Resistance Gene
BLOSUM	Blocks of Amino Acids Substitution Matrix
bn	Billion
B.S.	Baccalaureate of Science
BSA	Bovine Serum Albumin
Bt toxin	*Bacillus thuringiensis* Toxin
C. reinhardtii	*Chlamydomonas reinhardtii*
CEA	Carcinoembryonic Antigen
C_{Hx}	Constant Heavy Chain of Immunoglobulin
C_L	Constant Light Chain of Immunoglobulin
CW2	Chief Warrant Officer 2
dNTPs	Deoxynucleoside Triphosphates
E. coli	*Escherichia coli*
EGFR	Epidermal Growth Factor Receptor

EMBL	European Molecular Biology Laboratory
EPSPS	5-Enolpyruvyl-Shikimate-3-Phosphate Synthase
ER	Endoplasmic Reticulum
EtBr	Ethidium Bromide
FOR	Forward Primer
GEO	Genetically Engineered Organisms
GFP	Green Fluorescent Protein
Gly^{67}	Glycine, 67^{th} Amino Acid Residue
GMOs	Genetically Modified Organisms
GRP 78	Glucose Regulated Protein, 78 Kilodaltons
GRPHT	Glucose Regulated Protein, Chinese Hamster
GUS	Beta-D-Glucuronidase
HAMA	Human Anti-Mouse Antibody
HuCC49	Humanized Pancarcinoma Antibody CC49
HSP or Hsp	Heat Shock Protein
IgA	Immunoglobulin Type A
IgG	Immunoglobulin Type G
kDa	Kilodaltons
KDEL	Lysine-Aspartate-Glutamate-Leucine
M	Molar
MB	Mega Bases

ml	Milliliters
M.S.	Masters of Science
MT	Medical Technologist
mt+	Mating Type Positive, Wild Type Strain
NCBI	National Center for Biotechnology Info
NE	Nebraska
ng	Nanograms
OD	Optical Density
PCR	Polymerase Chain Reaction
PELS	Perkin Elmer Luminescence Spectrometer
Pfu	*Pyrococcus furiosis*
P. syringae	*Pseudomonas syringae*
®	Registered
Rbcs2	RubisCO Small Subunit 2
RFU	Relative Fluorescent Units
rGFP	Recombinant Green Fluorescent Protein
rpm	Revolutions Per Minute
S. mutans	*Streptococcus mutans*
scFvs	Single Chain Antibody Variable-Region Fragments
Ser^{65}	Serine, 65^{th} Amino Acid Residue
TAG 72	Tumor Associated Glycoprotein 72

TAP	Tris Acetate Phosphate
Tyr^{66}	Tyrosine, 66^{th} Amino Acid Residue
™	Trademark
U	International Units
µF	Micro Farad, Unit Of Capacitance
µg	Microgram
µl	Microliter
µm	Micrometer
µM	Micro molar
µmol	Micromole
UNC	University of North Carolina
V_L	Variable Light Chain of Immunoglobulin
V_H	Variable Heavy Chain of Immunoglobulin

Abstract

BIOINFORMATIC ANALYSIS OF A MAMMALIAN BIP GENE FOR INSERTION INTO GREEN ALGAE AND COMPARISON OF ITS POSSIBLE EFFECTS ON THE SYNTHESIS OF A MAMMALIAN ANTIBODY

By Katrina Patricia Ghazanfar, PhD

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, Microbiology and Immunology with concentration in Molecular Biology and Genetics at Virginia Commonwealth University.

Virginia Commonwealth University, 2004

Major Director: Dr. Fang-Sheng Wu,
Associate Professor, Department of Microbiology and Immunology

This dissertation describes a study utilizing bioinformatics to analyze homologues of a molecular chaperone, glucose-regulated protein 78 (g*rp* 78), also known as BiP. The selected homologous proteins originate from organisms of infinitely diverse genera. Comparisons of protein sequence yielded the first clues of a common ancestry among these proteins. Furthermore, protein molecular weights, isoelectric points, N-terminal amino acids and half-lives of a known homolog and a non-homologous protein were examined. Additionally, electroporation, a state-of-the-art plasmid insertion technique, was explored

using *Chlamydomonas reinhardtii*, a green alga, as the recipient of a parent plasmid, pSP124S. Distinctive hypertonic solutions and three separate field strengths were used in the plasmolysis of the cell wall of *C. reinhardtii* and subsequent electroporation, respectively. The number of transformants was tallied to evaluate which electroporation condition would yield the most transformed colonies.

We had two discrete hypotheses: 1) that a structurally and functionally similar protein to glucose-regulated protein 78 exists across a wide spectrum of organisms and 2) that *Chlamydomonas reinhardtii* could be successfully transformed with pSP124S under certain electroporation conditions.

The bioinformatics investigation revealed that analogous proteins to Human GRP 78 existed in *Mus musculus* (mouse), *Rattus norvegicus* (rat), *Gallus domesticus* (chicken), *Gallus domesticus* (chicken), *Mesocricetus auratus* (golden hamster), *Bos taurus* (cow), *Xenopus laevis* (frog), and *Spinacia oleracea* (spinach). Moreover, these homologous proteins more likely have a common evolutionary origin.

Additionally, we discovered that alteration of the hypertonic plasmolysis solution as well as electroporation field strength revealed differing rates of transformed colonies in *C. reinhardtii.* Using sucrose, sorbitol, ultrapure water, and mannitol with three unique field strengths, led to the discovery that sucrose was the best hypertonic solution to use to achieve the highest transformation efficiency rate in conjunction with a field strength comprised of 10 uF capacitance and a voltage of 2.5 kV/cm.

BACKGROUND

A. Bioinformatics and Molecular Biology

The term, "Bioinformatics," was coined by a Dutch theoretical biologist, Paulien Hogeweg and her colleague, Ben Hesper over three decades ago (**Hogeweg and Hesper, 1978**) . Essentially the term was applied to the mathematical, namely statistical, and computational approaches which were used to systematically collect and logically assemble information on innumerable biological processes (**Hogeweg and Hesper, 1984**). Later, the remarkable growth of shared databases and algorithms allowed these information technologies to be applied to the field of molecular biology. This rapid advancement enabled researchers to solve problems pertaining to little known biological processes and to analyze an immense amount of biological data in a relatively short period of time.

Bioinformatics allows researchers to readily align and compare multiple deoxyribonucleic acid (DNA) and protein sequences. Integration of information technology and molecular biology enables scientists to map and analyze sequences of DNA and proteins with the simple click of a keyboard. Databases advantageously allow investigators to map genes, predict gene expression, uncover genome construction, and measure biodiversity using phylogenetic trees.

Comparably for proteins, the information is just as ample. Researchers can quickly ascertain three-dimensional (3-D) protein structures, predict hydrophilic or hydrophobic components, and protein-protein interactions using bioinformatics (**Dandekar et al, 1998;**

Enright et al, 1999; and Marcotte et al, 1999). Analysis of phylogenetic trees can aid in the discovery of homologous proteins (**Pazos and Valencia, 2001**) and identify proteins that may have evolved simultaneously (**Pellegrini, 1999 and Tan, 2004**).

Data can be input and analyzed using several formats. The most well-known is the FASTA format. In this format, non-DNA sequences have been removed from the final sequence. The European Molecular Biology Laboratory (EMBL), GenBank, and DNA Data Bank of Japan (DDBJ), are world-renowned databases that comprise the International Nucleotide Sequence Database Collaboration. This partnership allows the public sharing of genetic information and manipulation and updates its latest versions every other month. EMBL and GenBank produce information in formats which still contains non-DNA sequences that are removed prior to analyses. Transferring between FASTA, EMBL, and GenBank formats takes mere seconds.

The most commonly used bioinformatic tool is the Basic Local Alignment Search Tool, (BLAST). BLAST enables investigators to rapidly compare sequences, either amino acids or nucleotides and is fifty times faster than dynamic programming, a proven method of using simpler steps to solve complex problems (**Altschul et al, 1997**). BLAST is most useful in the discovery of new genes. When a novel gene is revealed, its DNA sequence can be queried to see if a similar gene is known and carried by another organism, typically human. Similarity in sequence would suggest an ancestral connection and analogous function (**Altschul et al, 1990**).

The multiple sequence alignment tools give the percentage of identical amino acids, similar amino acids and the percentage of gaps. BLAST database makes use of the Blocks

of amino acid Substitution Matrix 62 (BLOSUM 62). This matrix considers every possible substitution and every possible identity of amino acids and assigns a score based on the frequencies of each as observed in alignments of related proteins. The protein building blocks that are identical are assigned the most positive scores. Substitutions that were observed to occur more frequently are also assigned positive scores; those occurring less often or rarely are assigned negative scores. This process is referred to as the compositional matrix adjust (**Altschul et al, 2005**).

B. Genetically Modified Organisms - Bacteria

Slightly more than 25 years ago, genetic engineering birthed the first recombinant bacteria. This genetically modified organism (GMO), also known as a genetically engineered organism (GEO), was the product of transferring genes from a species of *Salmonella* to produce a transgenic strain of *Escherichia coli* (**Cohen, 1973**). Upon learning of the modified microbe, concerns were raised throughout the scientific community about potential risks of genetic engineering and the organisms produced from these processes. The United States government soon had oversight over the research initiative (**Berg, et al, 1975 and Federal Register, 1976**).

Over time, scientists have modified various organisms' genomes in search of more desirable traits being expressed. Microbes, plants, and animals have all been experimented with in theory of creating microbes which can manufacture agents for gene therapy and immunotherapeutic compounds, plants that can resist pests and herbicides, and animals that can produce diagnostic and pharmaceutical substances.

As microorganisms were the simplest to alter, investigators were able to manipulate their genomes without much difficulty. A strain of *Escherichia coli* (*E. coli)* was manipulated to produce a structurally and functionally identical form of human insulin (**Tof, 1994**), now known as Humulin™. *E. coli* successfully produced a form of the hormone that would not illicit an antibody response such as the bovine and porcine varieties did. However, use of microbes has its limitations. For instance, *E. coli* cannot be used to generate more complex proteins containing disulfide bonds nor those whose functionality and utility are dependent on post-translational modifications (**Lee, 1996**).

There are countless successes in modern medicine from the use of genetically-altered microbes. Sometimes it is more beneficial to remove genes to garner more desirable traits in an organism. *Pseudomonas syringae* (*P. syringae)* is a stunning example of an organism with a deleted gene that has been used to assist in the proliferation of other organisms. *P. syringae* lacks the gene that enables ice crystals to form. When it is applied to crops, the crops survive cold weather and resist the formation of frost (**Lee et al, 1995**). However, over-use of this particular strain is thought to adversely affect ice formations in clouds and thereby negatively impacting rainfall.

Streptococcus mutans, a culprit in dental caries, has also been genetically mutated to construct a strain that fails to produce lactic acid, the chemical which initiates the breakdown of the hard tooth structures, thereby leading to dental cavities (**Hillman, 2002**).

C. Genetically Modified Organisms – Plants

On the heels of transgenic microbes, experimentation led to genetically engineered plants. Mainly to increase the yield of cash crops, the genomes of plants were genetically tailored to express resistance to insects by producing *Bacillus thuringiensis* toxin (Bt toxin), and resistance to herbicides, such as glyphosate, commonly known by the trade name of Roundup™. Later, lengthening shelf-life of crops led to the creation of the FlavrSavr™ tomatoes (**Martineau, 2001**). Soon, increasing quantities of certain vitamins, minerals, and/or proteins present in crops such as rice (**Ye, 2000**) and potatoes (**Chakraborty et al, 2000**), thereby boosting the nutritional value, became the priority.

Soybeans were the first plants custom-made to concomitantly tolerate glyphosate and to express resistance to the antibiotic, kanamycin (**Hinchee et al, 1988**). Others quickly ensued. Rice, corn, potato, and sweet potato plants (**Choi, 2007**) were soon followed by sugar cane and cotton. Tobacco plants were genetically programmed to produce human growth hormone. By far, the most lucrative and most documented venture was the creation of Bt maize, a corn which was genetically adapted to kill or sicken insects that try to devour it. This feat was accomplished as a result of the Bt toxin introduction into the corn DNA from bacterial DNA. This discovery allowed for an explosion in maize corn harvests and was thought to be the resolution to famine in Third World countries. However the corn, being genetically modified, was considered tainted and definitely not suitable for human consumption. This led to the abandonment of the donated corn.

With so many advantages to reengineering plant genomes such as to maximize crop yields, improve shelf-life, enhance nutritional worth, and decrease the need for pesticides,

there were, and still remain, several detrimental aspects. Insect populations, such as the monarch larvae are negatively impacted (**Losey et al, 1999**); cotton bollworms are exhibiting less sensitivity to the pesticide primarily due to integration of the Bt toxin (**Tabashnik and Carrière, 2003**); there is a realized potential to develop "superweeds," which are plants exhibiting multiple herbicide resistance (**Gressel, 1992 and Beckie, 2004**); and possible allergens being expressed (**Leary, 1996, and USDA website, 2002)** in prior non-allergenic plants and/or foods due to gene transfers from one allergenic source to a non-allergenic source. Hypothetically, a person, who has a known allergic response to ingesting tomatoes, could unknowingly eat soybeans that were genetically altered with genes from tomatoes. Theoretically, that person could experience an allergic reaction. Of all the negative elements ascribed to genetic reengineering of plants, the most inflexible one seems to be the lack of biodiversity, which could be attributable to the lack of genetic deviation in transgenic plants.

D. Genetically Modified Organisms - Animals

Varying the genomes of animals is undeniably more difficult when compared achieving the same with microbes and plants. The recombinant gene methodology must contain not only the structural gene, but also additional sequences to allow for the correct incorporation to the host's genome and sequences to allow for the gene to actually be expressed in that particular host.

Mice are usually the unsuspecting recipients of genetic manipulations. However rats, sheep, goats, cows, horses, rabbits, etc have also been used. Genes can be inserted

(knock-in) or deleted (knock-out) using straightforward technologies such as DNA microinjection, embryonic stem cell-mediated gene transfer, and retrovirus-mediated gene transfer (**Gordon and Ruddle, 1981; Gossel et al, 1986; and Donnelly et al, 1994**).

Animals have achieved remarkable success in production of human proteins in massive quantities. These proteins can be employed as immunotherapeutics for gene therapy and vaccine therapy, diagnostic agents useful in the determining if cancer or other debilitating disease processes are present, and pharmaceutics to treat hormone or vitamin deficiencies. Human growth hormone has successfully been produced in nude rats (**Bryant et al, 2007 and Baxter et al, 2007**). Goats have been genetically coaxed to produce an anticoagulant to treat patients suffering from antithrombin deficiency and to decrease the threat of clot formation in surgical patients (**Heavey, 2009**).

Gene function has been explored through the use of genetically modified animals. Genetic disorders such as Tay-Sachs disease, sickle cell anemia, cystic fibrosis, and Huntington disease represent several diseases in which ongoing research is being performed to unearth the precise animal model to manipulate in hopes of finding cures and/or treatments for these life-altering ailments (**Persons et al, 2003; Foster et al, 2006; and Rosenecker, 2006**).

E. Molecular "Pharming" of Mammalian Proteins

As early as November 1989, *Science News* printed an article on the use of plants to produce antibodies, which could be used as therapeutic and diagnostic agents. It was discovered that these Plantibodies™ were remarkably like the antibodies of an

animal's immune system and could strongly and selectively bind to only one or a few types of molecules. To encourage the production of foreign antibodies in plants, a group of scientists from the Research Institute of Scripps Clinic in La Jolla, California used a series of steps to shuttle two mouse genes, encoding an antibody molecule, into the nuclei of two different tobacco plant cells. Once in their perspective cells, the foreign genes were inserted into the genome of tobacco plant cells. Subsequently, the tobacco plant began producing the functional antibody. This process of utilizing plants to manufacture antibodies came to be known as "molecular pharming" according to the article downloaded from a *Science News* (**1989)** article entitled, "Turning Plants into Antibody Factories".

Nearly a decade or so later, it was established that a variety of proteins can be expressed in plants (**Blumenthal, 1999 and Borisjuk, 1999)** and that these proteins can retain their native properties **(van Engelen et al, 1994; Takeshi et al, 1997; Julian et al, 1998; Tacket et al, 1998; Borisjuk, 1999; Fischer et al, 1999; and Holger et al, 1999)**. Secreted mammalian proteins such as functional interleukin-2 and interleukin-4 **(Magnuson, 1998)**, high affinity monoclonal antibodies **(Julian et al, 1998)**, and human lactoferrin **(Salmon, 1998)** have been produced in transgenic plants. In addition, antigenic proteins of bacterial or viral origin have been manufactured using the natural machinery of plants. These antigenic proteins can be used to confer immunity in livestock and humans **(Tacket et al, 1998)**.

According to Bill Tuckey, "Tobacco, a plant responsible for the death of millions, is also the subject of experiments to produce antibodies, or "plantibodies", against diseases including, ironically cancer. The stakes are high, with the antibody drug market expected

to be worth some pounds 5bn [£5 billion] by 2004…." **(Tuckey, 2002)**. This is equivalent to approximately $9 billion US dollars.

Tobacco has been used to create antibodies against dental caries and colon cancer **(Daniell et al, 2001)**. *Nicotiana tobaccum*, a species of tobacco, has been genetically modified to produce a chimeric IgG-IgA antibody against a surface protein of *Streptococcus mutans* (*S. mutans*), the major etiological agent in human dental caries. This surface protein, a *S. mutans* glucosyltransferase, is used by the bacteria to attach and adhere to the tooth surface and begin the pathogenic process of tooth decay. Once this gram-positive organism is affixed to the tooth, other organisms begin forming the biofilm that leads to the formation of plaque. The antibody against the surface protein prevents colonization of the *S. mutans*, thereby preventing that initial, but crucial step of decay.

Genetically-altered *Nicotiana benthamiana*, another tobacco plant, has produced an antibody against colorectal-cancer-associated-antigen, GA733-2 **(Szala et al, 1990)**. This antibody has shown remarkable localization and has effectively destroyed cancer cells displaying that particular tumor- associated surface antigen.

A diagnostic tool to detect anti-human IgG has been created in alfalfa (**Khoudi et al, 1999**). This genetically engineered antibody is commonly used as a blood banking reagent. Prior to receiving a transfusion, the potential recipient and unit of blood unit must be tested for serum antibodies and red cell antigens, respectively. The anti-human IgG is used in the final phase of that antibody/antigen testing. Failure to detect these clinically significant proteins can lead to severe transfusion reactions and possibly cause the death of the recipient of the blood or blood products.

Researchers have discovered an agent which can be used to prevent transmission of vaginal Herpes Simplex Virus-2. Genetically manipulated soybean plants have produced a humanized antibody which has successfully prevented transmission of the incurable infection in a mouse model **(Daniell et al, 2001)**.

B Cell Non-Hodgkins Lymphoma has been successfully treated with Rituxan®, an agent produced in corn **(Davis et al, 1999)**. Rituxan®, a true molecular pharming success, targets cells expressing CD20. The CD20 antigen is a 33 – 37 kDa, non-glycosylated, transmembrane protein that is expressed on lineage B cells from the pre-B cell stage to the B cell lymphoblast stage and most malignant B cells. CD20 is not found on early B cell progenitors or plasma cells. Oligomers of CD20 form a Ca^{2+} channel and might have a function in regulating a local response during B cell activation. The first monoclonal antibody therapy approved in the United States for the treatment of cancer, Rituxan® has been used widely and studied extensively since its approval by the Food and Drug Administration in 1997 **(Davis et al, 1999)**.

The genomes of wheat and rice have been genetically altered to produce antibodies to treat cancer and detect a carcinoembryonic antigen (CEA) **(Stoger et al, 2000)**. CEA is a cell surface glycoprotein and the best characterized tumor associated antigen. The best use of CEA is as a tumor marker, especially for cancers of the gastrointestinal tract. Cancers of the pancreas, stomach, breast, lung, and certain types of thyroid and ovarian cancer will have significantly elevated CEA levels. Benign conditions, such as smoking, infections, inflammatory bowel disease, pancreatitis, cirrhosis of the liver, and some other

benign conditions may cause an increase of CEA also. Chemotherapy and radiation therapy can cause a temporary rise in CEA due to the death of tumor cells and release of CEA into the blood stream. When the CEA level is abnormally high before surgery or other treatment, it is expected to fall to within normal range following successful surgery to remove the cancerous cells. A rising CEA level indicates progression or recurrence of the cancer. In addition, levels greater than 20 ng/ml before therapy are associated with metastasized cancer. Anti-CEA antibodies are used as diagnostic and prognostic tools for *in vivo* imaging and immunotherapy **(Stoger et al, 2000)**.

Since its initial demonstration, the expression of functional antibodies in transgenic plants has been considered highly promising for potential disease control and manipulation of metabolic pathways **(van Engelen et al, 1994)**. The costs are significantly less as compared to production of these same antibodies in mammalian cell lines or livestock. Therapeutic and diagnostic agents produced in green tissue plants, such as tobacco, alfalfa, and soybeans, tend to be more advantageous due to the sheer levels of productivity. The increase in productivity is due to the ability to have several crop cuttings per year. This advantage is not realized in corn, wheat, or rice.

Fifty dollars per gram is the cost expended to purify IgA which is needed for the *S. mutans* vaccine. In a cell culture, one gram of this same antibody would require expending $1000. Production in alfalfa is generally more expensive costing $500 - $600 to produce and purify of one gram of anti-human IgG. Using a hybridoma system would cost nearly ten times that amount **(Daniell et al, 2001)**.

The biggest expenditure with production of Plantibodies™ is undoubtedly the purification phase. To eliminate the expensive purification cost, the plant-produced antibodies can be expressed in seeds of certain grains. Wheat, rice, and corn are examples of grains in which this has been accomplished. This strategy opens up the possibility of oral administration of some therapeutic antibodies without the need for purification.

F. Introducing Mammalian Proteins into Plants

There is increasing awareness of the potential value of using transgenic plant systems for the inexpensive production of high-quality mammalian proteins. These proteins can be manufactured for pharmaceutical and diagnostic purposes **(Magnuson, 1998)**. Uncovering ways to mass-produce therapeutic agents, while minimizing the cost, is paramount in the process of making medicines more affordable and readily available to the general public and especially to developing nations. Although most interest during the past 15 – 20 years has been focused on using microbial and animal cell cultures to produce biological agents **(Doran, 2000)**, production systems based in vascular plants have been studied for their usefulness in making therapeutic proteins **(Julian et al, 1998)**.

As opposed to bacterial production systems, plants have several characteristics that make them ideal systems to inexpensively manufacture mammalian proteins. Plants are easily transformed and cultivated, and are capable of carrying out post-translational modifications such as acetylation, phosphorylation, and glycosylation **(Borisjuk, 1999)**. An important advantage is the fidelity with which plants can express, fold, assemble, and process foreign proteins. Moreover, there is a potentially significant cost benefit in

growing bulk quantities of recombinant proteins in plants as opposed to bacteria, fungal, or animal based production systems **(Julian et al, 1998)**. Molecular pharming has been reported to be 10 – 50 times cheaper than *Escherichia coli* fermentation, even though overall product yield in bacteria is higher than those in plants. Foreign protein production using greenhouse-cultivated plants is considerably more expensive than with field-grown crops due to the cost of maintaining the environment. The expense for protein extraction and purification appears to be equivalent when comparing greenhouse cultivated plants and their field-grown counterparts **(Doran, 2000 and Kusnadi et al, 1997)**.

G. Increasing Yield of Mammalian Proteins in Plants

Initially, plants were genetically modified with the simple goal in mind to increase the overall crop yield. Genes for resistance to specific herbicides, insects, and viruses were introduced into several species of plants to increase their production. Resistance to the herbicide, glyphosate, was a true breakthrough in the genetic engineering field. This herbicide is sold under the trade name of "Roundup™". Agricultural fields are sprayed with this herbicide and the roots of plants absorb the glyphosate from the soil. Glyphosate blocks the production of a key enzyme, 5-enolpyruvyl shikimate 3-phosphate synthase (EPSPS). Glyphosate inhibits EPSPS which is required for plants to synthesize necessary aromatic amino acids, vitamins and lignin **(Brake and Evenson, 2004)**. Glyphosate kills plants by attacking the roots of most plants. Only those plants, which have been genetically modified and successfully expressed the gene that encodes for glyphosate resistance, can survive after exposure to the herbicide.

Some tomato plants have also been genetically engineered to express resistance to insects, which otherwise would destroy the plant completely over time. Genetically altered tomato plants expressing the gene for the protein toxin derived from *Bacillus thuringiensis* (*Bt)* are toxic to the larvae of some moth species, namely *Plutella xylostella*, the diamondback moth **(Schuler et al, 2004)**. This protein toxin has also been successfully incorporated into the genome of cotton, corn, potatoes, canola, and broccoli. Healthier crops are produced and there is no apparent harm to humans, insects other than the targeted pests, fish, or animals that may ingest the vegetables or come into contact with the cotton fibers **(Fox, 2003)**.

Viruses have also been thrust into the molecular pharming arena and have been responsible for some outstanding genetic breakthroughs. Plants have developed resistance to certain viruses after the plants' genomic material have been modified to produce an antibody against the coat protein of the infecting virus. Plants react to a viral infection in the same way as humans. Once the plants are encouraged to produce the antibody against the coat protein, the plants maintain protection against the virus, which is passed on to subsequent generations through the seeds. A successful example of this has been observed in the protection of plants from the destructive Tobacco Mosaic Virus **(Asurmendi et al, 2004)**. Tobacco and tomato plants are the benefactors of these breakthroughs.

Even though increasing the overall crop yield was the initial goal of creating genetically modified plants, another goal was soon realized through actual molecular pharming. If there was a way to increase the protein yield, then maybe that would decrease the number of plants needed to produce the required quantity of the protein to be used

therapeutically. From this theory, many strategies were devised for maximizing protein yield. Promoter sequences were altered to enhance their ability to drive expression of the coding sequence downstream **(Wu et al, 2001)**. For example, the cauliflower mosaic virus 35s promoter and its derivatives were and still are among the most commonly used constitutive expression promoters for plants. The 35s promoter exhibits strong, constitutive expression in many different plant tissues and organs and it has been widely used to construct expression vectors for plant genetic engineering.

Targeting of the foreign proteins to be produced in specific organelles has been studied to maximize protein yield. For some recombinant proteins, highest accumulation is achieved by retention in the endoplasmic reticulum. For example, the carboxy-terminal fusion of the Lysine-Aspartate-Glutamate-Leucine (KDEL) signal peptide to single-chain antibody variable-region fragments (scFvs) resulting in endoplasmic reticulum retention has been found to increase antibody levels 10 to 100 times compared with either extracellular secretion or expression in the cytosol **(Conrad and Fiedler, 1998; Doran, 2000; Fischer et al, 1999; and Jefferis and Lund, 1997)**.

Additionally, antibodies can be engineered to specifically localize to tumors or other cell types **(Kashmiri, 1995)**. This localization is advantageous because the constant region of an antibody can be fused with a green, yellow, red, or blue fluorescent protein or even luciferase and then used in molecular, cellular, or medical diagnostics and imaging **(Blumenthal, 1999; Gerdes and Kaether, 1996)**.

H. Advantages of Using Green Algae

Waiting for vascular plants to grow from seedlings to flowering plants can markedly increase the time needed for the generation of a therapeutic or diagnostic agent. Perhaps if another organism very similar to plants, but with a shorter life cycle, could be used, the production time could be decreased from months to weeks. Green algae appeared to be the most reasonable, reliable, and inexpensive substitute for this investigation.

Chlamydomonas reinhardtii is a unicellular, biflagellate eukaryote, which is typically oval-shaped and measures approximately 10 μm in length and 3 μm in width. These ubiquitous organisms have been discovered in soil, fresh water, oceans, and even more amazingly, in the snowcaps of mountains. The cells of these green algae contain a single chloroplast that occupies nearly 40 percent of the total cell volume **(Rochaix, 2001)**. The two anterior flagella are usually 10 μm in length and are used in a breaststroke motion to propel the algae toward or away from a particular stimulus. Surprisingly, the green algae have an "eye" that actually perceives light. The genome of this organism is 100 MB, and there are 17 chromosomes.

"Green Yeast", as this alga is commonly referred, has been used to elucidate aspects of photosynthesis as well as to study the different processes of cell wall biogenesis, flagella assembly, gametogenesis, cell cycle events, and phototaxis. *C. reinhardtii* has also been employed to investigate mating processes and nuclear-chloroplast interactions **(Rochaix, 1995)**. The advantages of using *Chlamydomonas reinhardtii* as a model eukaryotic system are numerous.

When compared to vascular plants, *C. reinhardtii* grows rapidly and has a doubling time of six to eight hours whereas vascular plants can take weeks to months to flower. The medium, in which the green algae grow, Tris-Acetate-Phosphate (TAP), is inexpensive to acquire ingredients and prepare. The algae can be easily cultured in liquid and/or solid media at room temperature, thereby eliminating the requirement for an incubator. However, growth in liquid medium is enhanced when placed on a shaker and aeration is added. Ordinary fluorescent lights are enough to support photosynthesis in this organism **(Lefebvre and Silflow, 1999)**. These eukaryotes can be grown in a minimal medium with light and CO_2 as its sole carbon source (phototrophically), without light (heterotrophically) or in an acetate-containing medium with light (mixotrophically).

Another great advantage to using the "cockroach of the algae world" is that *Chlamydomonas* can be easily transformed. Exogenous DNA can be introduced into the nuclear, mitochondrial, and chloroplast genomes **(Boyton et al, 1988; Kindle et al, 1989; Newman et al, 1991; Sodeinde and Kindle, 1993; Schnell and Lefebvre, 1993; and Davies et al, 1994; and Davies et al, 1996)**. Two direct gene transfer methods have proven to work rather well within this genus. Electroporation and vortexing with glass beads have both yielded significantly positive transformation frequencies, but the former consistently yields higher results (about 10^6 transformants per microgram of DNA) as seen in yeasts **(Manivasakam, 1993).** In another experiment, the electroporation conditions for green algae had been optimized to yield approximately 6.6 $x10^4$ transformants per microgram of DNA for a cell wall deficient mutant strain **(Shimogawara et al, 1997)**.

Chlamydomonas reinhardtii is a more sensible alternative to use in this study as compared to vascular plants. The short life cycle, easy manipulation, worry-free cultivation, and minimal laboratory costs all culminate to make this single-celled eukaryote a reasonable substitute.

I. Plasmolysis and Electroporation

The algal cell wall presents a slight obstacle for direct gene transfer into intact cells. However, plasmolysis (the drawing away of the plasmalemma, or cell membrane, from the cell wall due to cell shrinkage) can be used to enable the passage of DNA and protein molecules through the cell walls. These macromolecules accumulate between the cell wall and the plasma membrane once the cells are exposed to hypertonic solutions **(Wu and Cahoon, 1994; Wu et al, 1995)**. An electric pulse can then be applied to the cells creating more self-sealing pores in the cell membrane and cell wall.

After discovering plasmolysis before electropulsation increases the efficiency of DNA uptake, many laboratories embraced this methodology to produce transgenic cereal crops and tobacco plants. Now there are simpler procedures for development of genetically modified corn **(Sabri et al, 1996)**, tobacco **(Koscianska and Wypijewski, 2001),** rice, and wheat **(Sorokin et al, 2000)**, to name a few. This enabled the production of transient and stable genetic plant transformations to study gene expression and could possibly lead to discoveries of genetic manipulations that could increase the production and/or lifespan of these plants.

Electroporation immediately follows plasmolysis. This method utilizes a high voltage, up to 10,000 volts per centimeter, which is applied to cells for as little as one millisecond to as many as 99 milliseconds. When the target tissue experiences the high electrical pulses, plasmid DNA molecules are able to enter the cells by transient permeability of cell membrane. The pores formed in this process are self-sealing. Therefore, most of the transformed cells are located at the surface layer of the tissue, which leads to mosaic phenotypes of the electroporated embryo tissue **(Songstad et al, 1993)**. The plant cells, which have successfully taken up the plasmid DNA with the selectable marker, albeit an antibiotic resistance gene or a fluorescent fusion protein, will be recognized visually once plated on the selection media containing the antibiotic for the transformants with the former gene or under a fluorescent microscope for those transformed with the latter.

Electroporation is simple, but yet highly efficient if performed under the correct conditions, which must be determined experimentally for each type of plant tissue. Biological and physical parameters affecting electroporation must be optimized to maximize efficiency. Many cells can be transformed simultaneously in one electroporation experiment versus one cell at a time when using the microinjection technique. When utilizing the electroporation method, some factors must be taken into consideration. Energy input, electroporation buffer, and different DNA forms must be evaluated to improve efficiency **(Quecini et al, 2002)**. Energy input as combinations of electric field strengths discharged by different capacitors has been investigated. It has been determined that this sole factor has a critical influence on transgenic gene expression and achievement

of higher transformation efficiencies. A study conducted by researchers at North Dakota State University also demonstrated that linear plasmid DNA, the absence of chloride, and the presence of calcium ions in the electroporation buffer, also increased transient gene expression in the plant transformants from protoplasts, **(Tada et al, 1990)** which are plant cells whose cell walls have been removed by enzymatic digestion. However, it should be understood that the regenerated plants derived from these materials often showed various abnormalities and a reduced fertility rate due to difficulty in the mutation and/or regeneration caused by prolonged culture. Generally speaking, in protoplast regeneration, the high electric field pulse of electroporation was found to be harmful to the plant material. Therefore, a lower field strength and longer pulse time were usually adopted to produce a successful outcome of these electroporation experiments.

Pulse time and increased field strengths were also investigated. It was found that when the pulse time was about 13 milliseconds, along with increase of the field strength, the viability of plant tissue reduced gradually. Testing various field strengths and pulse times is necessary to obtain maximum efficiency and viability of the particular plant material used. Pollen, plant cells, protoplasts, tissue, proembryos, globular embryos, or mature embryos are different types of plant tissue that may be assessed to discover the optimal conditions to maximize cell viability and increase transformation efficiency.

The disadvantages of using electroporation are several: (1) the prolonged opening of the pores results in cell death; (2) the cell, even after absorbing the initial electrical pulse, may remain selective as to the molecules it lets in; (3) this method is very costly method as the sterile, plastic cuvettes can only be used once; (4) plants must be able to be

regenerated prior to applying this methodology, and this leads to the issue of plant tissue culture. That technique alone has its own disadvantages as it is challenging to maintain sterility; (5) plant tissue may require prior wounding, to remove the cell wall or make it permeable to larger macromolecules, before the tissue will uptake the DNA. The wounding can be accomplished enzymatically or mechanically.

Ke et al, **(1996)** discovered that electroporated maize embryos required prior wounding before the transient expression of two genes, beta-D-glucuronidase gene from *E. coli* and anthocyanin gene. The maize proembryos were heat-shocked, mechanically pretreated by dissection, or enzymatically pretreated prior to submitting the tissue to electroporation. In their experiments, the transformation frequency of enzymatic proembryos was approximately twice that of dissected proembryos, indicating that plasmid DNA molecules enter cells more easily after cell wall digestion by the enzyme.

Even though many varieties of full-grown adult plants can be regenerated from a single protoplast, there is a certain disadvantage to this pretreatment. When some species of plant cells are subjected to the removal of the cell wall by enzymatic treatment, they respond by synthesizing a new cell wall and eventually undergoing a series of cell divisions and developmental processes that result in the formation of a new adult plant. That adult plant is said to have been "cloned" from a single cell of the parent plant.

As for algae, the cell wall presents as much of a problem as in vascular plants. However, we hypothesized and subsequently showed (unpublished) that the algal cells can be plasmolysed with a hypertonic solution and pores can be introduced into the membrane with an electrical pulse. There are cell wall deficient strains that can be used for

electroporation; however, the viability rate of these mutants is drastically lower than the wild type algae or mutants with intact cell walls. Additionally regeneration in culture medium is also not a concern when using algae. That is the most beneficial facet of choosing green algae versus the true, green, vascular plants.

Electroporation is a fantastic technology that allows foreign DNA to be transferred into many different cells simultaneously and into a variety of sources. In plant cells there are several obstacles that must be overcome as compared to animal cells. The efficiency of this process is dependent upon the parameters of the electroporation instrumentation, the type of plant tissue used, and the electroporation buffer ingredients. Once maximum efficiency is obtained, the benefits of using this technology will far outweigh the disadvantages such as expense and sterility complications.

PROTEINS ANALYZED IN THIS INVESTIGATION

A. Glucose Regulated Protein 78

Glucose regulated protein 78 (GRP 78) is strikingly similar to the immunoglobulin heavy-chain binding protein (BiP), which associates with free immunoglobulin heavy chains in the endoplasmic reticulum (ER) until they are assembled with the light chains **(Munro and Pelham, 1986)**. The glucose regulated protein 78 is a member of the highly conserved family of heat shock proteins (Hsp70). GRP78 is a 78 kDa mammalian molecular chaperone found to localize in endoplasmic reticulum via the carboxy-terminal sequence, Lysine-Aspartate-Glutamate-Leucine (KDEL) **(Satoh 1993, and Holger et al, 1999)**. Phosphorylation of GRP78 is thought to be involved in the regulation of its binding function to immunoglobulin heavy chains **(Satoh, 1993)**.

It has been suggested that GRP78 is involved in several quality control mechanisms including recognizing, retaining, and degrading those secretory proteins within the endoplasmic reticulum that are misfolded and misassembled **(Satoh, 1993)**. GRP78 is hypothesized to be a molecular detergent, which shields the hydrophobic regions of folding proteins, and prevents them from aggregating **(Magnuson, 1998)**. It forms a stable association with some secretion-incompetent proteins, which suggests a role in retaining incorrectly folded proteins in the ER. GRP 78 is induced during times of adverse cell survival conditions such as glucose starvation, low pH, and hypoxia **(Mote et al, 1998 and Song et al, 2001)**.

There are innumerable journal articles written on this important molecular chaperone's purported functions in yeast, mammalian cells, transgenic plants, and particularly in human cells, while the cells are undergoing some physiological stress condition. Translocation of secretory proteins in yeast has been blocked when there is a loss of functional BiP/GRP78 **(Vogel et al, 1990)**. When BiP is upregulated in transgenic plants, the plants are more tolerant in drought conditions and during germination, these genetically altered plants are still able to tolerate tunicamycin, a glycosylation inhibitor **(Alvim et al, 2001)**. In mammalian cells the induction of GRP78 coincides with the G1 cell cycle arrest. In stressed cells, the epidermal growth factor receptor (EGFR) is underglycosylated and forms a more stable complex with GRP78 as compared to the mature form. The underglycosylated EGFR could not be translocated to the cell surface. This resulted in the epidermal growth factor's inability to induce the expression of cyclin D3, a G1 cyclin **(Cai et al, 1998)**. Overexpression of GRP 78 is also seen in malignant human breast lesions, which is primarily due to hypoxic conditions, low glucose, and low pH found in these tumors **(Fernandez et al, 2000)**. More recently, a study in China revealed that glucose regulated protein 78 along with another molecular chaperone, glucose regulated protein 94, could be used as prognostic indicators in gastric carcinomas. Both proteins were upregulated due to glucose starvation and the amount of expression of both correlated directly with tumor size, the degree of tumor metastasis especially when the lymphatic system was involved, and age of the tumor. Marked expression objectively indicated aggressiveness of the tumor and a poor clinical outcome for patients affected with gastric carcinomas (**Zheng et al, 2008**).

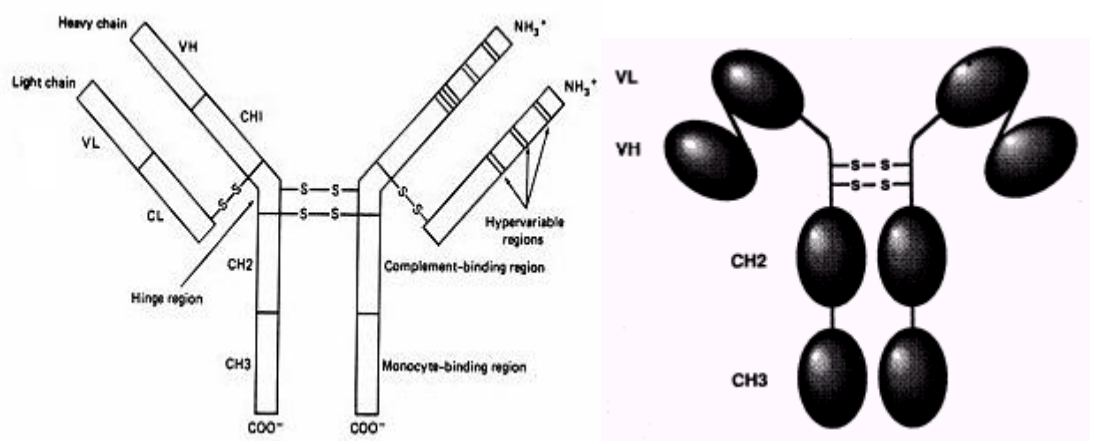
We hypothesize that the presence of GRP78 in the *Chlamydomonas reinhardtii* cells will aid the mammalian protein production in two ways. First, GRP78 may facilitate the association of the two chains of the antibody molecule. Secondly, GRP78 may prevent the plant from recognizing and degrading the foreign antibody, thereby helping to maximize its yield. With such assistance from GRP78, it is theorized that the antibody will be formed into its natural configuration as if it were made in mammalian cells, and this will result in normally folded, functional, and therapeutically valuable antibody. Plasmid pA78H is the source of the *grp 78* gene. See a plasmid map of pA78H in **Appendix A.**

B. Humanized Anti-Carcinoma Antibody 49

Humanized anti-carcinoma antibody 49 (HuCC49) is a derivative of CC49, a murine monoclonal antibody that is known to react with the Tumor-Associated glycoprotein 72 (TAG-72). TAG-72 is a human pancarcinoma antigen **(Kashmiri, 1995)**. CC49 has shown excellent tumor localization in clinical trials as it targets human colon carcinoma xenografts rather efficiently. The ability to reduce the growth of the xenografted tumors in mice also displays its tumor localization ability **(Kashmiri, 1995)**. HuCC49 is a single chain, humanized hybrid antibody created by grafting the mouse CC49 hypervariable regions onto the variable light and variable heavy frameworks of the human monoclonal antibodies LEN and 21/28'CL **(Kashmiri, 1995)**. A deleted constant 1 region and a linker peptide that connect the variable light (V_L), variable heavy (V_H), constant region 2, and constant region 3 characterize single chain antibodies. In **Figure 1**, the

differences in a typical IgG molecule (left) and a single chain IgG molecule, are depicted. The single chain IgG molecule has a deleted constant region one and a linker peptide that connects the V_L and V_H regions (**Figure 1**).

Figure 1 Typical IgG molecule (left) and Single chain IgG molecule (right). The single chain IgG molecule has a deleted constant region one (C_H1) and a linker peptide that connect the V_L and V_H regions.



www.bact.wisc.edu/ Bact303/IgG.jpeg Shu, C-F., et al, *Immunology*, 1993

A hybrid antibody was created to circumvent several problems involved in antibody-mediated therapy. First, HuCC49 reduces the human anti-mouse antibody (HAMA) response directed against CC49. Second, in humans, the immunogenic reaction to CC49 makes repeated treatments less effective due to more rapid clearance from serum, and CC49 also may elicit an allergic reaction in humans **(Kashmiri, 1995 and Kashmiri et al, 2001)**. HuCC49 was also created to bypass the difficulties in transfection and assembly of the heavy and light chains into a functional immunoglobulin, since *in vivo* the

genes coding the various domains of an antibody reside at different sites. This complicates the expression of an exogenously transferred complement of antibody genes, because the domains of the antibody must be coordinately expressed and assembled. Thus, the genes encoding the variable and constant regions of HuCC49 have been fused to form a single gene. The single gene construct approach provides a way of generating an immunoglobulin-like molecule that retains the specificity, binding properties, and catalytic activity of wild type antibodies **(Shu et al, 1993)**. The ability to express the heavy and light chains within the same cell due to the single gene construct is by far the most important property of *hucc49*. This property eliminates the need to design another experiment to incorporate and express both chains in a single algal cell. See **Appendix B** for map of plasmid containing this gene.

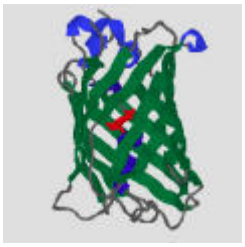
C. Green Fluorescent Protein

Green fluorescent protein (GFP) is a 27 kDa monomer protein, which autocatalytically forms a fluorescent pigment in the absence of additional proteins, substrates, or co-factors. This spontaneously fluorescent protein is isolated from coelenterates, such as the Pacific jellyfish, *Aequoria victoria* **(Morin and Hastings, 1971)** or from the sea pansy, *Renilla reniformis*. It has been expressed in bacteria, yeast, slime mold, plants, drosophila, zebrafish, and in mammalian cells. As a noninvasive fluorescent marker in living cells, it allows for a wide range of applications where it may function as a cell lineage tracer, reporter of gene expression, or as a measure of protein-protein

interactions. Its role is to transfer energy from the protein, aequorin, which is a blue chemiluminescent, into green fluorescent light **(Ward, 1979)**.

GFP is comprised of 238 amino acids. Its wild-type absorbance/excitation peak is at 395 nm with a minor peak at 475 nm. The emission peak is at 508 nm. The protein is in the shape of a cylinder, comprising 11 strands of β-sheet with an alpha helix inside and short helical segments on the ends of the cylinder (**Figure 2**).

Figure 2 Computerized depiction of the structure of the green fluorescent protein. There are 11 β-sheet strands comprising the barrel of the protein and an alpha helix inside with short helical segments at the ends of the cylinder.


www.glue.umd.edu/ ~nsw/ench482/gfp.gif

The fluorophores are protected inside the cylinder (red structure in the figure above). The structure of the fluorophore is consistent with the formation of aromatic systems made up of Tyr^{66} with reduction of its carbon-carbon bond coupled with cyclization of the neighboring glycine and serine residues. Studies of recombinant GFP expression in *E. coli* led to the discovery of the rapid cyclization between Ser^{65} and Gly^{67}, which forms the green fluorescent color when excited with blue light **(Heim et al, 1994)**.

Combinatorial mutagenesis suggests that the Gly^{67} is definitely required for formation of the fluorophores **(Delagrave et al, 1995)**.

Highly specific intracellular localization of this bioluminescent protein has been demonstrated in the nucleus, mitochondria **(Rizzuto et al, 1996)**, secretory pathway **(Kaether and Gerdes, 1995)**, plasma membrane **(Marshall et al, 1995)** and cytoskeleton **(Kahana et al, 1995)**. Visualization of this protein can be achieved repeatedly via fusions both to whole proteins and individual targeting sequences. GFP has an enormous flexibility as a noninvasive marker in living cells and this property allows for numerous other applications such as a cell lineage tracer, reporter of gene expression, and as a potential measure of protein-protein interactions **(Mitra et al, 1996)**.

One notable disadvantage concerning the use of the green fluorescent protein is that it is thermosensitive. The yield of fluorescently active protein decreases at temperatures greater than 30º C **(Lim et al, 1995)**. However, once produced GFP is quite thermostable. See **Appendix B** for map of plasmid p35S-49-GFP containing this gene.

D. Ble^R Gene

The *ble* gene originated from the *Actinomycetes* species, *Streptoalloteichus hindustanus* and encodes for a rather small protein of about 14 kDa and approximately 355 amino acids. It encodes for resistance to the drug, tallysomycin and related antibiotics including zeomycin, bleomycin, pepleomycin, and phleomycin. These are glycopeptide antibiotics, which act by perturbing the plasma membrane and also by binding to DNA,

cleaving it, thereby causing cell death. These drugs "complexed with metal ions such as copper and iron, intercalate the DNA base pairs and then catalyze the reduction of molecular oxygen to free radicals that can break DNA strands and inhibit further DNA synthesis," according to **Enrenfeld et al, 1987**.

Bleomycin has proven to be a good candidate for treatment of cancer especially human liver cancer cells *in vitro* and in xenografts in nude mice. In clinical trials, bleomycin has shown to be effective against cells not in the cell cycle and is most toxic to cells within the G2 phase of the cell cycle. The manufacturers of each of these different versions of this drug have outlined the pros and cons of using only their product. For this investigation, our laboratory used Zeocin™ produced by Invitrogen (catalog # R250-01). The disadvantage to using this form of the drug is that it is irreversibly denatured in high and low pH or in the presence of a weak oxidant. Therefore, it is usually added to media at a neutral pH.

Successful transformants, whether mammalian, bacterial, or algal, will express the BLE protein. This protein prevents the breakdown of DNA by reversibly binding to the antibiotic with a very strong affinity and consequently preventing its cleavage of DNA **(Umezawa, 1976)**.

ORIGINS of the PROTEINS ANALYZED in THIS STUDY

A. p35s-49-GFP and 35s Promoter

The 35s promoter is active in a large number of plant species, including both monocots and dicots, but generally gives higher levels of transcription in dicots. The regulatory elements, which reside within the region from –343 to –46 with the transcriptional start site designated at plus (+) 1, are responsible for the strong activity of this promoter within a wide spectrum of tissues and organs of evolutionary diverse species **(Kyung-Tae et al, 1996)**.

This plasmid, p35s-49-GFP, was constructed by a previous student, Scott Taylor in Dr. Fang-Sheng Wu's laboratory. The *hucc49* anticarcinoma antibody gene is fused to the reporter gene, *gfp*. This fusion protein fluoresces when it is introduced into plant cells **(Taylor, 2001)**. See **Table 1** for information on plasmid origin and the **Appendix B** for p35S-49-GFP plasmid map.

B. pSP124S and RBCS2 Promoter

The productivity of plants is governed by the efficiency with which they use their resources of light, water, nitrogen, and phosphate. That efficiency depends considerably on the effectiveness of the plants' CO_2-fixing enzyme, D-ribulose-1,5-bisphosphate

carboxylase-oxygenase, also called RubisCO **(Morell et al, 1992)**. There are two major types of RubisCO, which are divided phylogenetically. RubisCO Form I, found in plants, algae, some bacteria, and certain dinoflagellates, is a hexadecameric protein composed of eight large subunits, which are 50 - 55 kDa in size and bear the catalytic sites. Additionally, there are eight 12 - 18 kDa small subunits.

RubisCO Form II is strictly found in some bacteria and dinoflagellates and has only large subunits that differ in degrees of oligomerization. There are no small subunits in Form II **(Roy et al, 2000)**. Within Form I RubisCOs, there is a further evolutionary divergence between the "green" subclass found in bacteria, cyanobacteria, green algae, and higher plants and the "red" subclass found in bacteria and non-green algae **(Read and Tabita, 1994; Delwiche and Palmer, 1996; Horken and Tabita, 1999)**. In green algae and higher plants, the small subunits are nucleus-encoded.

The *rbcS* gene family encodes the small subunit of ribulose-1,5-bisphosphate carboxylase-oxygenase. The plasmid pSP124S contains an rbcS2 promoter and terminator. The rbcS2 promoter drives the transgenic expression in *Chlamydomonas reinhardtii*. This enzyme is known to be a strong, constitutive promoter in algae. This plasmid also has the *ble* gene as the selectable marker (**Figure 3**). This figure shows the partial plasmid map of pSP124 RubisCO cassette containing the *ble* gene. This is the portion of plasmid that will be used to introduce the genes of interest. This plasmid was provided to Dr. Wu by Dr. Don Weeks of the University of Nebraska **(Table1)**. See **Appendix C** for cartoon

depicting the vector which later underwent transgenic manipulation and formed the pSP124S plasmid.

Figure 3 Illustration of the RubisCO cassette of plasmid pSP124S. RubisCO cassette contains the *ble* gene that confers for resistance to Zeocin. Cartoon adapted from Saul Purton website at www.ucl.ac.uk/ biology/prg/ble1.jpg

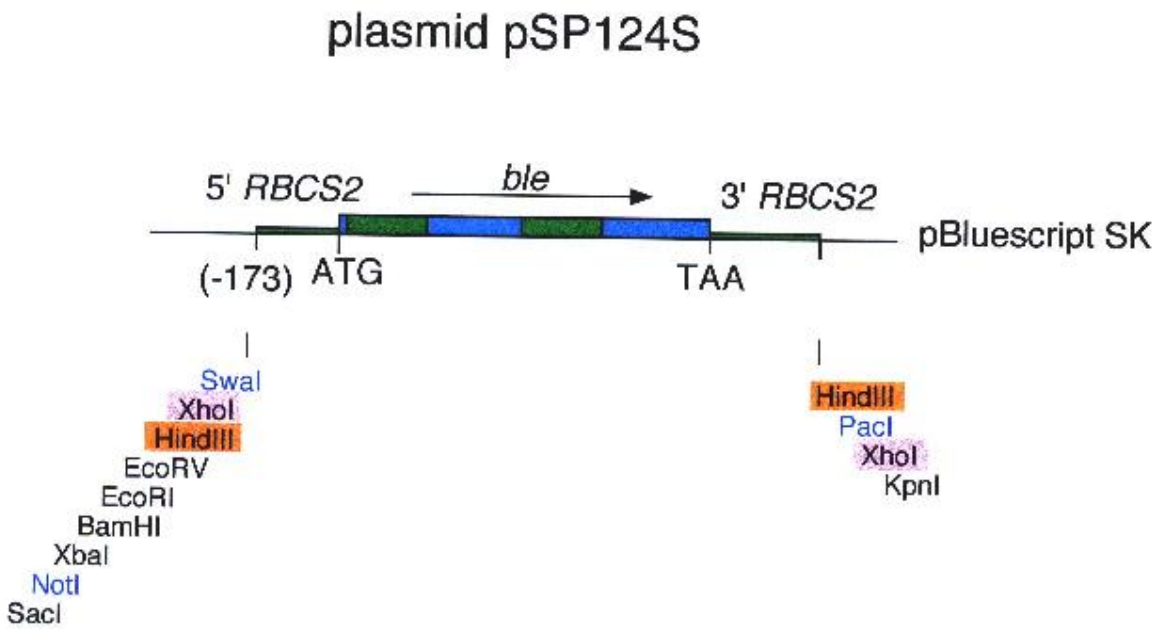


Table 1. Source of Proteins Analyzed

Plasmid Reconstructed in Dr. Wu's Lab	Original Plasmid	Source of Original Plasmid	Gene(s) Inserted	Promoter
p35S-49-GFP	p35-GFP	Jan Sheen, Harvard University	hucc49-gfp	35S
pA78H	pA8H	Academia Sinica, Taiwan	grp 78	AMY8

METHODS and MATERIALS

A. The National Center for Biotechnology Information (NCBI) Database

The Basic Local Alignment Search Tool (BLAST) located on the National Center for Biotechnology Information (NCBI) website was utilized to retrieve DNA sequences and protein sequences in FASTA format for eleven (11) potentially homologous Glucose Regulated Protein 78 (GRP 78), Heat Shock Protein 5 (HSP 5), HSP 70, and HSP 7C.

The eleven homologues were randomly selected from eukaryotic organisms. Evaluated sequence were from *Mus musculus*, the common house mouse; *Rattus norvegicus*, the Norway brown rat that was originally native to China; *Plasmodium falciparum*, a protozoan parasite that causes a deadly form of malaria in humans; *Aspergillus fumigatus*, a common fungal species to cause disease in immunocompromised patients; *Gallus domesticus*, the common domestic chicken; *Entamoeba histolytica*, a protozoan parasite that causes amoebic dysentery; *Drosophilia melanogaster*, the common fruit fly found primarily in the vicinity of unripened or rotted fruits; *Saccharomyces cerevisiae*, a species of budding yeast used in brewing and baking; *Mesocricetus auratus*, generally known as the golden hamster or Syrian hamster; *Bos Taurus*, the domestic cow; and lastly, *Xenopus laevis*, the South African clawed frog .

Each protein was paired with the human GRP 78 protein sequence. The protein sequences were then input into the query window and "BLASTed" to determine if the proteins were indeed homologues and if the protein potentially shared a similar ancestor; thereby concluding that the protein performed a similar function in the particular eukaryote from which it was derived.

B. Assessing Homology Using a Phylogeny Tree

Biodiversity can be measured through the use of phylogeny trees. These graphics provide the reviewer with a quick snapshot of how closely chosen organisms are related. The shorter the distance between organisms, the more closely those organisms are related. The greater the distance between organisms, the less likely those organisms are related. Once the sequences of the proteins of interest are retrieved from a database, phylogeny trees are very simple to generate. The HomoloGene function on the National Center for Biotechnology Information (NCBI) website is very user-friendly and the process to inputting the data is rather unsophisticated. The FASTA formats of the amino acid sequences for human Glucose Regulated Protein 78 and for each potential homologue were entered into the database.

C. CLC Main Workbench 5.1

CLC bio (http://www.clcbio.com/), a relatively young company, opened up its databases to the public in the Summer of 2005. For now, it allows free access without the hassle of registering on the site. A wealth of information is awaiting discovery. The site is

rather user friendly and not demanding of an inordinate amount of time to process queries. Through trial and error, it was discovered the best way to get the largest pay load in the shortest duration, one should run the complete protein report. The complete protein reports collate an abundance of information, far too much to use in this study, but yet good information to be familiar with for future investigations.

This bioinformatics database allows for the rapid analysis of both protein and DNA sequences. The overwhelming quantity of data provided for each protein includes the sequence information, half-life estimation, weight, isoelectric point, atomic composition focusing on hydrogen, carbon, nitrogen, oxygen and sulfur, and number of hydrophilic and hydrophobic regions and approximate locations. The count of charged residues, an amino acid residue table and histogram, number of di-peptides and how many of each combination present, are the other categories of information available on the website. The electrical charge as a function of pH is also obtainable along with secondary structures and location regions. Lastly, the database offers researchers information on the protein family to which each sequence belongs.

For the purpose of this investigation and using this bioinformatics database, human GRP 78 will be compared to one homologous protein, GRP 78 from *Mesocricetus auratus* (golden hamster) and a non-homologous protein, GRP 78 from *Plasmodium falciparum* (malarial parasite). Weight, isoelectric point, half-life, and atomic composition features will be compared and contrasted. Proteins most alike will have similarity in each of the observed characteristics. Conversely and theoretically, those proteins which are vastly different and unrelated will be markedly dissimilar.

D. Media and Growth Requirements for Algae

The wild type strain of *Chlamydomonas reinhardtii*, cc125 mt+, was used in this study and was grown mixotrophically on solid and liquid Tris-Acetate-Phosphate (TAP) media **(Table 2)**. The 100 ml liquid cultures grew under 200 μmol photons/m^s sec from ordinary fluorescent tube lights as the sole light source. The flasks were agitated on a gyratory shaker at a speed of 120 rpm at 26.5º C without aeration.

Table 2. Stock solutions for Preparation of TAP medium

Solution A	
Component	**for 500 ml**
NH_4Cl	20 g
$MgSO_4 \cdot 7H_2O$	5 g
$CaCl_2 \cdot 2H_2O$	2.5 g
Phosphate Buffer II	
Component	**For 100 ml**
K_2HPO_4	10.8 g
KH_2PO_4	5.6

Stock Solution	**For 1 L**
1M Tris base	20 ml
Phosphate Buffer II	1 ml
Hutner's Trace Metals	1 ml
Solution A	10 ml
Glacial acetic acid (pH to 7.0)	1 ml

E. Electroporation Techniques Used for Algae

Dr. Fang-Sheng Wu developed this electroporation methodology (unpublished). One hundred microliters (100 μL) of cc125 mt+ cells were harvested by centrifugation for seven (7) minutes at 2500 rpm in the Hermle Z320 centrifuge. The cells were resuspended in five (5) ml of TAP medium giving a final density of 2 – 8 x 10^8 cells/ml. 100 μL of cells were dispensed into a 24-well plate utilizing the 1st, 2nd, 5th, and 6th columns. The 3rd and 4th columns were left empty as it was noted that the cells did not spin down properly for the removal of the supernatant. In Row A: 2.0 M sucrose was added for an overall molarity of 0.4 M; in Row B, 2.0 M sorbitol, for final molarity of 0.4 M; in Row C autoclaved, ultra pure water; in Row D, 1.2 M Mannitol, for a overall molarity of 0.4 M. Prior to the additions of the sugar solutions to the algae, plasmid DNA was mixed with the sugars for overall concentration of DNA concentration of 100μg/ml. The hypertonic solutions allowed for the plasmolysis of the algae. After a four (4) minute incubation at room temperature, the 24-well plate was vacuum filtrated for another four (4) minutes and then reopened in the hood. The plate was then placed on ice for five (5) minutes prior to performance of electroporation.

Column 1 was electroporated with the BioRad Gene Pulser electroporation instrument under the following conditions: For Column 1, the capacitance, the property of an electric conductor that characterizes its ability to store an electric charge, was set at 10 μF, voltage of 2.5 kV/cm; Column 2 – 25 μF, 2.0 kV/cm; Column 5 was not electroporated and no plasmid added (negative control); Column 6 – 50 μF, 0.8 kV/cm. **Table 3** outlines the electroporation parameters used in this study.

Table 3. Electroporation Parameters Used for Green Algae Experiments

10 µF, 2.5 kV/cm	25 µF, 2.0 kV/cm		Negative Control	50 µF, 0.8 kV/cm	
A1 Sucrose 0.4M→	**A2**	Not Used	Not Used	**A5**	**A6**
uB1 Sorbitol 0.4M→	**B2**	Not Used	Not Used	**B5**	**B6**
C1 Water→	**C2**	Not Used	Not Used	**C5**	**C6**
D1 Mannitol 0.4M→	**D2**	Not Used	Not Used	**D5**	**D6**

Each well received two pulses approximately two seconds apart and lasting for anywhere from 2 – 51 milliseconds. After electroporation, the cells were allowed to remain on the ice for another five (5) minutes. Then the algal cells were incubated at room temperature for five (5) minutes. Every five (5) minute interval after then, 200µL of TAP were added to each well, including the wells in Column 5. This was done until 1 ml of TAP had been added. After the final five (5) minute incubation at room temperature, the 24-well plate was covered and centrifuged at 3000 rpm for seven (7) minutes in the Beckman Coulter Allegra™ 21R centrifuge. The supernatant was then removed under a laminar flow. A final volume of 600 µL of TAP per well was added to resuspend the

electroporated algae. The plate was incubated on a shaker overnight under previously described conditions.

After 18 – 24 hours had elapsed, the algae were spread on plates with TAP and plates with TAPZ (10 μg/ml of Zeocin™) with the starch embedding technique. Transformants were visible within 5 to 6 days.

F. Starch-Embedding Method

It is well known that transformed algae, especially the cell wall deficient strains, will have a higher plating efficiency if starch is used. Therefore, corn starch (10g) was washed sequentially with ultra pure water and then with 70 % ethanol. The washed starch was stored in 75 % ethanol to prevent bacterial contamination. Before each experiment, the ethanol was replaced with TAP medium by repeated centrifugations and resuspensions. The starch was finally resuspended to 20% (w/v) in TAP medium and polyethylene glycol (PEG) 8000 to 0.4% (w/v). PEG facilitates smooth and even spreading of the starch over the plate **(Shimogawara et al, 1997)**.

One milliliter (1ml) of starch was dispensed in the middle of each culture plate (TAP and TAPZ). The TAP plates were used as a control to ensure that the algae were not being killed during the electroporation. The TAPZ plates would help identify the positive transformants, the ones which had successfully taken up the DNA and expressed the BLE protein. Ten microliters (10 μL) of algae were dispensed onto TAP plates along with the starch. Two hundred microliters (200 μL) of algae were pipetted onto the TAPZ plates.

The plates were initially incubated in decreased light overnight, and then moved to the normal light conditions for the remaining growth period.

RESULTS

A. The National Center for Biotechnology Information (NCBI) Database

The BLAST bioinformatics tool on NCBI database was employed to rapidly compare the FASTA protein sequences for several proteins thought to be homologous to human Glucose Regulated Protein 78 (GRP 78). Four factors were readily considered in the final determination: length of protein sequences as compared to the human GRP 78; identity percentages; percentages of gaps; and overall alignment score. The following results were seen when human GRP 78 was paired with 11 potential homologues procured from a wide variety of eukaryotic organisms: *Mus musculus*, mouse; *Rattus norvegicus*, rat; *Plasmodium falciparum*, malarial parasite; *Aspergillus fumigatus*, fungi; *Gallus domesticus*, chicken; *Entamoeba histolytica*, parasitic amoeba; *Drosophilia melanogaster*, fruit fly; *Saccharomyces cerevisiae*, budding yeast; *Mesocricetus auratus*, golden hamster; *Bos taurus*, cow; *Xenopus laevis*, clawed frog, and finally, *Spinacia oleracea*, spinach, (**Figures 5 to Figure 16**).

Additionally the BLAST database was employed to construct a phylogenetic tree as a means to verify the numbers from the computational comparison (**Figure 17**). The phylogenetic tree verified the conclusions reached in the BLAST queries: the higher the identical amino acid number, the closer the other GRP 78 proteins were to the human GRP 78 on the phylogenic tree.

Table 4. Lists eukaryotic organisms used in comparison to human Glucose Regulated Protein 78. The Subject ID was assigned by BLAST and the length of the protein was determined from the FASTA amino acid sequence inserted in the program and retrieved from the NCBI database. The greater variation in number of amino acids the less likely the proteins will be homologous. The GRP78 from the malarial parasite, *Plasmodium falciparum*, is only 43% the length of the human GRP 78 which is 654 amino acids long.

Subject ID	Description	Length
52959	gi|2506545|sp|P20029.3|GRP78_*Mus musculus* (mouse)	655
52960	gi|25742763|ref|NP_037215.1| HSP 5 *Rattus norvegicus* (rat)	654
52961	gi|121573|sp|P12794.1|GRP78_*Plasmodium falciparum* (malaria)	279
52962	gi|70989035|ref|XP_749367.1| HSP70 chaperone *Aspergillus fumigatus (*fungi)	570
52963	gi|4033392|sp|Q90593.1|GRP78_*Gallus domesticus* (chicken)	652
52964	gi|67474975|ref|XP_653218.1| HSP70 family *Entamoeba histolytica*	660
52965	gi|55584057|sp|P29844.2|HSP7C_*Drosophilia melanogaster* (fruit fly)	656
52966	gi|121575|sp|P16474.1|GRP78_ *Saccharomyces cerevisiae* (yeast)	682
52967	gi|121570|sp|P07823.1|GRP78_*Mesocricetus auratus* (golden hamster)	654
52968	gi|122144501|sp|Q0VCX2.1|GRP78_*Bos taurus* (cow)	655
52969	gi|4033394|sp|Q91883.1|GRP78_*Xenopus laevis* (frog)	658
21051	gi|3913786|sp|Q42434.1|GRP78_*Spinacia oleracea* (spinach)	668

Figure 4. Below is the graphic color key for alignment scores that is generated upon using BLAST. The red lines represent the same organisms from **Table 1** in the same order. Notice the third line in the diagram represents the GRP78 extracted from the malarial parasite, *Plasmodium falciparum*. The alignment score is very low at the amine (NH_2) end, but improves dramatically at the carboxyl terminus. The scores for the other sequences were significantly higher suggesting, at first glance, a high probability of protein homology and similar evolutionary origin.

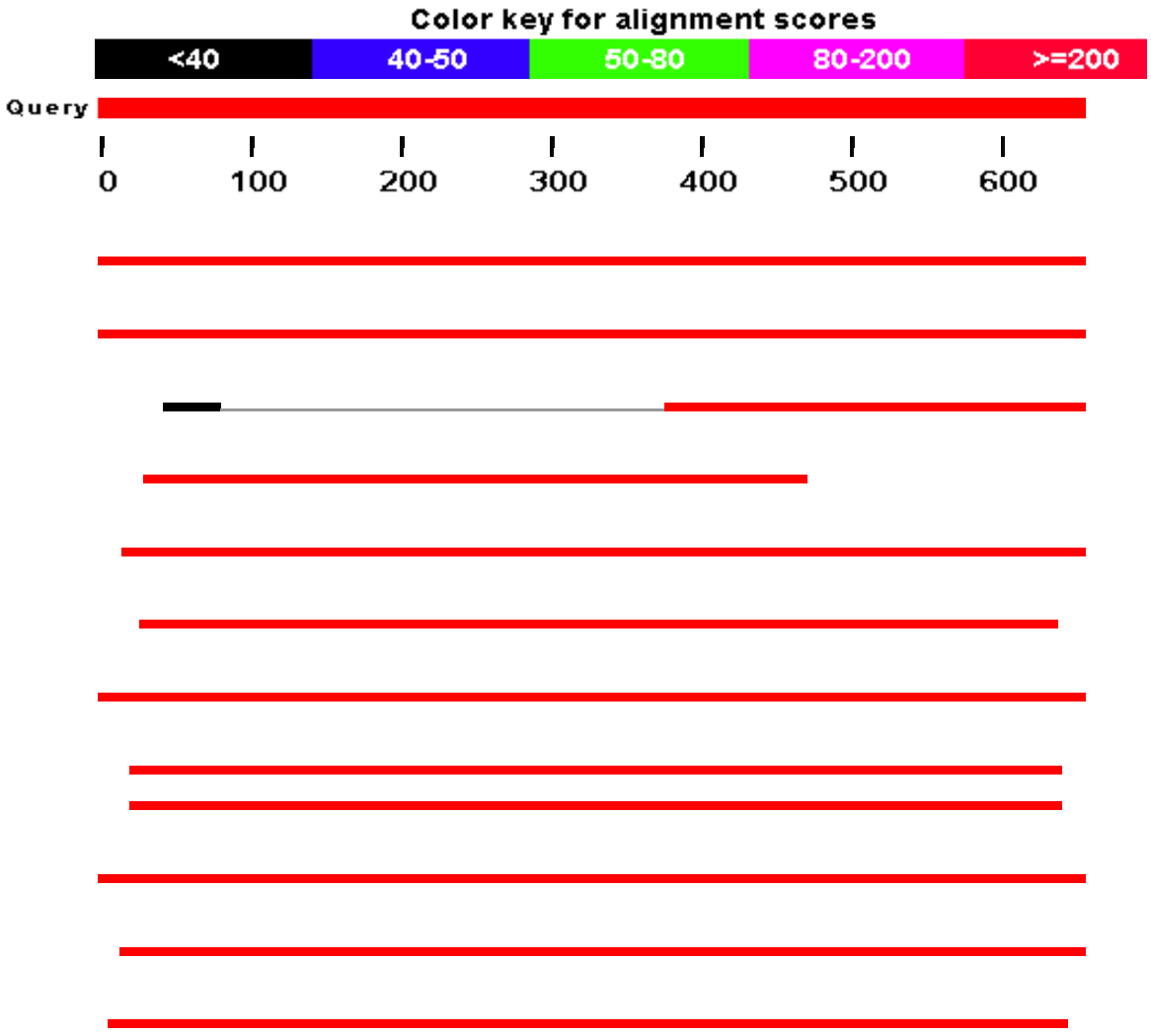


Table 5. Data retrieved after inputting the protein sequences in BLAST in the FASTA format. The scores were computed using the computational matrix adjust method. The higher the score, the increased probability that the proteins are homologous to the comparison protein, human GRP 78.

Sequences producing significant alignments:	Score (Bits)
lcl|52959 gi|2506545|sp|P20029.3|GRP78_ *Mus musculus* (mouse)	1310
lcl|52960 gi|25742763|ref|NP_037215.1| HSP5 *Rattus norvegicus* (rat)	1308
lcl|52961 gi|121573|sp|P12794.1|GRP78_ *Plasmodium falciparum* (malaria)	353
lcl|52962 gi|70989035|ref|XP_749367.1| Hsp70 chaperone *Aspergillus fumigatus* (fungi)	217
lcl|52963 gi|4033392|sp|Q90593.1|GRP78_*Gallus domesticus* (chicken)	1274
lcl|52964 gi|67474975|ref|XP_653218.1| HSP70 family *Entamoeba histolytica*	748
lcl|52965 gi|55584057|sp|P29844.2|HSP7C_*Drosophilia melanogaster* (fruit fly)	1061
lcl|52966 gi|121575|sp|P16474.1|GRP78_*Saccharomyces cerevisiae* (yeast)	859
lcl|52967 gi|121570|sp|P07823.1|GRP78_*Mesocricetus auratus* (golden hamster)	1312
lcl|52968 gi|122144501|sp|Q0VCX2.1|GRP78_*Bos taurus* (cow)	1290
lcl|52969 gi|4033394|sp|Q91883.1|GRP78_*Xenopus laevis* (frog)	1214
lcl|21051 gi|3913786|sp|42434.1|GRP78_*Spinacia oleracea* (spinach)	922

Figure 5. Protein sequences of Human GRP 78 and GRP 78 from *Mus musculus* (mouse) were aligned through the NCBI BLAST database. The percent of identical amino acids is 98% and there were no gaps in the sequences to adjust for a better alignment. The score is very high which is indicative of a similarity. From these data, it is concluded that these two proteins are homologous and very likely originated from the same ancestral lines.

```
>lcl|52959 gi|2506545|sp|P20029.3|GRP78_ Mus musculus (mouse)
Length=655
 Score = 1310 bits (3391),  Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 645/654 (98%), Positives = 649/654 (99%), Gaps = 0/654 (0%)

Query  1    MKLSLVAAMLLLLSAARAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNR  60
            MK ++VAA LLLL A RAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNR
Sbjct  2    MKFTVVAAALLLLGAVRAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNR  61

Query  61   ITPSYVAFTPEGERLIGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVV  120
            ITPSYVAFTPEGERLIGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVV
Sbjct  62   ITPSYVAFTPEGERLIGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVV  121

Query  121  EKKTKPYIQVDIGGGQTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQ  180
            EKKTKPYIQVDIGGGQTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQ
Sbjct  122  EKKTKPYIQVDIGGGQTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQ  181

Query  181  RQATKDAGTIAGLNVMRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNG  240
            RQATKDAGTIAGLNVMRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNG
Sbjct  182  RQATKDAGTIAGLNVMRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNG  241

Query  241  VFEVVATNGDTHLGGEDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALS  300
            VFEVVATNGDTHLGGEDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALS
Sbjct  242  VFEVVATNGDTHLGGEDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALS  301

Query  301  SQHQARIEIESFYEGEDFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIV  360
            SQHQARIEIESF+EGEDFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIV
Sbjct  302  SQHQARIEIESFFEGEDFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIV  361

Query  361  LVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVC  420
            LVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVC
Sbjct  362  LVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVC  421

Query  421  PLTLGIETVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLG  480
            PLTLGIETVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLG
Sbjct  422  PLTLGIETVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLG  481

Query  481  TFDLTGIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIER  540
            TFDLTGIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIER
Sbjct  482  TFDLTGIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIER  541

Query  541  MVNDAEKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEE  600
            MVNDAEKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEE
Sbjct  542  MVNDAEKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEE  601

Query  601  KIEWLESHQDADIEDFKAKKKELEEIVQPIISKLYGSAGPPPTGEEDTAEKDEL  654
            KIEWLESHQDADIEDFKAKKKELEEIVQPIISKLYGS GPPPTGEEDT+EKDEL
Sbjct  602  KIEWLESHQDADIEDFKAKKKELEEIVQPIISKLYGSGGPPPTGEEDTSEKDEL  655
```

Figure 6. Protein sequences of Human GRP 78 and Heat Shock Protein 5 from *Rattus norvegicus* (rat) were aligned through the NCBI BLAST database. The percent of identical amino acids is 98% and there were no gaps in the sequences to adjust for a better alignment. Again, the score was very high. From these data, it is concluded that these two proteins are homologous and very likely originated from the same ancestor.

```
>lcl|52960 gi|25742763|ref|NP_037215.1| HSP 5 Rattus norvegicus (rat)
Length=654
 Score = 1308 bits (3385),  Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 644/654 (98%), Positives = 648/654 (99%), Gaps = 0/654 (0%)

Query  1    MKLSLVAAMLLLLSAARAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNR  60
            MK ++VAA LLLL A RAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNR
Sbjct  1    MKFTVVAAALLLLCAVRAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNR  60

Query  61   ITPSYVAFTPEGERLIGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVV  120
            ITPSYVAFTPEGERLIGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVV
Sbjct  61   ITPSYVAFTPEGERLIGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVV  120

Query  121  EKKTKPYIQVDIGGGQTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQ  180
            EKKTKPYIQVDIGGGQTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQ
Sbjct  121  EKKTKPYIQVDIGGGQTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQ  180

Query  181  RQATKDAGTIAGLNVMRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNG  240
            RQATKDAGTIAGLNVMRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNG
Sbjct  181  RQATKDAGTIAGLNVMRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNG  240

Query  241  VFEVVATNGDTHLGGEDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALS  300
            VFEVVATNGDTHLGGEDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALS
Sbjct  241  VFEVVATNGDTHLGGEDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALS  300

Query  301  SQHQARIEIESFYEGEDFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIV  360
            SQHQARIEIESF+EGEDFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIV
Sbjct  301  SQHQARIEIESFFEGEDFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIV  360

Query  361  LVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVC  420
            LVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVC
Sbjct  361  LVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVC  420

Query  421  PLTLGIETVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLG  480
            PLTLGIETVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLG
Sbjct  421  PLTLGIETVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLG  480

Query  481  TFDLTGIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIER  540
            TFDLTGIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIER
Sbjct  481  TFDLTGIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIER  540

Query  541  MVNDAEKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEE  600
            MVNDAEKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLS EDKETMEKAVEE
Sbjct  541  MVNDAEKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSPEDKETMEKAVEE  600

Query  601  KIEWLESHQDADIEDFKAKKKELEEIVQPIISKLYGSAGPPPTGEEDTAEKDEL  654
            KIEWLESHQDADIEDFKAKKKELEEIVQPIISKLYGS GPPPTGEEDT+EKDEL
Sbjct  601  KIEWLESHQDADIEDFKAKKKELEEIVQPIISKLYGSGGPPPTGEEDTSEKDEL  654
```

Figure 7. Protein sequences of Human GRP 78 and GRP 78 from *Plasmodium falciparum* (malaria causing parasite) were aligned through the NCBI BLAST database. The variation in sequence length was remarkable, indicating the likelihood of these proteins being similar in composition and function was minute. The percent of identical amino acids is rather low (62%) and there was 1% gap in the sequences to achieve a better alignment. The score was very low in comparison to the other proteins BLASTed. From these data, it is concluded that these two proteins are not homologous and it is very unlikely that they have a common beginning.

```
>lcl|52961 gi|121573|sp|P12794.1|GRP78_Plasmodium falciparum (malaria)
Length=279

 Score =  353 bits (907),  Expect = 1e-101, Method: Compositional matrix adjust.
 Identities = 176/281 (62%), Positives = 221/281 (78%), Gaps = 5/281 (1%)

Query  377  EFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVCPLTLGIETVGGVMTKL  436
            EFFNGKEP+RGINPDEAVAYGAA+QAG++ G++   D+VLLDV PLTLGIETVGG+MT+L
Sbjct  1    EFFNGKEPNRGINPDEAVAYGAAIQAGIILGEE-LQDVVLLDVTPLTLGIETVGGIMTQL  59

Query  437  IPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLGTFDLTGIPPAPRGVPQ  496
            I RNTV+PTKKSQ FST  DNQP V I+V+EGER LTKDNHLLG F+L+GIPPA RGVP+
Sbjct  60   IKRNTVIPTKKSQTFSTYQDNQPAVLIQVFEGERALTKDNHLLGKFELSGIPPAQRGVPK  119

Query  497  IEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIERMVNDAEKFAEEDKKLK  556
            IEVTF +D NGIL V AEDKGTG    ITITND+ RL+ E+IE+M+NDAEKFA+EDK L+
Sbjct  120  IEVTFTVDKNGILHVEAEDKGTGKSRGITITNDKGRLSKEQIEKMINDAEKFADEDKNLR  179

Query  557  ERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEEKIEWLESHQDADIEDF  616
            E+++ +N L++Y  S+K  + DK+KL  K+  EDK T+  AV++  +WL ++ +AD E
Sbjct  180  EKVEAKNNLDNYIQSMKATVEDKDKLADKIEKEDKNTILSAVKDAEDWLNNNSNADSEAL  239

Query  617  KAKKKELEEIVQPIISKLYGSAG---PPPTGEEDTAEKDEL  654
            K K K+LE + QPII KLYG  G   P P+G+ED  + DEL
Sbjct  240  KQKLKDLEAVCQPIIVKLYGQPGGPSPQPSGDEDV-DSDEL  279
```

Figure 8. Protein sequences of Human GRP 78 and Heat Shock Protein 70 chaperone from *Aspergillus fumigatus* (an opportunistic fungi) were aligned through the NCBI BLAST database. The variation in sequence length was slightly remarkable, indicating the likelihood of these proteins being similar in composition and function was not probable. The percent of identical amino acids was extremely low (33%) and there was 15% gap in the sequences to attempt to achieve a better alignment. The score was very low in comparison to the other proteins BLASTed. From this data, it is concluded that these two proteins are not homologous and did not evolve from a common source. This is not surprising since a comparable protein with a similar function could not be located for *Aspergillus fumigatus*. *Note: The NCBI BLAST analyses will not yield results for amino acids 473 – 570 because there were no significant similarities found when compared to Human GRP 78. This protein did not have a HDEL/ KDEL C-terminus as seen in plants or mammalian homologues of GRP 78.

```
>lcl|52962 gi|70989035|ref|XP_749367.1| Hsp70 chaperone Aspergillus
fumigatus (fungi)
Length=570

 Score =  217 bits (552),  Expect = 1e-60, Method: Compositional matrix adjust.
 Identities = 155/466 (33%), Positives = 244/466 (52%), Gaps = 36/466 (7%)

Query  31   VGIDLGTTYSCVGVFKNGRVEIIANDQGNRITPSYVAFTPEGERLIGDAAKNQLTSNPEN  90
            +GI  G + S +     G+ E+IAN++G+R  P+ +++  +GE   G  AK QL  NP+N
Sbjct  17   IGISFGNSSSSIARLTPGKAEVIANEEGDRQIPTVLSYI-DGEEYHGTQAKAQLVRNPQN  75

Query  91   TVFDAKRLIGRTWN--DPS-VQQDIKFLPFKVVEKKTKPYIQVDIGGGQTKTFAPEEISA  147
            TV   +  +G+ +   DP+   Q     P +V    T  +   D       T    EI+
Sbjct  76   TVAYFRDYVGKNFKSIDPTPCHQSAH--PQQV--DSTVAFTIRDTASETPNTVTVSEITT  131

Query  148  MVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQRQATKDAGTIAGLNVMRIINEPTAAAI  207
              L ++K++A  YLGK V  AV+TVP  F D QR+A   A   AGL V+++I+EP AA +
Sbjct  132  RHLRRLKQSASDYLGKDVNAAVITVPTDFTDVQREALIAAAGAAGLEVLQLIHEPVAAVL  191

Query  208  AYGLDKRE----GEKNILVFDLGGGTFDVSLLTIDNGVFEVVATNGDTHLGGEDFDQRVM  263
            AY  D R      +K ++V D GG   D +++    G++ ++AT  D  LGG   DQ V+
Sbjct  192  AY--DARPEATVTDKLVVVADFGGTRSDAAVIACRGGMYTILATAHDYELGGASLDQIVI  249

Query  264  EHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALSSQHQARIEIESFYEGEDFSETLT  323
            +HF K + KK   D R++ R + KL+ E E  +RALS    A + IES  +G DFS T+
Sbjct  250  DHFAKEFIKKHKTDPRENARGLAKLKLEGEATRRALSLGTNASLSIESLADGIDFSSTIN  309

Query  324  RAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQQLVKEFFNGK-  382
            R ++E L+  +F    + +++V++ ++L   DIDE++  GG++  PKI QL +  F+ K
Sbjct  310  RTRYELLSGKVFAQFTRLIEQVVQKAELDVLDIDEVIFSGGTSHTPKIAQLARNMFSEKT  369

Query  383  ---EPS---RGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVCPL---------TLGIE  427
                PS     INP E    GAA+QA ++  + D  D +  ++ P+          +G+E
Sbjct  370  KILAPSTSASAINPSELAPRGAAIQASLIQ-EFDKED-IEQNIHPMVTATPHLRNAIGVE  427

Query  428  TVGGVMTKLIP---RNTVVPTKKSQIFSTASDNQPTVTIKVYEGER  470
             V G   +  P     T +P ++   +S   D    V ++V EG R
Sbjct  428  FVHGETVEFKPLLNAETALPARRVAQYSAPKDGG-DVLVRVCEGVR  472
```

Figure 9. Protein sequences of Human GRP 78 and GRP 78 from *Gallus domesticus* (domestic chicken) were aligned through the NCBI BLAST database. The variation in sequence length was unremarkable. The percent of identical amino acids was extremely high (97%) and there was no added gaps in the sequences. The score was very high which led to the conclusion that the proteins are homologous and probably were derived from a similar or quite possibly the same evolutionary source.

```
>lcl|52963 gi|4033392|sp|Q90593.1|GRP78_Gallus domesticus (domestic chicken)
Length=652
 Score = 1274 bits (3298),  Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 623/638 (97%), Positives = 635/638 (99%), Gaps = 0/638 (0%)

Query  17   RAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNRITPSYVAFTPEGERLI  76
            RA++E+KKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNRITPSYVAFTPEGERLI
Sbjct  15   RADDEEKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNRITPSYVAFTPEGERLI  74

Query  77   GDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVVEKKTKPYIQVDIGGGQ  136
            GDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIK+LPFKVVEKK KP+IQVD+GGGQ
Sbjct  75   GDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKYLPFKVVEKKAKPHIQVDVGGGQ  134

Query  137  TKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQRQATKDAGTIAGLNVM  196
            TKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQRQATKDAGTIAGLNVM
Sbjct  135  TKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQRQATKDAGTIAGLNVM  194

Query  197  RIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNGVFEVVATNGDTHLGGE  256
            RIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNGVFEVVATNGDTHLGGE
Sbjct  195  RIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNGVFEVVATNGDTHLGGE  254

Query  257  DFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALSSQHQARIEIESFYEGE  316
            DFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALSSQHQARIEIESF+EGE
Sbjct  255  DFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALSSQHQARIEIESFFEGE  314

Query  317  DFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQQLVK  376
            DFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQQLVK
Sbjct  315  DFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQQLVK  374

Query  377  EFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVCPLTLGIETVGGVMTKL  436
            EFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVCPLTLGIETVGGVMTKL
Sbjct  375  EFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVCPLTLGIETVGGVMTKL  434

Query  437  IPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLGTFDLTGIPPAPRGVPQ  496
            IPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLGTFDLTGIPPAPRGVPQ
Sbjct  435  IPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLGTFDLTGIPPAPRGVPQ  494

Query  497  IEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIERMVNDAEKFAEEDKKLK  556
            IEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIERMVNDAEKFAEEDKKLK
Sbjct  495  IEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIERMVNDAEKFAEEDKKLK  554

Query  557  ERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEEKIEWLESHQDADIEDF  616
            ERID RNELESYAYSLKNQIGDKEKLGGKLSSEDKET+EKAVEEKIEWLESHQDADIEDF
Sbjct  555  ERIDARNELESYAYSLKNQIGDKEKLGGKLSSEDKETIEKAVEEKIEWLESHQDADIEDF  614

Query  617  KAKKKELEEIVQPIISKLYGSAGPPPTGEEDTAEKDEL  654
            K+KKKELEE+VQPI+SKLYGSAGPPPTGEE+ AEKDEL
Sbjct  615  KSKKKELEEVVQPIVSKLYGSAGPPPTGEEEAAEKDEL  652
```

Figure 10. Protein sequences of Human GRP 78 and a protein from the HSP 70 family extracted from *Entamoeba histolytica* (amoeba) were aligned through the NCBI BLAST database. The variation in sequence length was not remarkable. The percent of identical amino acids was unsurprisingly low (59%) and number of gaps in the sequence was insignificant. With such a low score, it is more than likely not homologous and could not have resulted from a similar ancestral origin. *Note: The NCBI BLAST analysis will not yield results for amino acids 625 – 660 because there were no significant similarities found when compared to Human GRP 78. This C-terminus of the HSP 70 family protein extracted from E. histolytica is KDEL.

```
>lcl|52964 gi|67474975|ref|XP_653218.1| HSP70 family Entamoeba histolytica
Length=660
 Score =  748 bits (1930),  Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 363/608 (59%), Positives = 479/608 (78%), Gaps = 2/608 (0%)

Query  29   TVVGIDLGTTYSCVGVFKNGRVEIIANDQGNRITPSYVAFTPEGERLIGDAAKNQLTSNP  88
             ++GIDLGTT+S VG++++  VEIIANDQGNRITPS VAFT + + L+G+AA+NQ+T NP
Sbjct  19   VIIGIDLGTTFSAVGIYRDSGVEIIANDQGNRITPSVVAFT-DHDILVGEAARNQITENP  77

Query  89   ENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVVEKKTKPYIQVDIGGGQTKTFAPEEISAM  148
             +NT+F+ KRLIGRT++D  VQ+D+   PF ++ +  KP+I+V + G + K ++PEEISAM
Sbjct  78   KNTIFEIKRLIGRTYDDKEVQRDLHIFPFNIINQDNKPFIKVTLKG-EEKIYSPEEISAM  136

Query  149  VLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQRQATKDAGTIAGLNVMRIINEPTAAAIA  208
             ++ KM +TA  YLGK+V  AV+TVPAYFNDAQRQATKDAGTIAGL V+RI+NEPTAA++A
Sbjct  137  IIHKMAKTASDYLGKEVKKAVITVPAYFNDAQRQATKDAGTIAGLEVLRIVNEPTAASMA  196

Query  209  YGLDKREGEKNILVFDLGGGTFDVSLLTIDNGVFEVVATNGDTHLGGEDFDQRVMEHFIK  268
             +GL+  +GEK ILVFDLGGGTFDVSLL I+N VFEV+AT+GDTHLGG DFDQR+    ++
Sbjct  197  FGLNSFKGEKQILVFDLGGGTFDVSLLNIENNVFEVIATSGDTHLGGSDFDQRIALFLVE  256

Query  269  LYKKKTGKDVRKDNRAVQKLRREVEKAKRALSSQHQARIEIESFYEGEDFSETLTRAKFE  328
             + K+K  KD   + RA+ KLR+E EKAK ALSS+ Q +IEIE   EG DFS  LTRA+F
Sbjct  257  ICKRKFKKDPSDNPRAMSKLRKEAEKAKIALSSEEQTKIEIEGLMEGLDFSFVLTRARFN  316

Query  329  ELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQQLVKEFFNGKEPSRGI  388
             ELN+DLF+ T+ PV+ VL D+ L K D+DEIVLVGGSTRIPK+Q+L++EFFNGKEP++ +
Sbjct  317  ELNLDLFKKTLGPVRMVLSDAKLDKKDVDEIVLVGGSTRIPKVQELLQEFFNGKEPNKDV  376

Query  389  NPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVCPLTLGIETVGGVMTKLIPRNTVVPTKKS  448
             NPDEAVAYGAA+Q  VL+  + T D+VL+D  PLTLGI T GGVM  +IPR T VPTKKS
Sbjct  377  NPDEAVAYGAAIQGAVLNNSEGTNDVVLVDATPLTLGIMTAGGVMASIIPRGTHVPTKKS  436

Query  449  QIFSTASDNQPTVTIKVYEGERPLTKDNHLLGTFDLTGIPPAPRGVPQIEVTFEIDVNGI  508
             QIF+T +DNQ  V I+V+EGER LTKDNHLLG F L GI  APRG+P+IEVTF++DVNGI
Sbjct  437  QIFTTHADNQEQVEIQVFEGERSLTKDNHLLGKFMLEGIKRAPRGIPKIEVTFDVDVNGI  496

Query  509  LRVTAEDKGTGNKNKITITNDQNRLTPEEIERMVNDAEKFAEEDKKLKERIDTRNELESY  568
             LRV+A+DK +G K +ITIT+++ RLT E+I+RMV +A++ + ED K K+ I++RNELE+Y
Sbjct  497  LRVSAQDKKSGKKEEITITSEKGRLTEEQIQRMVKEAQERSGEDNKAKKMIESRNELENY  556

Query  569  AYSLKNQIGDKEKLGGKLSSEDKETMEKAVEEKIEWLESHQDADIEDFKAKKKELEEIVQ  628
             AY +++++ DK+KL  KL   DK+T+   V+E +++LE      IE  +   K+LE+IV
Sbjct  557  AYKVRDEVIDKDKLADKLQEGDKKTILDGVDEVLDFLEREMHPSIEKCEEMYKKLEQIVH  616
```

```
Query  629  PIISKLYG  636
            PI+ +  G
Sbjct  617  PILRRYGG  624
```

Figure 11. Protein sequences of Human GRP 78 and a purified protein, HSP 70 Cognate 3, from *Drosophilia melanogaster* (fruit fly) were aligned through the NCBI BLAST database. The variation in sequence length was not remarkable. The percent of identical amino acids was moderate (80%) and there were an insignificant number of gaps needed to achieve better sequence alignment. With a mediocre score, it is difficult to concretely conclude that these proteins are homologous and/or originated from the same or very similar source.

```
>lcl|52965 gi|55584057|sp|P29844.2|HSP7C_Drosophilia melanogaster
Length=656
 Score = 1061 bits (2745),  Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 529/657 (80%), Positives = 593/657 (90%), Gaps = 4/657 (0%)

Query  1    MKLSLVAAMLLLLSAARAEEEDKKED-VGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGN  59
            MKL ++ A++  +  +  EE+ +K+  +GTV+GIDLGTTYSCVGV+KNGRVEIIANDQGN
Sbjct  1    MKLCILLAVVAFVGLSLGEEKKEKDKELGTVIGIDLGTTYSCVGVYKNGRVEIIANDQGN  60

Query  60   RITPSYVAFTPEGERLIGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKV  119
            RITPSYVAFT +GERLIGDAAKNQLT+NPENTVFDAKRLIGR W+D +VQ DIKF PFKV
Sbjct  61   RITPSYVAFTADGERLIGDAAKNQLTTNPENTVFDAKRLIGREWSDTNVQHDIKFFPFKV  120

Query  120  VEKKTKPYIQVDIGGGQTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDA  179
            VEK +KP+I VD   G  K FAPEEISAMVL KMKETAEAYLGKKVTHAVVTVPAYFNDA
Sbjct  121  VEKNSKPHISVDTSQG-AKVFAPEEISAMVLGKMKETAEAYLGKKVTHAVVTVPAYFNDA  179

Query  180  QRQATKDAGTIAGLNVMRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDN  239
            QRQATKDAG IAGL VMRIINEPTAAAIAYGLDK+EGEKN+LVFDLGGGTFDVSLLTIDN
Sbjct  180  QRQATKDAGVIAGLQVMRIINEPTAAAIAYGLDKKEGEKNVLVFDLGGGTFDVSLLTIDN  239

Query  240  GVFEVVATNGDTHLGGEDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRAL  299
            GVFEVVATNGDTHLGGEDFDQRVM+HFIKLYKKK GKD+RKDNRAVQKLRREVEKAKRAL
Sbjct  240  GVFEVVATNGDTHLGGEDFDQRVMDHFIKLYKKKKGKDIRKDNRAVQKLRREVEKAKRAL  299

Query  300  SSQHQARIEIESFYEGEDFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEI  359
            S  HQ RIEIESF+EG+DFSETLTRAKFEELN+DLFRST+KPVQKVLED+D+ K D+ EI
Sbjct  300  SGSHQVRIEIESFFEGDDFSETLTRAKFEELNLDLFRSTLKPVQKVLEDADMNKKDVHEI  359

Query  360  VLVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDV  419
            VLVGGSTRIPK+QQLVK+FF GKEPSRGINPDEAVAYGAAVQAGVLSG+QDT  +VLLDV
Sbjct  360  VLVGGSTRIPKVQQLVKDFFGGKEPSRGINPDEAVAYGAAVQAGVLSGEQDTDAIVLLDV  419

Query  420  CPLTLGIETVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLL  479
             PLT+GIETVGGVMTKLIPRNTV+PTKKSQ+FSTASDNQ TVTI+VYEGERP+TKDNHLL
Sbjct  420  NPLTMGIETVGGVMTKLIPRNTVIPTKKSQVFSTASDNQHTVTIQVYEGERPMTKDNHLL  479

Query  480  GTFDLTGIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIE  539
            G FDLTGIPPAPRG+PQIEV+FEID NGIL+V+AEDKGTGNK KI ITNDQNRLTPE+I+
Sbjct  480  GKFDLTGIPPAPRGIPQIEVSFEIDANGILQVSAEDKGTGNKEKIVITNDQNRLTPEDID  539
```

```
Query  540  RMVNDAEKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVE  599
            RM+ DAEKFA+EDKKLKER+++RNELESYAYSLKNQIGDK+KLG KLS ++K  +E A++
Sbjct  540  RMIRDAEKFADEDKKLKERVESRNELESYAYSLKNQIGDKDKLGAKLSDDEKNKLESAID  599

Query  600  EKIEWLESHQDADIEDFKAKKKELEEIVQPIISKLYGSAG--PPPTGEEDTAEKDEL  654
            E I+WLE + DAD E++K +KK+LE IVQP+I+KLY  AG  PPP G +D   KDEL
Sbjct  600  ESIKWLEQNPDADPEEYKKQKKDLEAIVQPVIAKLYQGAGGAPPPEGGDDADLKDEL  656
```

Figure 12. Protein sequences of Human GRP 78 and a GRP 78 derived from the budding yeast, *Saccharomyces cerevisiae* were aligned through the NCBI BLAST database. The variation in sequence length was not remarkable. The percent of identical amino acids was moderately remarkable (68%) and there was a typical number of gaps added to achieve better sequence alignment. With this low score, it is concluded that the probability of these two proteins being homologous and having descent from the same source in very unlikely. *Note: The NCBI BLAST analyses will not yield results for amino acids 658 – 682 because there were no significant similarities found when compared to Human GRP 78. This protein of *S. cerevisiae* has an HDEL C-terminus.

```
>lcl|52966 gi|121575|sp|P16474.1|GRP78_Saccharomyces cerevisiae (yeast)
Length=682

 Score =  859 bits (2219),  Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 426/619 (68%), Positives = 509/619 (82%), Gaps = 7/619 (1%)

Query  22   DKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNRITPSYVAFTPEGERLIGDAAK  81
            D  E+ GTV+GIDLGTTYSCV V KNG+ EI+AN+QGNRITPSYVAFT + ERLIGDAAK
Sbjct  44   DDVENYGTVIGIDLGTTYSCVAVMKNGKTEILANEQGNRITPSYVAFTDD-ERLIGDAAK  102

Query  82   NQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVVEKKTKPYIQVDIGGGQTKTFA  141
            NQ+ +NP+NT+FD KRLIG  +ND SVQ+DIK LPF VV K  KP ++V + G + K F
Sbjct  103  NQVAANPQNTIFDIKRLIGLKYNDRSVQKDIKHLPFNVVNKDGKPAVEVSVKG-EKKVFT  161

Query  142  PEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQRQATKDAGTIAGLNVMRIINE  201
            PEEIS M+L KMK+ AE YLG KVTHAVVTVPAYFNDAQRQATKDAGTIAGLNV+RI+NE
Sbjct  162  PEEISGMILGKMKQIAEDYLGTKVTHAVVTVPAYFNDAQRQATKDAGTIAGLNVLRIVNE  221

Query  202  PTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNGVFEVVATNGDTHLGGEDFDQR  261
            PTAAAIAYGLDK + E  I+V+DLGGGTFDVSLL+I+NGVFEV AT+GDTHLGGEDFD +
Sbjct  222  PTAAAIAYGLDKSDKEHQIIVYDLGGGTFDVSLLSIENGVFEVQATSGDTHLGGEDFDYK  281

Query  262  VMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALSSQHQARIEIESFYEGEDFSET  321
            ++   IK +KKK G DV  +N+A+ KL+RE EKAKRALSSQ   RIEI+SF +G D SET
Sbjct  282  IVRQLIKAFKKKHGIDVSDNNKALAKLKREAEKAKRALSSQMSTRIEIDSFVDGIDLSET  341

Query  322  LTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQQLVKEFFNG  381
            LTRAKFEELN+DLF+ T+KPV+KVL+DS L+K D+D+IVLVGGSTRIPK+QQL++ +F+G
Sbjct  342  LTRAKFEELNLDLFKKTLKPVEKVLQDSGLEKKDVDDIVLVGGSTRIPKVQQLLESYFDG  401

Query  382  KEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVCPLTLGIETVGGVMTKLIPRNT  441
            K+ S+GINPDEAVAYGAAVQAGVLSG++   D+VLLDV  LTLGIET GGVMT LI RNT
```

```
Sbjct  402  KKASKGINPDEAVAYGAAVQAGVLSGEEGVEDIVLLDVNALTLGIETTGGVMTPLIKRNT  461
Query  442  VVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLGTFDLTGIPPAPRGVPQIEVTF  501
             +PTKKSQIFSTA DNQPTV IKVYEGER ++KDN+LLG F+LTGIPPAPRGVPQIEVTF
Sbjct  462  AIPTKKSQIFSTAVDNQPTVMIKVYEGERAMSKDNNLLGKFELTGIPPAPRGVPQIEVTF  521

Query  502  EIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIERMVNDAEKFAEEDKKLKERIDT  561
             +D NGIL+V+A DKGTG    ITITND+ RLT EEI+RMV +AEKFA ED  +K ++++
Sbjct  522  ALDANGILKVSATDKGTGKSESITITNDKGRLTQEEIDRMVEEAEKFASEDASIKAKVES  581

Query  562  RNELESYAYSLKNQI-GDKEKLGGKLSSEDKETMEKAVEEKIEWLESHQDADI-EDFKAK  619
            RN+LE+YA+SLKNQ+ GD   LG KL  EDKET+  A  + +EWL+ + +  I EDF  K
Sbjct  582  RNKLENYAHSLKNQVNGD---LGEKLEEEDKETLLDAANDVLEWLDDNFETAIAEDFDEK  638

Query  620  KKELEEIVQPIISKLYGSA  638
             + L ++  PI SKLYG A
Sbjct  639  FESLSKVAYPITSKLYGGA  657
```

Figure 13. Protein sequences of Human GRP 78 and a GRP 78 purified from *Mesocricetus auratus* (golden hamster) were aligned through the NCBI BLAST database. The variation in sequence length was not remarkable. The percent of identical amino acids was exceptionally high (98%) and there were absolutely no gaps added to achieve better sequence alignment. This exceedingly high score cements the conclusion that these two proteins are indeed homologous and have descended from a common ancestor.

```
>lcl|52967 gi|121570|sp|P07823.1|GRP78_Mesocricetus auratus (golden hamster)
Length=654
 Score = 1312 bits (3395),  Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 646/654 (98%), Positives = 649/654 (99%), Gaps = 0/654 (0%)

Query  1    MKLSLVAAMLLLLSAARAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNR  60
            MK  +VAA LLLL A RAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNR
Sbjct  1    MKFPMVAAALLLLCAVRAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNR  60

Query  61   ITPSYVAFTPEGERLIGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVV  120
            ITPSYVAFTPEGERLIGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVV
Sbjct  61   ITPSYVAFTPEGERLIGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVV  120

Query  121  EKKTKPYIQVDIGGGQTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQ  180
            EKKTKPYIQVDIGGGQTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQ
Sbjct  121  EKKTKPYIQVDIGGGQTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQ  180

Query  181  RQATKDAGTIAGLNVMRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNG  240
            RQATKDAGTIAGLNVMRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNG
Sbjct  181  RQATKDAGTIAGLNVMRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNG  240

Query  241  VFEVVATNGDTHLGGEDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALS  300
            VFEVVATNGDTHLGGEDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALS
Sbjct  241  VFEVVATNGDTHLGGEDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALS  300

Query  301  SQHQARIEIESFYEGEDFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIV  360
            SQHQARIEIESF+EGEDFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIV
Sbjct  301  SQHQARIEIESFFEGEDFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIV  360
```

```
Query  361  LVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVC  420
            LVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVC
Sbjct  361  LVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVC  420

Query  421  PLTLGIETVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLG  480
            PLTLGIETVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLG
Sbjct  421  PLTLGIETVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLG  480

Query  481  TFDLTGIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIER  540
            TFDLTGIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIER
Sbjct  481  TFDLTGIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIER  540

Query  541  MVNDAEKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEE  600
            MVNDAEKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEE
Sbjct  541  MVNDAEKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEE  600

Query  601  KIEWLESHQDADIEDFKAKKKELEEIVQPIISKLYGSAGPPPTGEEDTAEKDEL  654
            KIEWLESHQDADIEDFKAKKKELEEIVQPIISKLYGSAGPPPTGEEDT+EKDEL
Sbjct  601  KIEWLESHQDADIEDFKAKKKELEEIVQPIISKLYGSAGPPPTGEEDTSEKDEL  654
```

Figure 14. Protein sequences of Human GRP 78 and a GRP 78 extracted from *Bos taurus* (cow) were aligned through the NCBI BLAST database. The variation in sequence length was not remarkable. The percent of identical amino acids was exceptionally high (99%) and there were absolutely no gaps added to achieve better sequence alignment. This exceedingly high score strengthens the conclusion that these two proteins are homologues and were derived from a common antecedent.

```
>lcl|52968 gi|122144501|sp|Q0VCX2.1|GRP78_Bos taurus (cow)
Length=655
 Score = 1290 bits (3337),  Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 634/639 (99%), Positives = 637/639 (99%), Gaps = 0/639 (0%)

Query  16   ARAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNRITPSYVAFTPEGERL  75
            ARAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNRITPSYVAFTPEGERL
Sbjct  17   ARAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNRITPSYVAFTPEGERL  76

Query  76   IGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVVEKKTKPYIQVDIGGG  135
            IGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVVEKKTKPYIQVD+GGG
Sbjct  77   IGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVVEKKTKPYIQVDVGGG  136

Query  136  QTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQRQATKDAGTIAGLNV  195
            QTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQRQATKDAGTIAGLNV
Sbjct  137  QTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQRQATKDAGTIAGLNV  196

Query  196  MRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNGVFEVVATNGDTHLGG  255
            MRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNGVFEVVATNGDTHLGG
Sbjct  197  MRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNGVFEVVATNGDTHLGG  256

Query  256  EDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALSSQHQARIEIESFYEG  315
            EDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALSSQHQARIEIESFYEG
Sbjct  257  EDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALSSQHQARIEIESFYEG  316
```

```
Query  316  EDFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQQLV  375
            EDFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQQLV
Sbjct  317  EDFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQQLV  376

Query  376  KEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVCPLTLGIETVGGVMTK  435
            KEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVCPLTLGIETVGGVMTK
Sbjct  377  KEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVCPLTLGIETVGGVMTK  436

Query  436  LIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLGTFDLTGIPPAPRGVP  495
            LIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLGTFDLTGIPPAPRGVP
Sbjct  437  LIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLGTFDLTGIPPAPRGVP  496

Query  496  QIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIERMVNDAEKFAEEDKKL  555
            QIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIERMVNDAEKFAEEDKKL
Sbjct  497  QIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIERMVNDAEKFAEEDKKL  556

Query  556  KERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEEKIEWLESHQDADIED  615
            KERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEEKIEWLESHQDADIED
Sbjct  557  KERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEEKIEWLESHQDADIED  616

Query  616  FKAKKKELEEIVQPIISKLYGSAGPPPTGEEDTAEKDEL  654
            FKAKKKELEEIVQPIISKLYGSAGPPPT EE+ A+KDEL
Sbjct  617  FKAKKKELEEIVQPIISKLYGSAGPPPTSEEEAADKDEL  655
```

Figure 15. Protein sequences of Human GRP 78 and a GRP 78 recovered from *Xenopus laevis* (frog) were aligned through the NCBI BLAST database. The variation in sequence length was not remarkable. The percent of identical amino acids was high (94%) and there were an insignificant number of gaps added to achieve better sequence alignment. This high score lends credence to the conclusion that these two proteins are quite possibly homologues and evolved from the same ancestral beginning. *Note: The NCBI BLAST analyses will not yield results for amino acids 644 – 658 because there were no significant similarities found when compared to Human GRP 78. The C-terminal sequence is KDEL.

```
>lcl|52969 gi|4033394|sp|Q91883.1|GRP78_Xenopus laevis (frog)
Length=658
 Score = 1214 bits (3142),  Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 602/637 (94%), Positives = 623/637 (97%), Gaps = 3/637 (0%)

Query  8    AMLLLLSAA--RAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNRITPSY  65
            A++LL+SA+   ++++DKK+D+GTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNRITPSY
Sbjct  8    ALVLLVSASVFASDDDDKKDDIGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNRITPSY  67

Query  66   VAFTPEGERLIGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKVVEKKTK  125
            VAFTPEGERLIGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIK+LPFKV+EKKTK
Sbjct  68   VAFTPEGERLIGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKYLPFKVIEKKTK  127

Query  126  PYIQVDIGGGQTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQRQATK  185
            PYI+VDIG  Q KTFAPEEISAMVL KMKETAEAYLG+KVTHAVVTVPAYFNDAQRQATK
Sbjct  128  PYIEVDIGD-QMKTFAPEEISAMVLVKMKETAEAYLGRKVTHAVVTVPAYFNDAQRQATK  186
```

```
Query  186  DAGTIAGLNVMRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNGVFEVV  245
            DAGTIAGLNVMRIINEPTAAAIAYGLDK+EGEKNILVFDLGGGTFDVSLLTIDNGVFEVV
Sbjct  187  DAGTIAGLNVMRIINEPTAAAIAYGLDKKEGEKNILVFDLGGGTFDVSLLTIDNGVFEVV  246

Query  246  ATNGDTHLGGEDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRALSSQHQA  305
            ATNGDTHLGGEDFDQRVMEHFIKLYKKKTGKDVR D RAVQKLRREVEKAKRALS+QHQ+
Sbjct  247  ATNGDTHLGGEDFDQRVMEHFIKLYKKKTGKDVRADKRAVQKLRREVEKAKRALSAQHQS  306

Query  306  RIEIESFYEGEDFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGS  365
            RIEIESF+EGEDFSETLTRAKFEELNMDLFRSTMKPVQKVL+DSDLKKSDIDEIVLVGGS
Sbjct  307  RIEIESFFEGEDFSETLTRAKFEELNMDLFRSTMKPVQKVLDDSDLKKSDIDEIVLVGGS  366

Query  366  TRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVCPLTLG  425
            TRIPKIQQLVKE FNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVCPLTLG
Sbjct  367  TRIPKIQQLVKELFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVCPLTLG  426

Query  426  IETVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLGTFDLT  485
            IETVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLGTFDLT
Sbjct  427  IETVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLGTFDLT  486

Query  486  GIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIERMVNDA  545
            GIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIERMV DA
Sbjct  487  GIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIERMVTDA  546

Query  546  EKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEEKIEWL  605
            EKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKET+EKAVEEKIEWL
Sbjct  547  EKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETIEKAVEEKIEWL  606

Query  606  ESHQDADIEDFKAKKKELEEIVQPIISKLYGSAGPPP  642
            ESHQDADIEDFKAKKKELEEIVQPI+ KLYG AG PP
Sbjct  607  ESHQDADIEDFKAKKKELEEIVQPIVGKLYGGAGAPP  643
```

Figure 16. Protein sequences of Human GRP 78 and a GRP 78 isolated from *Spinacia oleracea* (spinach) were aligned through the NCBI BLAST database. The miniscule variation in sequence length was not remarkable. The percent of positive amino acids was moderately high (82%) and there were an insignificant number of gaps added to achieve better sequence alignment. This score allows the conclusion that human GRP 78 and GRP 78 of spinach may in fact have similar evolutionary beginnings.

```
>lcl|21051 gi|3913786|sp|Q42434.1|GRP78_Spinacia oleracea (spinach)
Length=668
 Score =  922 bits (2384),  Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 458/663 (69%), Positives = 550/663 (82%), Gaps = 12/663 (1%)

Query  2    KLSLVAAMLLLLSA--ARAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGN
59           + S +A  ++LL +  A    +D+   +GTV+GIDLGTTYSCVGV+K+G+VEIIANDQGN
Sbjct  8    RASSIAFGIVLLGSLFAFVSAKDEAPKLGTVIGIDLGTTYSCVGVYKDGKVEIIANDQGN
67

Query  60   RITPSYVAFTPEGERLIGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDIKFLPFKV
119          RITPS+VAFT + ERLIG+AAKNQ  +NPE T+FD KRLIGR + D  VQ+D+K +P+K+
```

```
Sbjct  68   RITPSWVAFTND-ERLIGEAAKNQAAANPERTIFDVKRLIGRKFEDKEVQKDMKLVPYKI
126

Query  120  VEKKTKPYIQVDIGGGQTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDA
179          V +  KPYIQV +  G+TK F+PEEISAM+LTKMKETAE +LGKK+  AVVTVPAYFNDA
Sbjct  127  VNRDGKPYIQVKVQEGETKVFSPEEISAMILTKMKETAETFLGKKIKDAVVTVPAYFNDA
186

Query  180  QRQATKDAGTIAGLNVMRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDN
239          QRQATKDAG IAGLNV RIINEPTAAAIAYGLDKR GEKNILVFDLGGGTFDVS+LTIDN
Sbjct  187  QRQATKDAGVIAGLNVARIINEPTAAAIAYGLDKRGGEKNILVFDLGGGTFDVSVLTIDN
246

Query  240  GVFEVVATNGDTHLGGEDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAKRAL
299          GVFEV+ATNGDTHLGGEDFDQR+ME+FIKL KKK  KD+ KDNRA+ KLRRE E+AKRAL
Sbjct  247  GVFEVLATNGDTHLGGEDFDQRLMEYFIKLIKKKHTKDISKDNRALGKLRRECERAKRAL
306

Query  300  SSQHQARIEIESFYEGEDFSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEI
359          SSQHQ R+EIES ++G DFSE LTRA+FEELN DLFR TM PV+K ++D+ L+K+ IDEI
Sbjct  307  SSQHQVRVEIESLFDGVDFSEPLTRARFEELNNDLFRKTMGPVKKAMDDAGLEKNQIDEI
366

Query  360  VLVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGD--QDTGDLVLL
417          VLVGGSTRIPK+QQL+KEFFNGKEPS+G+NPDEAVA+GAAVQ  +LSG+  ++T +++LL
Sbjct  367  VLVGGSTRIPKVQQLLKEFFNGKEPSKGVNPDEAVAFGAAVQGSILSGEGGEETKEILLL
426

Query  418  DVCPLTLGIETVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNH
477          DV PLTLGIETVGGVMTKLIPRNTV+PTKKSQ+F+T  D Q TVTI+V+EGER LTKD
Sbjct  427  DVAPLTLGIETVGGVMTKLIPRNTVIPTKKSQVFTTYQDQQTTVTIQVFEGERSLTKDCR
486

Query  478  LLGTFDLTGIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEE
537          LLG FDLTGI PAPRG PQIEVTFE+D NGIL V AEDK +G   KITITND+ RL+ EE
Sbjct  487  LLGKFDLTGIAPAPRGTPQIEVTFEVDANGILNVKAEDKASGKSEKITITNDKGRLSQEE
546

Query  538  IERMVNDAEKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKA
597          IERMV +AE+FAEEDKK+KE+ID RN LE+Y Y++KNQI D +KL  KL S++KE +E A
Sbjct  547  IERMVREAEEFAEEDKKVKEKIDARNSLETYIYNMKNQISDADKLADKLESDEKEKIEGA
606

Query  598  VEEKIEWLESHQDADIEDFKAKKKELEEIVQPIISKLYGSAGPPPTGE-----EDTAE-K
651          V+E +EWL+ +Q A+ ED+  K KE+E +  PII+ +Y  +G  P+GE     ED+ E
Sbjct  607  VKEALEWLDDNQSAEKEDYDEKLKEVEAVCNPIITAVYQRSG-GPSGESGADSEDSEEGH
665

Query  652  DEL  654
             DEL
Sbjct  666  DEL  668
```

In summary, from analyses of the data garnered from NCBI's BLAST database, we discovered which proteins were closely related to human GRP 78. This assessment was accomplished by comparing the identities percentages and the overall alignment score. This study revealed and confirmed the following proteins are homologous to, and most likely have an evolutionary link, to Human Glucose Regulated Protein 78:

1) GRP 78 from *Mus musculus* (mouse),
2) Heat Shock Protein 5 (HSP 5) from *Rattus norvegicus* (rat),
3) GRP 78 from *Gallus domesticus* (chicken);
4) GRP 78 from *Mesocricetus auratus* (golden hamster)
5) GRP 78 from *Bos taurus* (cow)
6) GRP 78 from *Xenopus laevis* (frog)
7) GRP 78 from *Spinacia oleracea* (spinach)

Contrarily and also from the analyses of the data gathered from the BLAST database, the following proteins did not appear to be homologous or have an ancestral linkage to Human Glucose Regulated Protein 78:

1) GRP 78 from *Plasmodium falciparum* (malaria)
2) Heat Shock Protein 70 chaperone from *Aspergillus fumigatus* (fungi)
3) Heat Shock Protein 70 family from *Entamoeba histolytica* (amoeba)
4) GRP78 from *Saccharomyces cerevisiae* (yeast)

Four of the selected proteins did not have results that included the C-terminal ending of either KDEL as seen in mammalian cells or HDEL as seen in plants and lower organisms. This is a limitation of the NCBI BLAST program in that if no significant similarities are found, the output ends at the last group of significantly comparable amino acids. The protein evaluation ranges were changed to try and elicit some statistical response from the program, but no manipulation of the parameters would yield the complete protein sequence analyzed. Each sequence in which this occurred, *A. fumigatus, E. histolytica, S. cerevisiae* and *X. laevis,* was verified and the C-terminus was added to the corresponding figure's annotation.

Lastly, as every investigation would have, there is one outlier or one indeterminable variable, Heat Shock Protein 7C purified from *Drosophilia melanogaster,* (fruit fly). With the other proteins evaluated in this study, it was relatively simple to decide on homology and possible relationship between the proteins. In the case of HSP7C from the fruit fly, it is virtually impossible to look at the data generated and make a definitive decision. Even though the score was moderately high for identical amino acids in the protein sequence at 80%, it could not be determined with reasonable certainty and accuracy that this particular protein was indeed homologous to human GRP 78 or shared any common ancestry.

B. Assessing Homology Using a Phylogeny Tree

A phylogeny tree was generated from the HomoloGene portion of the NCBI website. The FASTA format of the protein sequences was submitted for each of the eleven

homologous proteins. The phylogeny tree would provide a graphical representation of ancestry, versus a calculated one as in sequence alignment analyses (**Figure 17**).

Figure 17. Below is a phylogeny tree generated through the use of the HomoloGene function of the NCBI database. As suspected from the sequence alignment data performed previously, human Glucose Regulated Protein 78 is most closely related to GRP78 from *Mesocricetus auratus* (golden hamster), *Bos taurus* (cow), *Rattus norvegicus* (Norway rat), and *Mus musculus* (common mouse). Additionally, as determined from the sequence alignment data, similar proteins from *Entamoeba histolytica* and *Plasmodium falciparum* are not structurally similar and are at best, distantly related to human GRP 78.

As reported from the protein sequence alignment data, human GRP 78 is most closely related to corresponding proteins in the common house mouse, the Norway rat, the cow, and the golden hamster. Its ancestral relationship with similar proteins in chickens is hinted at from this phylogenetic tree. Unsurprisingly, human GRP 78 is not very closely related to either of the protozoan parasites, *Entamoeba histolytica* or *Plasmodium falciparum.* There appears to be no relationship whatsoever with the fungi, *Aspergillus fumigatus,* which is so far away, it almost failed to be pictured on the phylogeny tree. Its placement is at the top far right corner of the graphic representation in **Figure 17**. *Spinacia oleracea* was appended to this study and according to the bioinformatics data obtained in the BLAST analyses, spinach would be very close to *Saccharomyces cerevisiae* on the phylogeny tree.

C. CLC Main Workbench 5.1

Human GRP 78 was compared to one homologous protein, GRP 78 from *Mesocricetus auratus* (golden hamster) and one non-homologous protein, GRP 78 from *Plasmodium falciparum* (malarial parasite). The protein molecular weight, isoelectric point, half-life, atomic composition and frequency features were evaluated. Proteins most alike will have similar characteristics in each of these areas; whereas those which are vastly different will be markedly dissimilar (**Table 6**).

Table 6. Comparison and contrast of a few selected biochemical difference between human GRP78 in a protein, GRP 78 from *Mesocricetus auratus* (golden hamster), a known homologue and a non-homologous protein, GRP 78 from *Plasmodium falciparum* (malarial parasite). Human GRP 78, hamster GRP 78, and spinach GRP 78 are comparable in molecular weight, isoelectric point, N-terminal residues, and differ slightly in half-life. Whereas the non-homologous GRP 78 from the malarial parasite is very different in that it has only half of the molecular weigh of the other two compounds; 0.5 pH units different making it soluble in more acidic solutions. Human and hamster GRP 78 would have net positive charges at the isoelectric point of the Malarial GRP78. Its net charge would be zero.

	Human GRP78	**Hamster GRP78**	**Malarial GRP78**	**Spinach GRP78**
Molecular Weight	72.332 kDa	72.378 kDa	30.657 kDa	73.600 kDa
Isoelectric Point	5.31	5.31	5.00	4.76
N-Terminal Amino Acid	Methionine	Methionine	Glutamic acid	Methionine
Half-life	30 hours	30 hours	1 hour	10-20 hours

Human GRP 78, hamster GRP 78, and spinach GRP 78 are nearly identical when molecular weight, isoelctric point, N-terminal amino acid and half-life of the compounds are compared. This is what is expected with homologous proteins. If these proteins were subject to polyacrylamide gel analysis, one would expect to find them in close vicinity of each other. The opposite holds true for the malarial parasite GRP 78.

Human GRP 78 and hamster GRP 78 are nearly identical when atomic composition and atom count frequency are compared. Again, this is the expectation for homologous proteins. The more closely related the proteins are the more similar they react biochemically and biologically. On the contrary, when a comparable assessment is done with the malarial GRP 78 with the same variables, atomic composition and atom count frequency, again are numerous differentiations among the atoms reviewed. The malarial protein contains 1/3 to ½ of the atomic compositions.

Table 7. Comparison and contrast of the Atomic Composition and Count Frequency by analyzing hydrogen (H), carbon (C), nitrogen (N), oxygen (O), and sulfur (S). As expected, there are negligible differences between human GRP78 and an evolutionary relative, GRP 78 from *Mesocricetus auratus* (golden hamster). When comparing and contrasting GRP 78 from *Plasmodium falciparum* (malarial parasite), it is readily noted the difference in count and frequency of the selected atoms. Human GRP78 and hamster GRP 78 are comparable in atomic composition and count frequency. Conversely, GRP 78 from the malarial parasite is very different in that it has

ATOMIC COMPOSITION AND ATOM COUNT FREQUENCY										
Protein	**H**	**H freq**	**C**	**C freq**	**N**	**N freq**	**O**	**O freq**	**S**	**S freq**
Human GRP78	5,153	0.503	3,189	0.311	865	0.084	1,019	0.100	13	0.001
Hamster GRP 78	5,151	0.503	3,193	0.312	865	0.084	1,017	0.099	14	0.001
Malarial GRP 78	2,177	0.503	1,342	0.310	365	0.084	444	0.102	4	0.001

D. Electroporation Techniques Used for Algae

The wild type strain of *Chlamydomonas reinhardtii*, cc125 mt+, was used in this study and was successfully grown mixotrophically on solid and liquid Tris-Acetate-Phosphate (TAP) media.

Several different electroporation conditions were used in this study until the optimal conditions were discovered. Initially, we varied the hypertonic solutions and then the field strengths and only used the parent plasmid pSP124S as our electroporation DNA. After seven experiments, the results were computed **(Figures 18, 19, 20 and 21)** and we concluded that the best results without question were obtained when we used 0.4 M Sucrose in our electroporation media and had a field strength with a total capacitance of 10μF and a voltage of 2.5 kV/cm.

Figure 18. Effect of Field Strength on Transformation Efficiency in 0.4 M Sucrose. A 100-ml culture *Chlamydomonas reinhardtii* cells were centrifuged at 3000 rpm for seven minutes. The supernatant TAP was decanted and the pellet was resuspended in 5 ml of fresh TAP solution. The green algae cells (100 µL) were aliquoted into a 24-well plate. The specified concentration of sucrose was used as the hypertonic solution to produce plasmolysis in algae and allow the pSP124S plasmid DNA to enter the cells. Each aliquot of algal cells was electroporated with two pulses at the designated voltage. The results from seven experiments were totaled. At least a four-fold increase of transformed colonies was obtained with the use of sucrose for plasmolysis as compared to the other hypertonic solutions or ultra pure water.

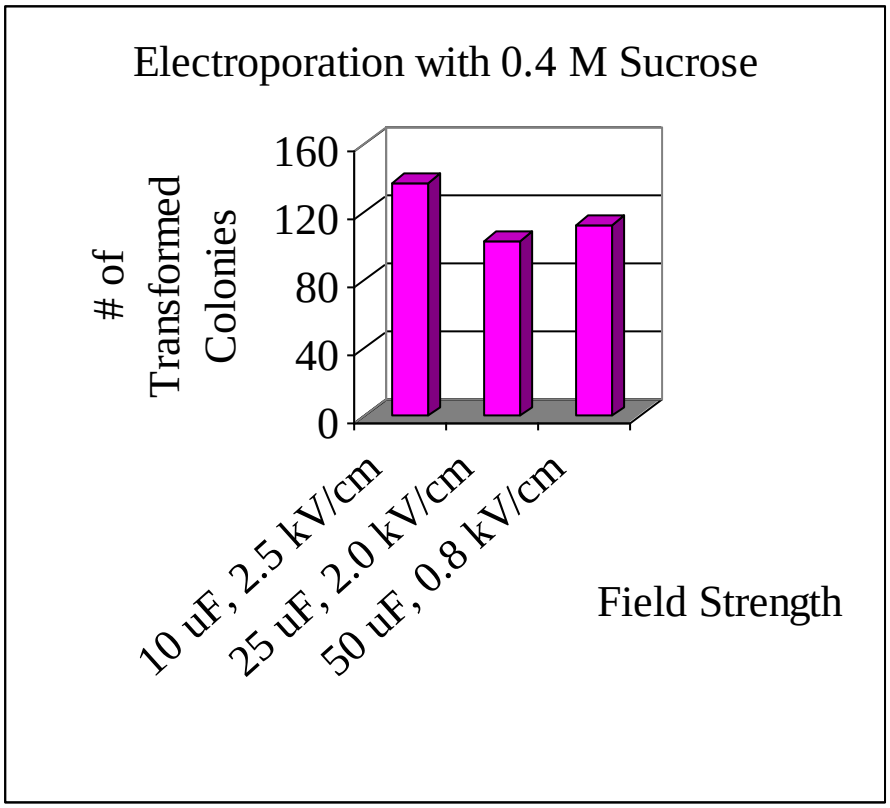


Figure 19. Effect of Field Strength on Transformation Efficiency in 0.4 M Sorbitol. The same procedure from the previous figure was used and sucrose was replaced with sorbitol. The results from seven experiments were totaled. Sorbitol is an effective plamolysis agent, however not as efficient as sucrose when transformation efficiency is compared. We obtained a total of 1/3 the number of transformed colonies as compared with the number of transformed colonies submerged in sucrose. Variation in the field strengths did prove to produce slightly higher transformation efficiency among the cells exposed to sorbitol. Overall, this sugar did not produce as many transformed colonies as sucrose.

.

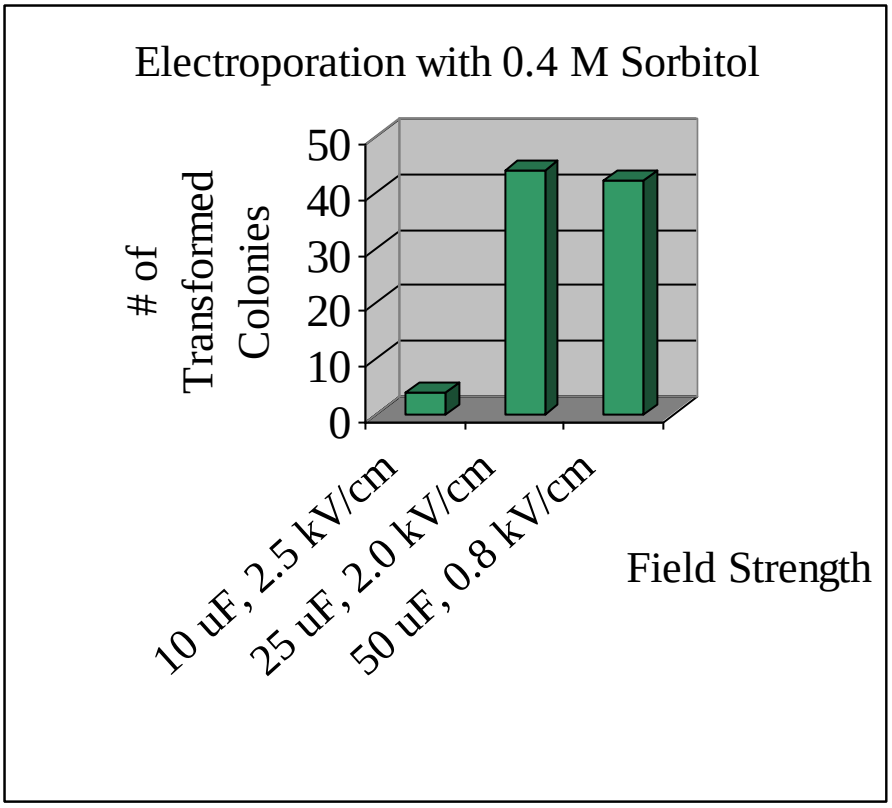


Figure 20. Effect of Field Strength on Transformation Efficiency in Ultra Pure Water. A 100-ml culture *Chlamydomonas reinhardtii* cells were centrifuged at 3000 rpm for seven minutes. The supernatant TAP was decanted and the pellet was resuspended in 5 ml of fresh TAP solution. 100 µL of algal cells were pipetted into a 24-well plate. The pSP124S plasmid DNA was mixed with ultra pure water. Each aliquot of algal cells was exposed to the indicated field strengths and electroporated with two pulses at the designated voltage. The sums of colonies from seven experiments were graphed. The use of ultra pure water allowed us to successfully transform colonies, but not nearly as many when compared to sucrose or sorbitol. It appears that the change in field strengths did not affect the number of colonies successfully transformed.

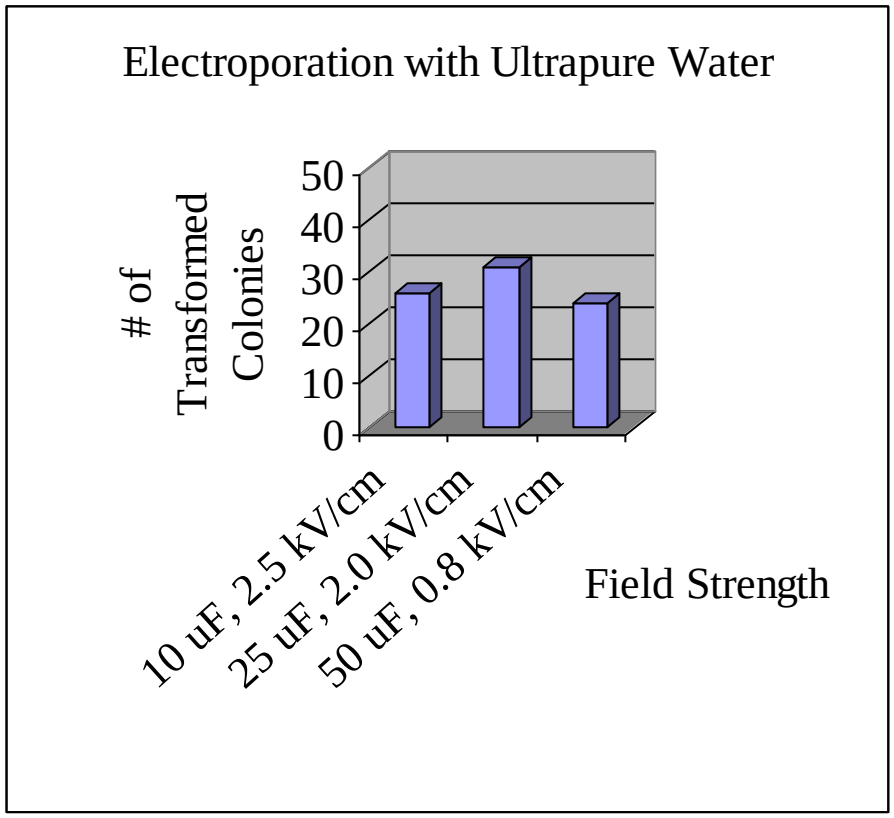
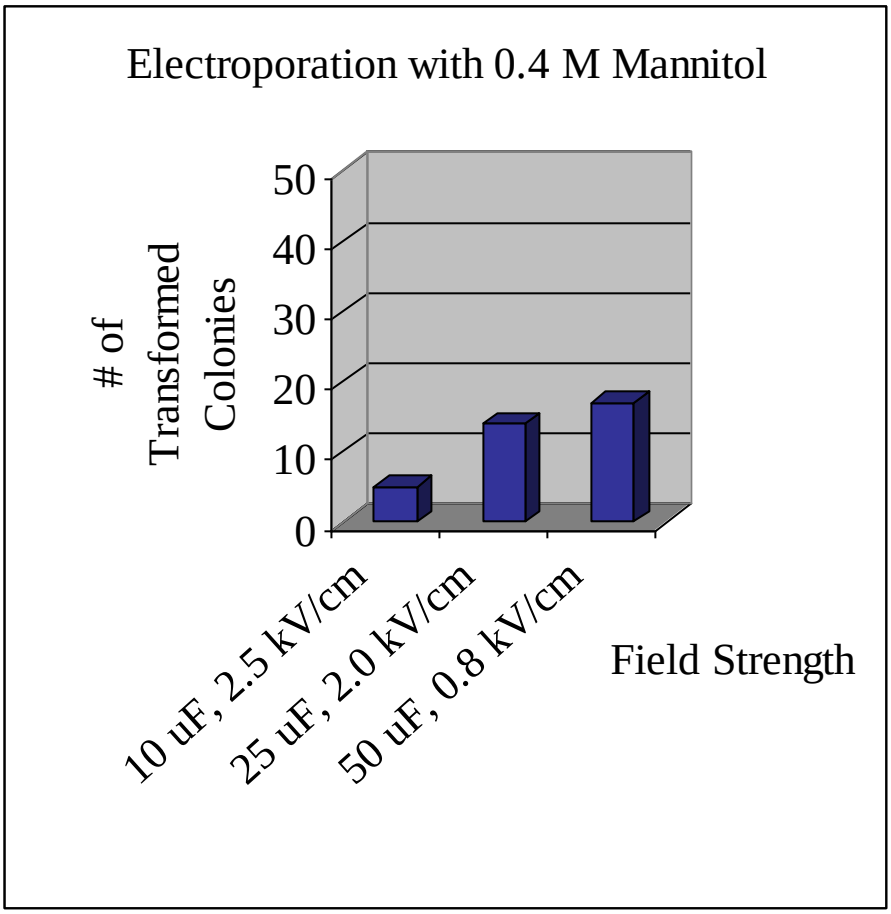


Figure 21. Effect of Field Strength on Transformation Efficiency in 0.4 M Mannitol. A 100-ml culture *Chlamydomonas reinhardtii* cells were centrifuged at 3000 rpm for seven minutes. The supernatant TAP was decanted and the pellet was resuspended in 5 ml of fresh TAP solution. The green algae cells (100 µL) were aliquoted into a 24-well plate. Mannitol in the concentration of 0.4 M was used as the hypertonic solution to produce plasmolysis in algae and allow the pSP124S plasmid DNA to enter the cells. Each aliquot of algal cells was electroporated with two pulses at the designated field strength. The results from seven experiments were totaled. The use of 0.4 M Mannitol was not a feasible option as compared to sucrose, sorbitol or ultra pure water. The worse transformation efficiencies were received with this sugar.



DISCUSSION

Bacteria, plants, and animals have been genetically modified for several decades now. These genetically modified organisms (GMOs) have been used to mass-produce mammalian proteins. The proteins can be purified and employed as immunotherapeutic, diagnostic, pharmaceutic, and prognostic agents in the treatment of several forms of adenocarcinoma, hormone deficiencies, and vaccine therapy.

Glucose Regulated Protein 78, a molecular chaperone, which has the primary of function of associating with free immunoglobulin heavy chains in the endoplasmic reticulum until they are assembled with the light chains, has been the focus of this study. There are innumerable homologues of this protein among eukaryotes, but not all of them originated from the same source or common ancestor. Human GRP 78 is nearly identical to the same protein expressed in *Mus musculus,* the ordinary house mouse, and in the golden hamster, also called the Syrian hamster, *Mesocricetus auratus.* When comparing the glucose regulated protein 78 of spinach, it was remarkably similar to the human homolog. Human GRP 78 was most dissimilar to the opportunistic fungi, *Aspergillus fumigatus* or the budding yeast, *Saccharomyces cerevisae*.

The phylogenetic tree generated reiterated the same fact, but by using graphics rather that computational studies. The distance between the human GRP 78 and

that of the mouse and hamster were very close and this was as expected. Farther away were the fungi and budding yeast. However, this study was unable to determine exactly, if at all the GRP 78 from the *Drosophilia melanogaster*, the pesky fruit fly, was actually a homologue or no. The computational analyses suggest it may be, but the phylogenetic tree hints otherwise. The bioinformatics databases are unbelievable in the amount of work they can perform in little time. There is a wealth of information at the seekers fingertips. The databases used in this study were extremely user friendly and took a very time-consuming task and made them quick and easy. NCBI, CLC bio and STRING 8.1 had free access to their databases and only required that credit be given for their use. These sites can only be improved by incorporating more protein-protein interaction data. The data are abundant, but lacking in the ability to compare proteins across various genera. As much research as has been completed on *Chlamydomonas reinhardtii*, very little, if any, was accessible to compare with organisms that were vastly different from the green alga. The biggest challenge was realized when trying to complete protein-protein interaction studies. The bioinformatics databases and search engines were unable to successfully analyze protein sequences that originated from genera. The sites did include information stating that this information would be readily available in the future, but was a work in progress. NCBI, by far, led the others in this arena. However, none of the programs were able to ascertain, even theoretically, how any of the studied proteins would interact once introduced within *C. reinhardtii.* That was a true disappointment.

Bacteria, plants, and animals have been genetically modified for several decades now. These genetically modified organisms (GMOs) have been used to mass-produce

mammalian proteins. The proteins can be purified and employed as immunotherapeutic, diagnostic, pharmaceutical, and prognostic agents in the treatment of several forms of adenocarcinoma, hormone deficiencies, and vaccine therapy.

Chlamydomonas reinhardtii, a unicellular eukaryote, is easily maintained in the laboratory environment. This member of the family Chlorophyta, was easily transformed by electroporation with the parent pSP124S as detailed in this study. Once the conditions and parameters for electroporation, a direct gene transfer method, were revealed, the chloroplast, nuclear, and mitochondrial genomes could be easily modified. In this study, we attempted to modify the nuclear genome only. However, further testing must be accomplished to ensure that is where the DNA was transferred to after subjecting the algae to plasmolysis and electroporation.

Using 0.4 M sucrose in the electroporation media enabled the most remarkable increase in the final number of transformants. This result reflects the increase in survival of the cells and increased efficiency of introducing exogenous DNA into the cells. Incorporating the starch-embedding technique during plating also may have contributed to the dramatic increase in survival of the algae post electropulsation. Previously, this technique was usually applied strictly to cell wall deficient mutants.

To attain the highest transformation frequency, three different field strength parameters were studied. The capacitance and voltage were varied and the results were nearly the same for 10uF, and 2.5 kV/cm, 25uF, 2.0 kV/cm. However, more experienced scientists were able to obtain their best results using the third parameter, 50uF, 0.8 kV/cm. The efficiency in which we were able to introduce foreign DNA into the green algae is still

lower than what is expected for yeast. Further experiments may be conducted to investigate the differences that may be due to varying the amount of exogenous DNA, utilizing linear DNA versus plasmid DNA, adding carrier DNA, or using another method, such as the glass beads vortexing method, to introduce the foreign DNA.

Once the electroporation conditions and the field strength parameters that worked best for *Chlamydomonas reinhardtii* wild type strain, cc125 mt+, were discovered, introduction of the mammalian molecular chaperone (*grp 78*) and/or the humanized pancarcinoma antibody (*hucc 49*) could be investigated. We hypothesize that the glucose regulated protein 78 would shield the HuCC49 from degradation until the antibody could be properly folded, thereby increasing the production of this mammalian protein. The *grp78* gene could be extracted from the pA78H plasmid; the *hucc49*-gfp from the p35S-49-GRP plasmid. Both of these genes could, in separate experiments be inserted into a duplicated RubisCO cassette of pSP124S, thereby leaving the *ble*^{r} gene, as a selectable marker. Upon success of those experiments, the "cockroach of the algae world" could take the Plantibody® industry by storm and produce exceedingly more antibodies in a mere fraction of the time it takes true, green vascular plants to produce the same agent. The pharmaceutical world and medical community could be inundated with a variety of immunotherapeutic, diagnostic, and prognostic agents from which to choose for modern day diagnoses and treatments.

Further studies would need to be performed to analyze the functionality of the anticarcinoma antibody once it is produced in the algae. Further down the road, binding, specificity, and cytotoxicity assays would be the next logical steps to take to ensure that

the antibody produced in *Chlamydomonas reinhardtii* is comparable in functionality as the original humanized mouse monoclonal antibodies is.

References

References

Alvim, F.C., Carolino, S.M.B., Cascardo, J.C.M., Nunes, C.C., Martinez, C.A., Otoni, W.G., and Fontes, E.P.B. 2001. Enhanced accumulation of BiP in transgenic plants confers tolerance to water stress. *Plant Physiology* **126**: 1042-1054.

Asurmendi, S., Berg, R.H., Koo, J.C., Beachy, R.N. 2004. Coat protein regulates formation of replication complexes during tobacco mosaic virus infection. *Proceedings of the National Academy of Sciences of the United States of America* **101**(5): 1415-1420.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. *J Mol Biol* **215** (3): 403-410.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J. Zheng, Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl Acid Res* **25** 3389-3402.

Altschul, S.F., Wootton, J.C., Gertz, M.E., Agarwala, R., Morgulis, A., Schaffer, A.A., Yu, Y.K. 2005. Protein database searches using compositionally adjusted substitution matrices. *FEBS Journal* **272**: 5101-5109.

Baxter, L., Bryant, J., Cave, C.B., Milne, R. 2007. Recombinant growth hormone for children and adolescents with turner syndrome. *Cochrane Database Syst Rev* (1): CD003887.

Beckie, H. J., Séguin-Swartz, G., Nair, H., Warwick , S. I. , Johnson, E. 2004. Multiple herbicide–resistant canola can be controlled by alternative herbicides. *Weed Science* **52:** 152-157.

Berg, P et al. 1975. Summary statement of the asilomar conference on recombinant DNA molecules. *Proc. Natl. Acad. Sci. USA* **72** (6): 1981–1984.

Blumenthal, A. 1999. Measurement of green fluorescence protein in plants: quantification, correlation to expression, rapid screening and differential gene expression. *Plant Science* **142**: 93-99.

Borisjuk, N. 1999. Production of recombinant proteins in plant root exudates. *Nature Biotechnology* **7**: 466-469.

Boynton, J.E., Gillham, N.W., Harris, E.H., Hosler, J.P., Johnson, A.M. 1988. Chloroplast transformation in *Chlamydomonas* with high–velocity microprojectiles. *Science* **240**: 1534-1537.

Brake, D.G. and Evenson D.P. 2004. A generational study of glyphosate-tolerant soybeans on mouse fetal, postnatal, pubertal and adult testicular development. *Food Chemistry and Toxicology.* **42**(1):29-36.

Bryant, J., Baxter, L., Cave, C.B., Milne, R. 2007. Recombinant growth hormone for idiopathic short stature in children and adolescents. *Cochrane Database Syst Rev* (3): CD004440.

Cai, B., Tomida, A., Mikami, K., Nagata, K., and Tsuruo, T. 1998. Down-regulation of epidermal growth factor receptor-signaling pathway by binding of GRP78/BiP to the receptor under glucose-starved stress conditions. *Journal of Cellular Physiology* **177**: 282-288.

Chakraborty, S., Chakraborty, N. and Datta, A. 2000. Increased nutritive value of transgenic potato by expressing a nonallergenic seed albumin gene from *amaranthus hypochondriacus. Proc Natl Acad Sci USA* **97** (7): 3724-3729.

Choi, H.J., Chandrasekhar,T., Lee, H.Y. and Kim, K.M. 2007. Production of herbicide-resistant transgenic sweet potato plants through *agrobacterium tumefaciens* method. *Journal of Plant Cell, Tissue and Organ Culture* **91** (3): 235-242.

CLC bio Database, Protein Reports. http://www.clcbio.com. Accessed 16 June 2009.

Cohen, S.N. et al. 1973. Construction of biologically functional bacterial plasmids in vitro. *Proc. Natl. Acad. Sci. USA* **70** (11): 3240–3244.

Conrad U., Fiedler, U. 1998. Compartment-specific accumulation of recombinant immunoglobulins in plant cells: an essential tool for antibody production and immunomodulation of physiological functions and pathogen activity. *Plant Molecular Biology* **38**: 101-109.

Dandekar,T., Snel, B., Huynen, M. and Bork P. 1998. Conservation of gene order: a fingerprint of proteins that physically interact. *Trends Biochem. Sci.***23**: 324-328.

Daniell, H., Streatfield, S., Wycoff, K. 2001. Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. *TRENDS in Plant Science* **6**(5): 219–226.

Davies, J.P., Yildiz, F.H., and Grossman, A.R. 1994. Mutants of *Chlamydomonas* with aberrant responses to sulfur deprivation. *Plant Cell* **6**: 53-63.

Davies, J.P., Yildiz, F.H., and Grossman, A.R. 1996. SacI,m a putative regulator that is critical for survival of *Chlamydomonas reinhardtii* during sulfur deprivation. *EMBO Journal* **15**: 2150-2159.

Davis, T.A., White, C. A., Grillo-López, A. J., Velásquez, W. S., Link, B., Maloney D. G., R. Dillman, O., Williams, M. E., Mohrbacher, A., Weaver, R., Dowden, S., and Levy, R. 1999. Single-Agent Monoclonal Antibody Efficacy in Bulky Non-Hodgkin's Lymphoma: Results of a Phase II Trial of Rituximab. *Journal of Clinical Oncology* **17** (6):1851-1858.

Delagrave, S., Hawtin, R., Silva, C., Yang, M. and Youvan, D. 1995. Red-shifted excitation mutants of the green fluorescent protein. *Biotechnology* **13**: 151-154.

Delwiche, C. F. and Palmer, J. D. 1996. Rampant horizontal transfer and duplication of RubisCO genes in eubacteria and plastids. *Molecular Biology and Evolution* **13**: 873–882.

Donnelly, S., McCarthy, C.R. and Singleton, R. Jr. 1994. The brave new world of animal biotechnology. Special Supplement, Hastings Center Report

Doran, P. 2000. Foreign protein production in plant tissue cultures. *Current Opinion in Biotechnology* **11**(2): 199-204.

Ehrenfeld, G.M., et al. 1987. Copper-dependent cleavage of DNA by bleomycin. *Biochemistry* 26: 931-942.

Enright ,A.J., Iliopoulos, I., Kyripides, N.C. and Ouzounis, C.A. 1999. Protein interaction maps for complete genomes based on gene fusion events. *Nature* **402**: 86-90.

Fernandez, P.M., Tabbara, S.O., Jacobs, L.K., Manning, F.C., Tsangaris, T.N., Schwartz, A.M., Kennedy, K.A., and Patierno, S.R. 2000. Overexpression of the glucose-regulated stress gene GRP78 in malignant but not benign human breast lesions. *Breast Cancer Research and Treatment* **59**(1): 15-26.

Fischer, R., Schumann, D., Zimmermann, S., Drossard, J., Sack, M., and Schillberg, S. 1999. Expression and characterization of bispecific single chain Fv fragments produced in transgenic plants. *European Journal of Biochemistry* **262**: 810-816.

Foster, K., Foster, H., and Dickson, J.G. 2006. Gene therapy progress and prospects: Duchenne muscular dystrophy. *Gene Ther.* **13** (24): 1677–85.

Fox, J.L. 2003. Resistance to *Bt* toxin surprisingly absent from pests. *Nature Biotechnology* **21**(9): 958 – 959.

Gerdes, H., Kaether, C. 1996. Green Fluorescent Protein: Applications in Cell Biology. *FEBS Lett.* **389**: 44-47.

Gordon, J.W. and Ruddle, F.H. 1981. Integration and stable germ line transformation of genes injected into mouse pronuclei. *Science* **212**: 1244-1246.

Gossler, A. et al. 1986. Transgenesis by means of blastocyst-derived embryonic stem cell line. *Proc Natl Acad Sci* **86**: 9065-9069.

Gressel, J. 1992. Addressing real weed science needs with innovations. *Weed Technology* **6**:509-525.

Guidelines for research involving recombinant DNA molecules," *Federal Register* **41**, no. 131, pp. 27911-27943 (1976).

Heavey, S. 2009. US approves first drug from dna-altered animals. http://www.reuters.com/article/domesticNews/idUSTRE5154OE20090206?sp=true Retrieved on 24 Mar 2009

Heim, R., Prasher, D.C. and Tsien, R.Y. 1994. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 12501- 12504.

Hillman, J.D. 2002. Genetically modified *streptococcus mutans* for the prevention of dental caries. *Antonie Van Leeuwenhoek* **82** (1-4): 361–366.

Hinchee, M.A.W., Connor-Ward, D.V., Newell, C.A., McDonnell, R.E., Sato, S.J., Gasser, C.S., Fischhoff, D.A., Re, D.B., Fraley, R.T., and Horsch, R.B. 1988. Production of transgenic soybean plants using *agrobacterium*-mediated DNA transfer. *BioTechnology* **6**: 915-922.

Hogeweg, P. and Hesper, B. 1978. Interactive instruction on population interactions. *Comput Biol Med* **8**:319-27.

Hogeweg, P. and B. Hesper B. 1984. The alignment of sets of sequences and the construction of phyletic trees: an integrated method. *J Mol Evol* **20**: 175-186.

Holger S., Stegan S., Markus S., Achim H., Jorg M., Michael M., Yu-Cai L., and Rainer, F. 1999. Accumulation of Antibody fusion proteins in the cytoplasm and ER of plant cells. *Plant Science* **149**: 63-71.

Horken, K.M. and Tabita, F.R. 1999. The "green" form I ribulose 1,5-bisphosphate carboxylase/oxygenase from the nonsulfur purple bacterium Rhodobacter capsulatus. *Journal of Bacteriology* **181**(13): 3935-3941.

Jefferis, R., and Lund, J.. 1997. Glycosylation of antibody molecules: structural and functional significance. *Chemical Immunology and Antibody Engineering* **65**: 111-128.

Jensen, L.J., Kuhn, M., Stark M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M., Bork, P. and von Mering, C. 2009. STRING 8– a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res.* **37** (Database issue): D412-416.

Julian, Ma., Ban, Y., Hitmat, K., Wycoff, D., Vine, D., Lloyd, Y., Mich, B., and Hein, T. 1998. Characterization of a recombinant plant monoclonal secretory antibody and preventive immunotherapy in humans. *Nature Medicine* **4**(5): 601-606.

Kaether, C. and Gerdes, H. 1995. Visualization of protein transport along the secretory pathway using green fluorescent protein. *FEBS Lett***. 369**: 267-271.

Kahana, J., Schapp, B. and Silver, P. 1995. Kinetics of spindle pole body separation in budding yeast. *Proceedings of the National Academy of Sciences of the United States of America* **92**: 9707-9711.

Kashmiri, S. 1995. Generation, Characterization, and *In Vivo* Studies of Humanized Anticarcinoma Antibody CC49. *Hybridoma* **14**(5): 461-471.

Kashmiri, S., Iwahashi, M., Tamura, M., Padlan, E., Milenic, D., and Schlom, J. 2001. Development of a minimally immunogenic variant of humanized anti-carcinoma monoclonal antibody CC49. *Critical Reviews in Oncology/Hematology* **30**: 3-16

Ke, X., Zhu, X., and Shi, H., et al. 1996. Electroporation of immature maize zygotic embryos and regeneration of transgenic plants. *Transgenic Research* **5**(3): 219.

Khoudi, H., Laberge, S., Ferullo, J.M., Bazin, R., Darveau, A., Castoguay, Y., Allard, G., Lemieux, R., and Vezina, L.P. 1999. Production of a diagnostic monoclonal antibody in perennial alfalfa plants. *Biotechnology and Bioengineering* **64**: 135-143.

Kindle, K.L., Schnell, R.A., Fernandez, E. and Lefebvre, P.A. 1989. Stable transformation of *Chlamydomonas* using *Chlamydomonas* gene for nitrate reductase. *Journal of Cell Biology* **109**: 2589-2601.

Koscianska, K. and Wypijewski, K. 2001. Electroporated intact BY-2 tobacco culture cells as a model of transient expression study. *Acta Biochimica Polonica* **48**(3): 3/2001.

Kusnadi A., Nikolov Z., and Howard J. 1997. Production of recombinant proteins in transgenic plants: practical considerations. *Biotechnology and Bioengineering* **56**: 473-484

Kyung-Tae, P., JunHi, Y., Fosket, D., and In-Seob, H. 1996. A comparison of the activity of three cauliflower mosaic virus 35S promoters in rice seedlings and tobacco (BY-2) protoplasts by analysis of GUS reporter gene transient expression. *Plant Science* **114**(2): 141-148.

Leary, W. 1996. Genetic engineering of crops can spread allergies, study shows. *New York Times,* Published 14 Mar 1996.

Lefebvre, P.A. and Silflow, C.D. 1999. *Chlamydomonas*: The cell and its genomes. *Genetics* **151**: 9-14.

Lee, R. E. Jr., Warren, G.J., and Gusta, L.V. (Editors) 1995. "Chapter 4, "Biochemistry of bacterial ice nuclei by Ray Fall and Paul K. Wolber." *Biological Ice Nucleation and Its Applications*. St. Paul, Minnesota: APS PRESS (The American Phytopathological Society). pp. 63-83.

Lee, S.Y. 1996. High cell-density culture of *escherichia coli*. *Trends Biotechnol.* **14** (3): 98–105.

Lim, C., Kimata, K., Oka, M., Nomaguchi, K., and Kohno, K. 1995. Thermosensitivity of a green fluorescent protein utilized to reveal novel nuclear-like compartments. *Journal of Biochemistry (Tokyo)* **118**: 13-17.

Losey, J.E., Rayor, L.S., and Carter, M.E. 1999. Transgenic pollen harms monarch larvae. *Nature* **399** (6733): 214.

Magnuson, N. 1998. Secretion of biologically active human interleukin-2 and interleukin-4 from genetically modified tobacco cells in suspension culture. *Protein Expression and Purification* **13**: 45-52

Manivasakam, P. and Schiestl, R.H. 1993. High efficiency transformation of *saccharomyces cerevisiae* by electroporation. *Nucleic Acids Research* **21**: 4414-4415.

Marcotte, E.M., Pellegrini, M., Ng, H.L., Rice, D.W., Yeates, T.O., Eisenberg, D., Yeates, T.O. and Eisenberg, D. 1999. Detecting protein function and protein-protein interactions from genome sequences. *Science* **285**: 751-753.

Marshall, J., Molloy, R., Moss, G., Howe, J. and Hughes, T. 1995. The jellyfish green fluorescent protein: a new tool for studying ion channel expression and function. *Neuron* **14**: 211-215.

Martineau, B. 2001. First fruit: The creation of the flavr savr tomato and the birth of biotech food. McGraw-Hill.

Mitra, R., Silva, C. and Youvan, D. 1996. Fluorescence resonance energy transfer between blue-emitting and red-shifted excitation derivatives of the green fluorescent protein. *Gene* **173**: 13-7.

Morell, M. K., Paul, K., Kane, H. J., and Andrews, T. J. 1992. RubisCO: maladapted or misunderstood. *Australian Journal of Botany* **40:** 431-441.

Morin, J. and Hastings, J. 1971. Energy transfer in a bioluminescent system. *Journal of Cell Physiology* **77**: 313-8.

Mote,P.L., Tillman, J.B., and Spindler, S.R. 1998. Glucose regulation of GRP78 gene expression. *Mechanisms of Aging and Development* **104**(2): 149-158.

Munro, S. and Pelham, B. 1986. An hsp 70-like protein in the ER: Identity with the 78 kDa glucose-regulated protein and immunoglobulin heavy chain binding protein. *Cell* **46**: 291–300.

Napier, R.M., Fowke, L.C., Hawes, C., Lewis, M., and Pelham, H.R. 1992. Immunological evidence that plants use both HDEL and KDEL for targeting proteins to the endoplasmic reticulum. *Journal of Cell Science* **102**(2): 261-271.

National Center for Biotechnology Information (NCBI) Database. GRP 78 Sequence Analyses. http://www.ncbi.nlm.nih.gov. Accessed on 24 Jun 2009.

Newman, S. M., Gillham, N.W., Harris, E.H., Johnson, A.M. and Boynton, J.E. 1991. Targeted disruption of chloroplast genes in *Chlamydomonas reinhardtii*. *Molecular Genetics* **230**: 65-74.

Pazos, F. and Valencia, A. 2001. Similarity of phylogenetic trees as indicator of protein-protein interaction. *Protein Engineering*, **9** (14), 609-614.

Pellegrini, M., Marcotte, E.M., Thompson, M.J., Eisenberg, D., and Yeates, T.O. 1999. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proc Natl Acad Sci U S A.*, **96**, 4285-8.

Persons, D.A., and Nienhuis, A.W. 2003. Gene therapy for the hemoglobin disorders. *Curr. Hematol. Rep.* **2** (4): 348–55.

Quecini, V.M., de Oliveira, C.A., Alves, A.C., Vieira, M.L.C. 2002. Factors influencing electroporation-mediated gene transfer to *Stylosanthes guianensis* (Aubl.) Sw. protoplasts. *Genetic and Molecular Biology* **25**(1)

Read, B. A. and Tabita, F. R. 1994. RubisCOs from some non-green algae, in particular, have higher specificities for CO_2, compared with O_2, than the higher-plant enzyme. *Arch. Biochem. Biophys.* **312**: 210–218.

Rizzuto, R., Brini, M., De Giorgi, F., Rossi, R., Heim, R., Tsien, R. and Pozzan, T. 1996. Double labeling of subcellular structures with organelle-targeted GFP mutants *in vivo*. *Current Biology* **6**: 183-188.

Rochaix, J.D. 1995. *Chlamydomonas reinhardtii* as photosynthetic yeast. *Annual Review in Genetics* **29**: 209-230.

Rochaix, J.D. 2001. Assembly, Function, and Dynamics of the Photosynthetic Machinery in *Chlamydomonas reinhardtii*. *Plant Physiology* **127**: 1394-1398.

Rosenecker, J., Huth, S., and Rudolph, C. 2006. Gene therapy for cystic fibrosis lung disease: current status and future perspectives. *Curr. Opin. Mol. Ther.* **8** (5): 439–45.

Roy, H., Andrews, T. J., Leegood, R. C., Sharkey, T. D., and von Caemmerer, S. 2000. eds. *Photosynthesis: Physiology and Metabolism* (Kluwer, Dordrecht, The Netherlands).

Sabri, N., Pelissier, B. and Teissie, J. 1996. Transient and stable electrotransformations of intact black Mexican sweet maize cells are obtained after preplasmolysis. *Plant Cell Reports* **15**: 924-928.

Salmon, V. 1998. Production of human lactoferrin in transgenic tobacco plants. *Protein Expression and Purification* **13**: 127-135.

Satoh, M. 1993. Modulation of the Phosphorylation of Glucose-Regulated Protein, GRP78 by Transformation and Inhibition of Glycosylation. *Experimental Cell Research* **205**: 76-83.

Schnell, R.A. and Lefebvre, P.A. 1993. Isolation of the *Chlamydomonas* regulatory gene NIT2 by transposon tagging. *Genetics* **134**: 737-747.

Schuler, T.H., Denholm, I., Clark, S.J., Stewart, C.N., and Poppy, G.M. 2004. Effects of Bt plants on the development and survival of the parasitoid *Cotesia plutellae* (Hymenoptera: Braconidae) in susceptible and Bt-resistant larvae of the diamondback moth, *Plutella xylostella* (Lepidoptera: Plutellidae). *Journal of Insect Physiology* **50**(5): 435- 443.

Shimogawara, K., Fujiwara, S., Grossman, A. and Usuda, H. 1997. High efficiency transformation of C*hlamydomonas reinhardtii* by electroporation. *Genetics* **148**: 1821-1828.

Shu Chen-F., Schlom.A. and Kashmiri S. 1993. Secretion of a single-gene encoded immunoglobulin from myeloma cells. *Immunology* **90**: 7995-7999.

Sodeinde, O.A. and Kindle, K.L. 1993. Homologous recombination in the nuclear genome of *Chlamydomonas reinhardtii*. *National Academy of Science USA* **90**: 9119-9203.

Song, M.S., Park, Y.K., Lee, J.-H., and Park, K. 2001. Induction of glucose-regulated protein 78 by chronic hypoxia in human gastric tumor cells through a protein kinase C-ε/ERK/AP-1 signaling cascade. *Cancer Research* **61**: 8322-8330.

Songstad, D.D., Halaka, F.G., and Deboer, D.L. et al. 1993. Transient expression of GUS and anthocyanin constructs in intact maize immature embryos following electroporation. *Plant Cell, Tissue and Organ Culture* **33**: 95.

Sorokin, A.P., Ke, X., Chen, D., and Elliott, M.C. 2000. Production of fertile transgenic wheat plants via tissue electroporation. *Plant Science* **156**(2): 227-233.

Stoger, E., Vaquero, C., Torres, E., Sack, M., Nicholson, L., Drossard, J., Williams, S., Keen, D., Perrin, Y., Christou, P., and Fischer, R. Cereal crops as viable production and storage systems for pharmaceutical scFv antibodies. *Plant Molecular Biology* **42**(4): 535-590.

STRING 8.1: Search Tool for the Retrieval of Interacting Genes/Proteins. http://www.string.embl.de. Protein-protein interaction study. Accessed 12 Jul 2009.

Szala, S., Froehlich, M., Scollon, M., Kasai, Y., Steplewski, Z., Koprowski, H., and Linnenbach, A.J. 1990. Molecular cloning of cDNA for the carcinoma-associated antigen GA733-2. *Proceedings of the National Academy of Sciences* **87**(9): 3542-3546.

Tacket, C., Mason, H., Losonsky, G., Clements, J., Levine, M., and Arntzen, C. 1998. Immunogenicity in humans of a recombinant bacterial antigen delivered in a transgenic potato. *Mature Medicine* **4**: 607-613.

Tada, Y., Sakamoto, M., and Fujimura, T. 1990. Efficient gene introduction into rice by electroporation and analysis of transgenic plants: use of electroporation buffer lacking chloride ions. *Theoretical Applications of Genetics* **80**: 475.

Tabashnik, B.E. and Carriere.Y. 2003. Bt transgenic crops do not have favorable effects on resistant insects. *Journal of Insect Science* **4**: 4.

Tan, S.H., Zhang, Z., and Ng, S.K. 2004. ADVICE: Automated detection and validation of interaction by co-evolution. *Nucl. Ac. Res.*, **32** (Web Server issue):W69-72.

Takeshi, A., Daniel, K., Chong, J., Lawrence, M., and William, H. 1997. Expression on cholera toxin B subunit oligomers in transgenic potato plants. *Transgenic Research* **6**: 403-413.

Tof, I. 1994. Recombinant dna technology in the synthesis of human insulin. Little Tree Pty. Ltd . http://www.littletree.com.au/dna.htm. Retrieved on 24 Jan 2005.

Tuckey, B. 2002. Phreak of nature. *The Independent* **2**: 124

Turning plants into antibody factories. 1989. *Science News*.

Umezawa, H. 1976. Structure and action of bleomycin. *Prog Biochem Pharmacology* **11**:18-27

van Engelen, F., Schouten, A., Molthoff, J., Roosien, J., Salinas, J., Dirkse, W., Schots, A., Bakker, J., Gommers, F. and Jongsma, M. 1994. Coordinate expression of antibody subunit genes yields high levels of functional antibodies in roots of transgenic tobacco. *Plant Molecular Biology* **26**: 1701-1710.

Vogel, J.P., Misra, L.M., and Rose, M.D. 1990. Loss of BiP/GRP78 function blocks translocation of secretory proteins in yeast. *Journal of Cell Biology* **110**: 1885-1895.

von Mering, C., Huynen, M., Jaeggi, D., Schmidt, S., Bork, P., and Snel, B. 2003. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res.* **31** (1): 258-261.

von Mering, C., Jensen, L.J., Kuhn, M., Chaffron, S., Doerks, T., Kruger, B., Snel, B. and Bork, P. 2007. STRING 7 – recent developments in the integration and prediction of protein interactions. *Nucleic Acids Res.* **35** (Database issue): D358 -362.

Ward, W. 1979. *Photochemical and Photobiological Reviews*, K. Smith, Editor. Plenum: NY. p. 1-57

Wu, F.-S., Cahoon, A.B. 1994. Plasmolysis facilitates the accumulation of protein and DNA intro extra-plasmalemma spaces of intact plant cells. *Plant Science* **104**: 201-214.

Wu, F.-S., Cahoon, A.B., Shulleeta, M. 1995. Passage of DNA and protein molecules into intact plant cells and direct gene transfer. Plant Genome IV Conference. San Diego, CA, January 1995.

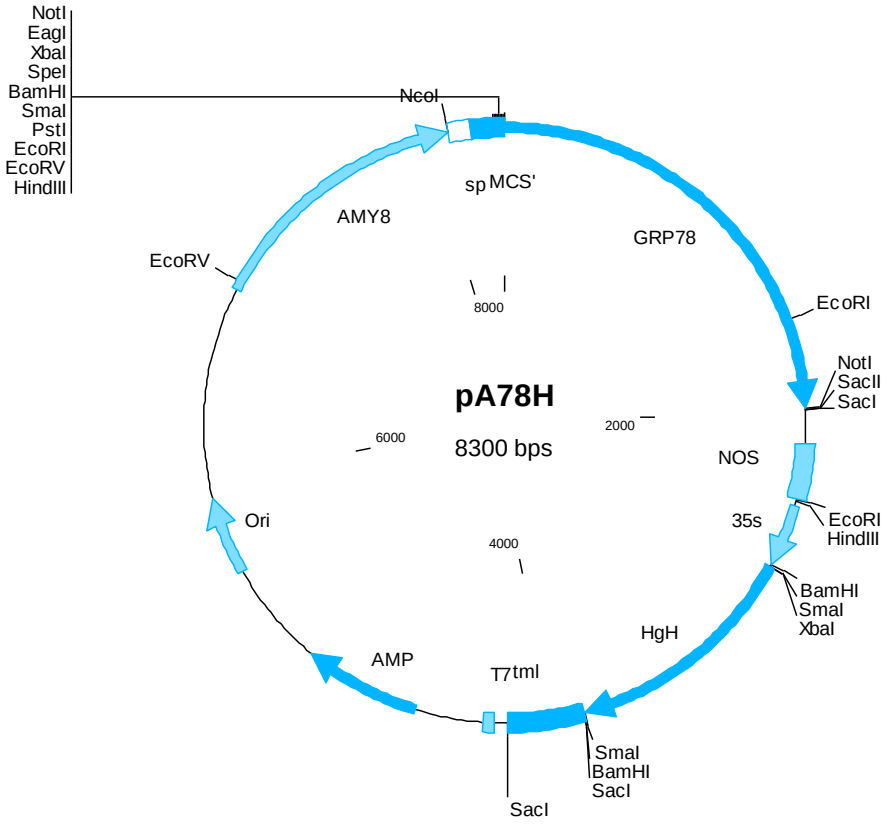
Wu, K., Malik, K., Tian, L., Hu, M., Martin, T., Foster, E., Brown, D., and Miki B. 2001. Enhancers and core promoter elements are essential for the activity of a cryptic gene activation sequence from tobacco, tCUP. *Molecular Genetics and Genomics* **265**(5): 763-770

Ye, et al. 2000. Engineering the provitamin a (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* **287** (5451): 303-305.

Zheng, H.C., Takahashi, H., Li, X.H., Hara, T., Masuda, S., Guan, Y.F., and Takano, Y. 2008. Overexpression of GRP78 and GRP94 are markers for aggressive behavior and poor prognosis in gastric carcinomas. *Human Pathology* **39** (7): 1042-1049.

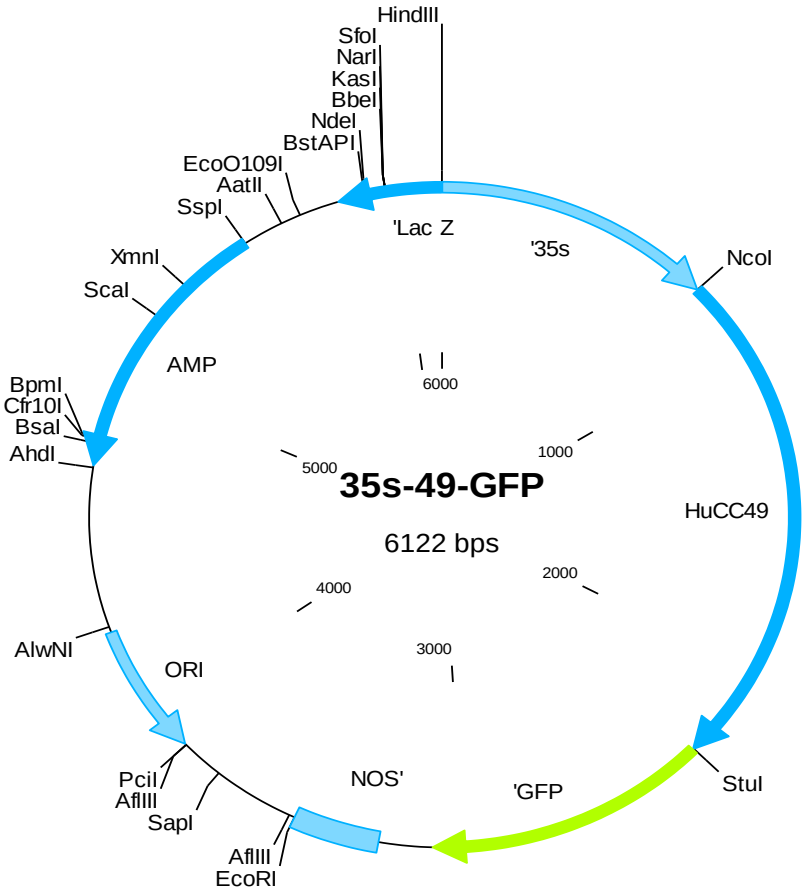
Appendices

APPENDIX A



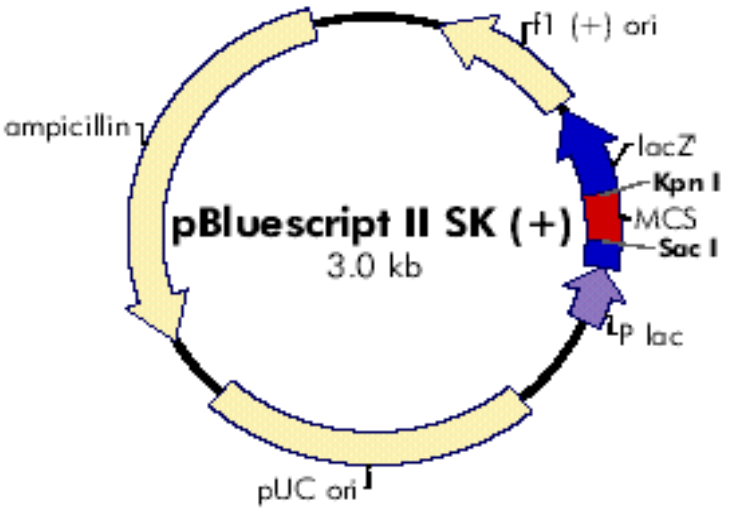
Originally constructed by Scott Taylor, 2001, Dr. Wu's Laboratory

APPENDIX B



Originally constructed by Scott Taylor, 2001, Dr. Wu's Laboratory

APPENDIX C



From: Stratagene.com Vector restriction map of pBluescript II SK. The cassette containing the RubisCO promoter followed by the *ble* gene was inserted into this vector at the *Sac*I and *Kpn*I sites to become pSP124S plasmid.

VITA

Katrina Patricia Ghazanfar was born in Brooklyn, New York on Thanksgiving Day of 1969. After completion of her primary education in NY, she joined the United States Army in 1987 and became a Medical Laboratory Specialist with the 44th Medical Brigade at Fort Bragg, North Carolina. Following her 1991 Army combat tour to Dhahran, Saudi Arabia in support of Operation Desert Storm, Katrina enrolled in Fayetteville State University (UNC) to pursue her B.S. in Medical Technology (Clinical Laboratory Sciences). She graduated Magna Cum Laude in 1994.

After receiving her undergraduate degree, Katrina rejoined the active duty Army and attended Officer Candidate School in Fort Benning, Georgia. Her last tour of duty with the Army was completed at Walter Reed Army Institute of Research, Dental Research Detachment where she worked as a Research Technologist.

In October of 1997, Katrina was commissioned in the United States Air Force as a Clinical Laboratory Officer, with a first duty assignment at Fort Dix/McGuire AFB, New Jersey and managed a small hospital laboratory. After selection for an Air Force Institute of Technology Civilian School Program scholarship, Capt Ghazanfar began her graduate studies at Medical College of Virginia Campus, Virginia Commonwealth University. She plans to do her postdoctoral research at an Air Force Medical Research Laboratory.