IDENTIFYING PROTEIN-PROTEIN INTERACTIONS

OF A CELL CYCLE REGULATOR
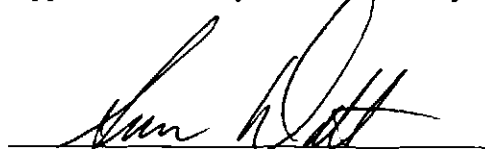
A Senior Honors Thesis
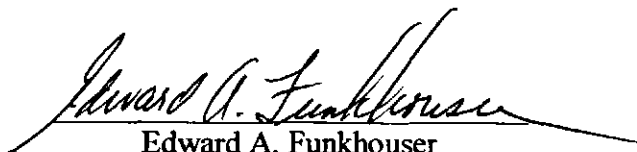
by

JOSEPH EDWARD AMOS

Submitted to the Office of Honors Programs
& Academic Scholarships
Texas A&M University
in partial fulfillment for the designation of

UNIVERSITY UNDERGRADUATE
RESEARCH FELLOW

Approved as to style and content by:

_____   _____
Sumana Datta   Edward A. Funkhouser
(Fellows Advisor)   (Executive Director)

April 2001

Group: Biochemistry

ABSTRACT

Identifying Protein-Protein Interactions

of a Cell Cycle Regulator. (April 2001)

Joseph Edward Amos
Department of Biochemistry
Texas A&M University

Fellows Advisor: Dr. Sumana Datta
Department of Biochemistry and Biophysics

The role of *anachronism (ana)* protein in stem cell division of *Drosophila melanogaster* was examined. Synthesis of identifiable *ana* protein was necessary. The identifying method exploited was that of antibody tagging using a myc epitope or a poly-Histidine region.

Analysis of previously utilized complementary DNA (cDNA) sequences showed errors in the alignment of the reading frame. Methods were employed to clone the required *ana* cDNA sequence into an expression vector (pBAD/Myc-HisA) within the correct reading frame. Restriction endonucleases were then used to confirm that the cDNA sequence and the expression vector had ligated properly.

Analysis of the results from the single- and double-restriction digests showed that the *ana* cDNA (1.6 kilobases) was successfully cloned and ligated into the pBAD/Myc-HisA (4.1 kilobases) expression vector. The *ana* cDNA was also successfully cloned in frame with the sequence coding for the myc epitope and the poly-Histidine region. This data provides groundwork for the expression of identifiable *ana* protein.

DEDICATION

This thesis is dedicated to my family and friends who have shown support in all of my endeavors. Without you none of this would be possible.

ACKNOWLEDGMENTS

I would like to recognize all people who made this project possible. Of course, Dr. Suma Datta, in whose laboratory I conducted the research. Thank you for teaching me the world of research and explaining things through example. Dr. Youngji Park, for showing me the most efficient ways to experiment and obtain results. Megan Reynolds, for answering my endless list of questions and always helping with the legwork.

I would also like to thank Dr. Efthimos Skoulakis for donating Leo DNA, and Dr. Max Summers for donating Histidine antibody.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

Importance

Cancer is caused by the uncontrollable division of diseased cells, while Alzheimer's

disease is caused by the loss of cells in certain areas of the brain. Although some of the

mechanisms that cause these diseases are known, efficient treatments have yet to be

discovered. By finding mechanism(s) that regulate cell division in *Drosophila*, we can

develop ways to treat diseases in higher eukaryotes (i.e., humans). Understanding these

molecular mechanisms will have tremendous potential for drug design that could allow for

the regeneration of damaged tissues, the treatment of neurological disorders, and the

treatment of cancer. Directed control of cell division would allow treatments for diseases

such as cancer and Alzheimer's.

Objective

The broad purpose of this research is to further characterize the cell cycle in

Drosophila melanogaster (fruit fly). The specific goal of my project is to synthesize a

protein that can be readily identified by a number of different techniques (Western blotting,

immunohistochemistry, biochemical techniques). Also, this synthesis and identification

must be easily reproducible.

This thesis follows the style and format of Cell.

Stem Cell Background

Stem cells are of great interest to researchers because of their unique qualities: they can make a variety of tissues and they have the power of self-renewal. Upon division, they are able to renew or make almost all the tissues in the body (blood, brain, skin, germline, and other internal organs) and at the same time regenerate a second stem cell capable of the same function. Stem cells have an infinite capacity to renew our cellular count and their cell division is critical for fertility and replenishment of tissues and organs. The fact that a single transplanted stem cell can restore all of the blood cells in a mouse (Ikuta et al., 1992) demonstrates this. Insufficient and/or uncontrolled stem cell division in humans can have detrimental effects, such as mental retardation (Miller, 1988), birth defects (Copp et al., 1988), and cancer (Ross et al., 1995). In *Drosophila*, insufficient stem cell proliferation leads to behavioral defects, abnormal brain structure, and death (Datta and Kankel, 1992; Datta, 1995), while premature stem cell proliferation results in loss of olfactory sensitivity and unusual axonal patterning (Ebens et al., 1993; Park et al., 1997). This stem cell division has been shown to be controlled by the *anachronism (ana)* protein.

Anachronism (ana) **Protein Background**

Ana is a secreted glycoprotein expressed in glial cells of the central and peripheral nervous systems (Ebens et al., 1993; Park et al., 1997). The *ana* protein is required to maintain the mitotic quiescence of stem cells, but does not appear to regulate the identity or quantity of the stem cells (Ebens et al., 1993). Interestingly, *ana* expression is not down

regulated when the quiescent stem cells begin to divide (Ebens et al., 1993), suggesting the existence of a mechanism(s) to inactivate *ana* or bypass *ana*-mediated repression.

Analysis of the sequence (474 amino acids, 55kD) shows that *ana* is a novel protein: it shows no obvious homology to any other protein in the current database. Additionally, the current anti-*ana* antibody for identifying the protein is unsatisfactory and does not work well in biochemical techniques. As a result, we need to develop a way to make a currently existing antibody recognize *ana*. To do this, we are employing a system called epitope tagging.

An epitope is that part of the structure that reacts with a specific antibody. Epitope tagging is simply adding an epitope to the structure of interest (protein, carbohydrate, etc.). For optimal results, we want to use an antibody that is reliable, commercially available, has been shown to work in biochemical techniques, and is unique to *Drosophila*. The epitope that we are tagging *ana* with is a myc epitope (part of the myc oncoprotein). The myc epitope has readily available antibodies that have been shown to work in Western blot analysis, immunoprecipitation, and immunohistochemistry. Furthermore, the coding sequence for the myc epitope is widely accessible and the myc epitope does not occur naturally in *Drosophila*.

OBJECTIVES

I. To synthesize *ana*-myc protein

    • Protein must be functional, identifiable, able to be secreted

II. Determine protein-protein interactions of *ana*

    • What does *ana* interact with

    • Find and characterize what proteins interact with *ana*

III. Understand how the protein-protein of ana interactions influence the regulation of stem cell division

MATERIALS AND METHODS

Sources of DNA

The *ana* cDNA used in the experiments was readily available in the laboratory as a result of other's work (Colleen Hough). The *ana* cDNA was contained in a carrier vector termed pUC and was given the designation pUC *ana* SB 39 (Figure 1). This vector has two restriction sites, Eco R1 and Hind III, surrounded the *ana* cDNA.

The other vector used, pBAD/*Myc*-His A, was commercially available through Invitrogen™ Corporation (Invitrogen™ #V440-01). This vector has the coding region for both the Myc epitope and for a poly-Histidine region (Figure 2), as well as restriction sites for Eco R1 and Hind III (Figure 3). This allows two methods of identification: myc antibodies and poly-Histidine antibodies. These antibodies can also be used for purification techniques.

Reticulocyte Lysate System

The primary and simplest strategy for making a tagged *ana* protein is to use a coupled transcription/translation system. The system we used was the Promega TNT® Coupled Reticulocye Lysate System (Promega #L4610). This system used a T7 RNA polymerase promoter and radiolabeled (via ^{35}S) our protein using only the protein's cDNA and the cellular components of rabbit blood. The advantages of this system are that it simplifies the process (transcription/translation are coupled), produces more protein than standard reticulocyte systems, and incorporates significantly more ^{35}S than standard
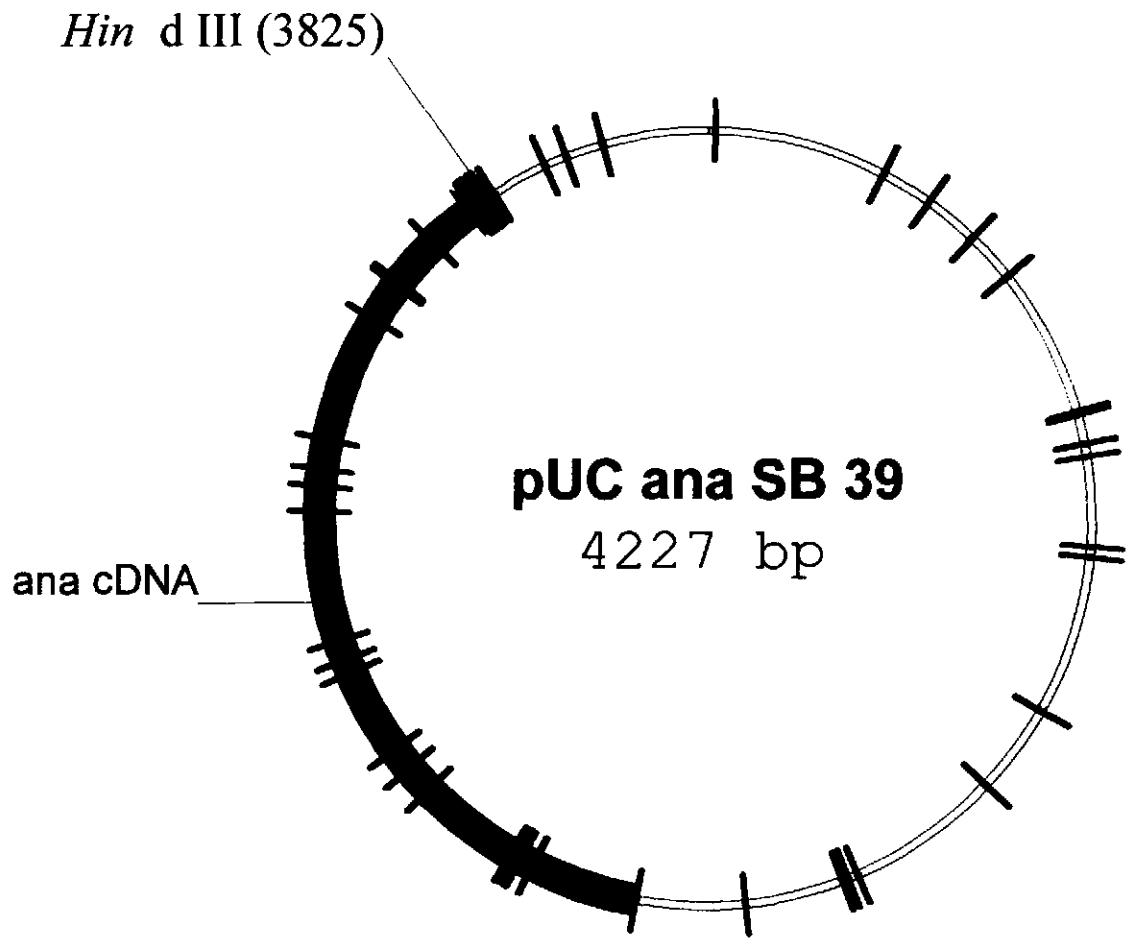
translation reactions. We also used an additional DNA (Leo DNA) as a positive control for our reticulocyte system. Leo is a protein that is naturally in *Drosophila*, and is only used as a control.

Transformation

The vectors used were first transformed into *E.coli* cells. Once transformed, the cells were allowed to grow and multiply on media plates that contained ampicillin, which allowed us to screen for vectors that had ampicillin resistance (pUC *ana* SB 39 and pBAD/*Myc*-His A). Once grown, the DNA was then recovered from the *E.coli* cells using a plasmid mini-prep and phenol chloroform cleanup. This straightforward procedure allowed us to produce an extremely large quantity of our DNA's, which were then used in subsequent experiments. The DNA was stored at -20°C until use.

Restriction Digests

The DNA recovered from the *E.coli* transformations was treated with restriction endonucleases Eco R1 and Hind III. The restriction fragments from these digests were then run on a 0.6% LM agarose electrophoresis gel to check the size of the DNA and verify that we were working with the correct copies. The DNA fragments were gel purified using QIAquick® Gel Extraction Kit (QIAGEN® #28706) and stored at -20°C until use.

Hin d III (3825)

ana cDNA

pUC ana SB 39

4227 bp

Eco R I (2233)

Figure 1. Carrier vector containing *ana* cDNA

EcoR1   Hind III   myc epitope

357   GGAATTCGAA   GCTTGGGCCC   GAA CAA AAA CTC ATC TCA GAA GAG
                                                 Glu   Gln   Lys   Leu   Ile   Ser   Glu   Glu

                                   Polyhistidine Region

GAT CTG AAT AGC GCC GTC GAC CAT CAT CAT CAT CAT CAT TGA
Asp   Leu   Asn   Ser   Ala   Val   Asp   His   His   His   His   His   His   ***

Figure 2. Coding region for Myc epitope and for poly-Histidine region

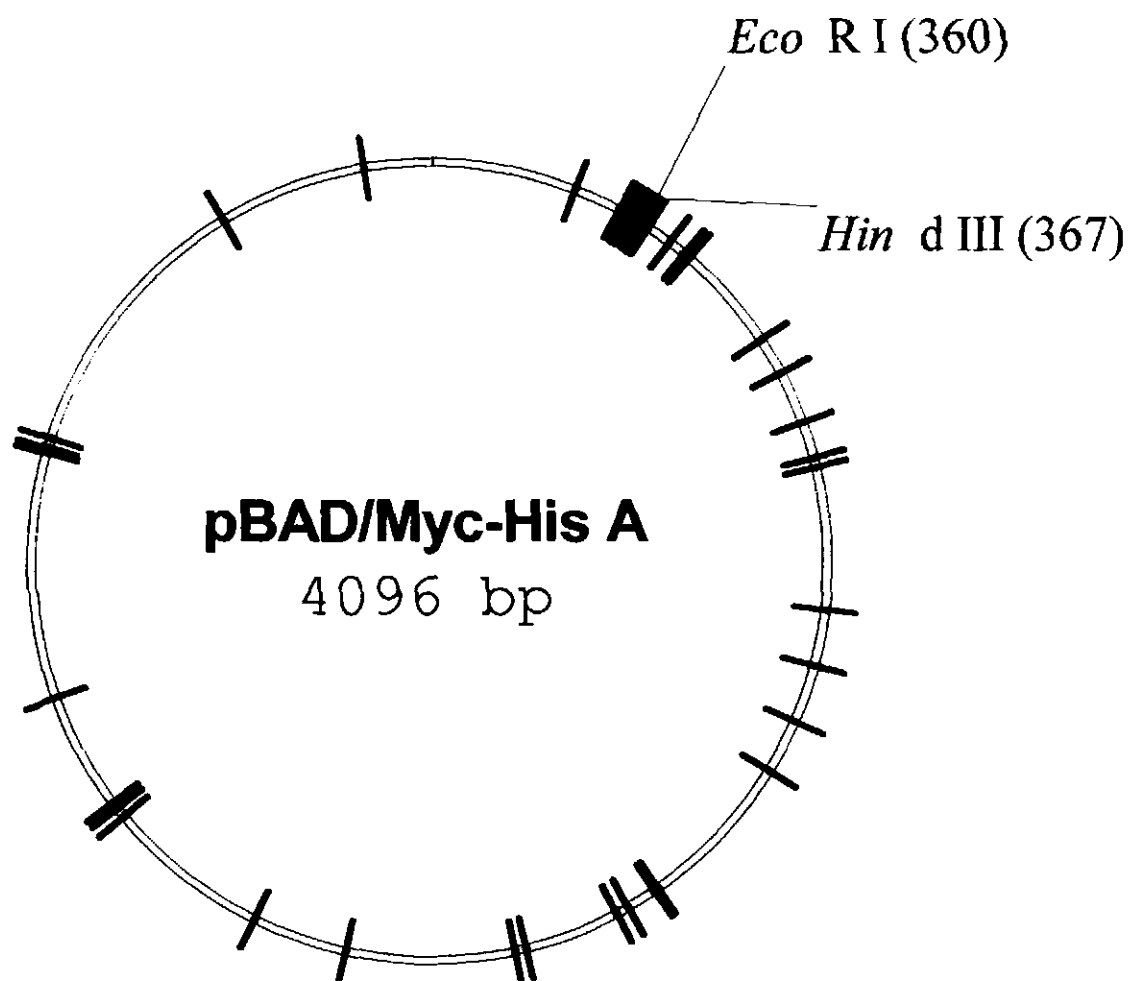(taken from Invitrogen™ Instruction Manual)

Figure 3. Invitrogen™ pBAD/*Myc*-His A vector

Cloning DNA

Once both the *ana* cDNA fragment and pBAD/*Myc*-His A fragment were purified,

they were ligated together. This ligation produced a new vector pBAD *Myc*-His/*ana* cDNA

(Figure 4). This new vector used pBAD as the carrier and had cloned in to it the *ana*

cDNA. The *ana* cDNA was meticulously placed in front of the coding region for the Myc

epitope and for the poly-Histidine region. Double restriction digests (Eco R1 and Hind III) were used to confirm the presence of the DNA fragments (*ana* cDNA and pBAD).
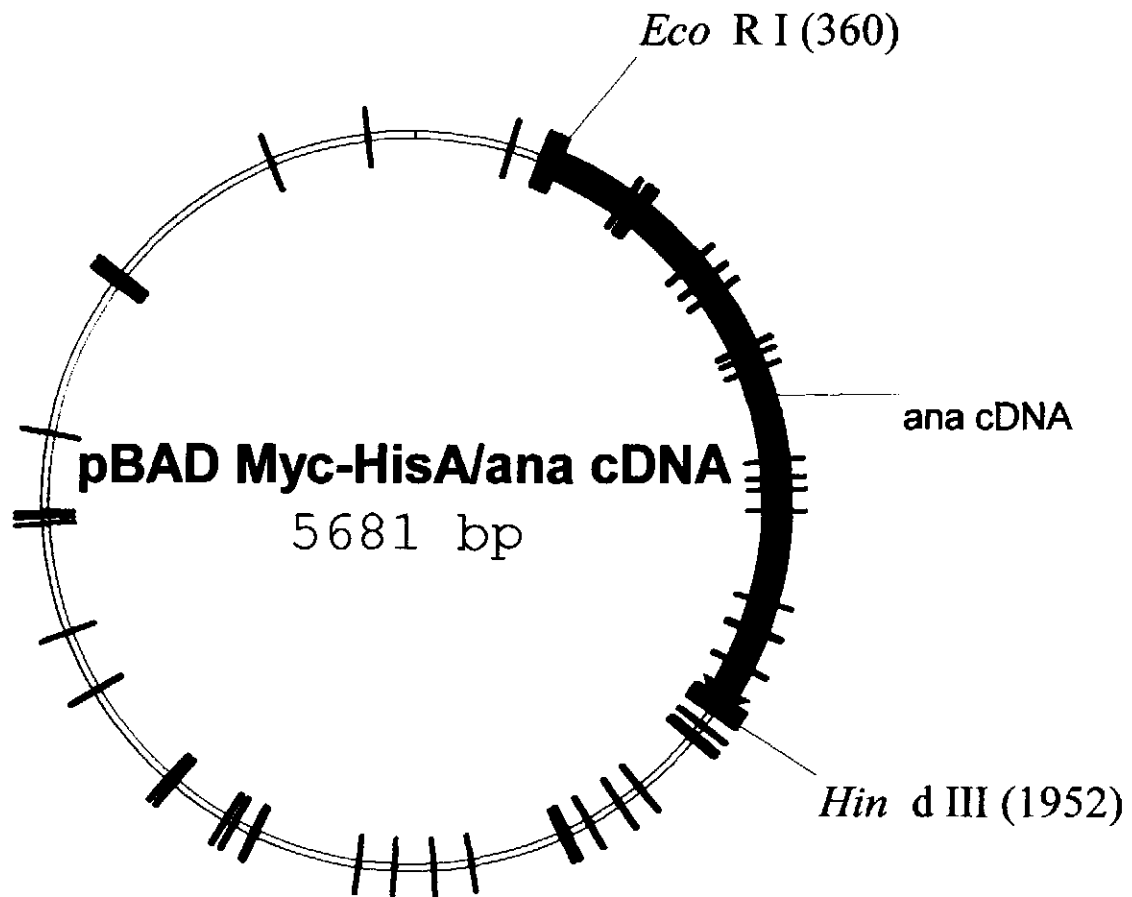

Figure 4. pBAD *Myc*-His/*ana* cDNA vector

Protein Expression & Visualization

The protein expression vector was expressed using *E.coli* TOP 10 according to the

Invitrogen™ protocol. The samples (before and after induction) were analyzed on an SDS-

PAGE gel and stained with Coomassie blue. The samples used were: 1) Kaleidoscope

Marker, 2,3) pBAD/*Myc*-His/*lacZ* (before/after induction, positive control), 4) pBAD *Myc*-

His/*ana* cDNA (before induction), 5) pBAD *Myc*-His/*ana* cDNA (after induction, 0.0002%

L-arabinose), 6) pBAD *Myc*-His/*ana* cDNA (after induction, 0.002% L-arabinose), 7)

pBAD *Myc*-His/*ana* cDNA (after induction, 0.02% L-arabinose), 8) pBAD *Myc*-His/*ana*

cDNA (after induction, 0.2% L-arabinose), 9,10) pBAD/*Myc*-His A (before/after

induction).

RESULTS

Our preliminary experiments involved protein production via the Reticulocyte Lysate System and antibody tagging through Western Blotting. It is crucial for us to be able to easily mark and identify our *ana*-myc protein for further experimentation. First, experiments were run to ensure that our reticulocyte system was transcribing and translating our cDNA and that it was utilizing the methionine radioactively labeled with sulfur (^{35}S). The cDNA used in the experiments is pBS11 SK *ana* myc and was constructed by Colleen Hough. In this vector it is assumed that the cDNA of *ana* and the coding sequence of the myc epitope are correct and in frame. The results (Figure 5) show that our cDNA construct is able to undergo eukaryotic *in vitro* translation via the reticulocyte system. The ^{35}S was actively used in the translation of the nascent proteins. When this film was compared to the marker, the size of the bands in each lane closely corresponds to the estimated size of *ana*-myc protein (~57kD).

Next, the reticulocyte experiment was run with Leo DNA (as a positive control), *ana* cDNA (*ana* RX-7 #3, as a negative control), and *ana*-myc cDNA (pBS11 SK *ana* myc). These proteins were then transferred to a nitrocellulose membrane and treated with myc antibodies as well as a chemiluminescence solution. Leo protein can be used as a positive control because we have specific antibodies that have been shown to react with only Leo and our *ana* protein can be used as a negative control because it should have no affinity for the myc antibody. These results (Figure 6) were very poor. Our myc antibodies showed affinity for everything (proteins and miscellaneous things) on the membrane,

including the "No DNA" control lane. This shows that the myc-antibodies have affinity for

something found in the Reticulocyte Lysate System. To double check that *in vitro*

translation was working, we stripped the membrane and treated it with Leo antibody.

Figure 7 shows the Leo antibody having affinity for the Leo protein, while not cross-

reacting with other proteins present. This demonstrates that the Reticulocyte Lysate System

is functioning appropriately. From these results, we concluded that the myc antibodies

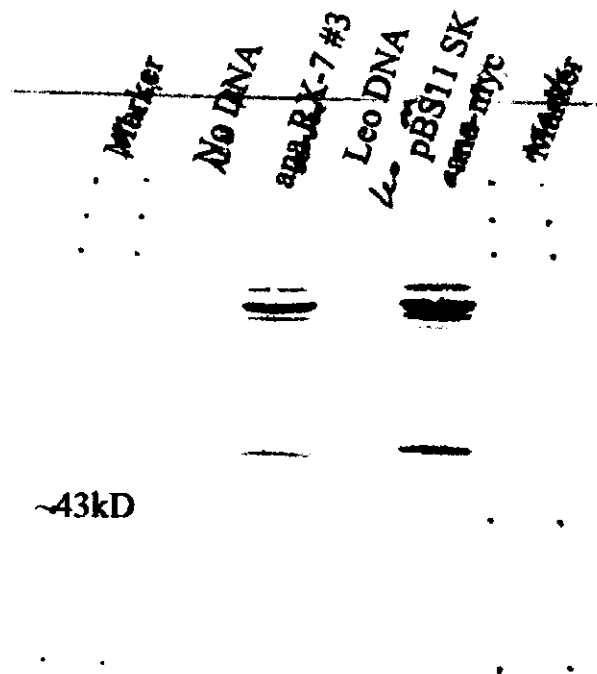themselves could have been bad, and decided to order new ones (Santa Cruz #sc-2004).

~43kD

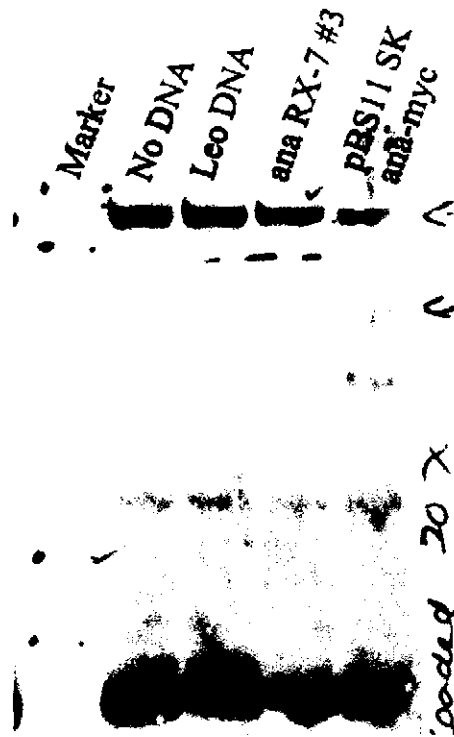Figure 5. Exposure of ^{35}S in nascent *in vitro* translation

Figure 6. Western blot of Leo DNA, *ana* RX-7#3, and pBS11SK*ana*-myc
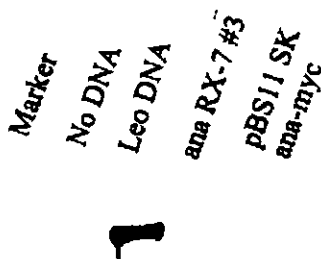
probed with anti-myc antibody

Figure 7. Antibody tagging using Leo antibody

We then proceeded to once again run Leo DNA, *ana* RX-7 #3, and pBS11 SK *ana*

myc in the Reticulocyte Lysate System. After we transferred this to a nitrocellulose

membrane, we treated with our new myc antibodies. This experiment produced exciting

results (Figure 8). The lane in which we expected to find our *ana*-myc protein stained with

the new myc antibody, and rough molecular weight calculations were close to expected

values (57Kd expected vs. 59Kd calculated). However, these results proved to be

irreproducible. Several subsequent Lysate System/Western blotting experiments showed no

reaction to the myc antibody, suggesting no *ana*-myc protein being made (Figures 9 and

10). However, Figure 10 does show weak bands across. In order to determine what these

bands could be, we used an antibody that recognizes Histidine repeats. We expected to see

the Histidine antibody react with *ana* and *ana*-myc proteins because *ana* has a 6-Histidine repeat at the c-terminal end. For this experiment, we stripped the membrane used in Figure 10 and treated with the Histidine antibody. Figure 11 shows that nothing on the membrane (also used for Figure 10) had affinity for the Histidine antibody. As a positive control, we ran a new gel using protein extracted from 1st, 2nd, and 3rd instar CS larvae (50 from each stage, 15uL samples) and *ana* protein samples previously made from the Reticulocyte Lysate System that we suspected had the protein. We know that *ana* protein is expressed in larvae stages and therefore should be present in the larvae samples. Figure 12 shows the two Reticulocyte Lysate samples of *ana* protein reacting with the Histidine antibody. However, the molecular weight determinations showed these bands to be around 40kD, illustrating the two samples could not have contained *ana* protein (~55kD). This shows that the weak bands in Figure 10 are not due to the presence of *ana* or *ana*-myc protein. Additionally, the Histidine antibody did not react with larvae extracts showing that either antibody (myc or Histidine) could be used in identification experiments.
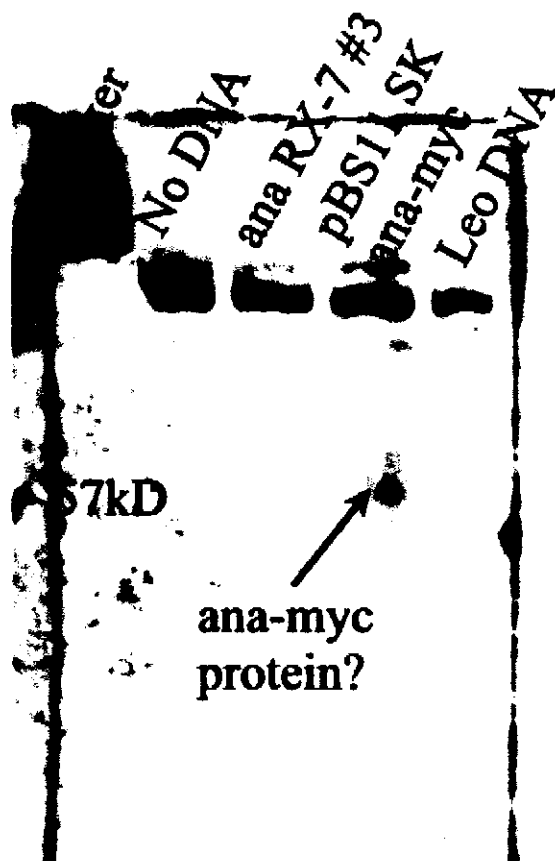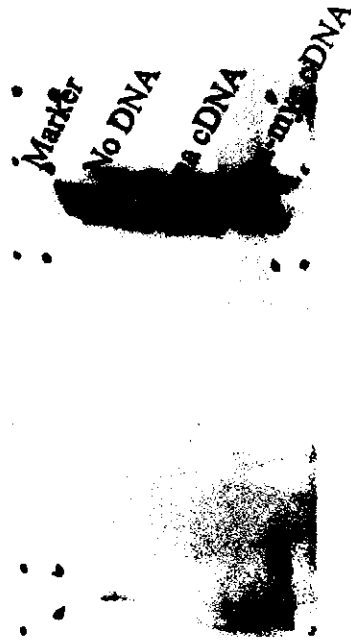
Figure 8. Positive results of Western Blot of *ana*-myc

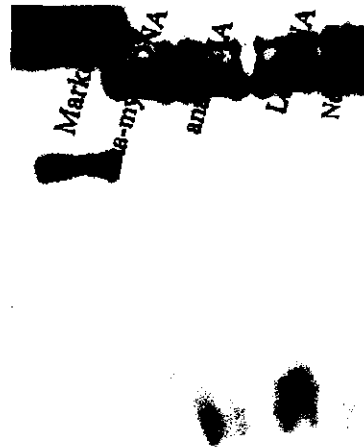Figures 9. Primary negative results of Western blot of *ana*-myc

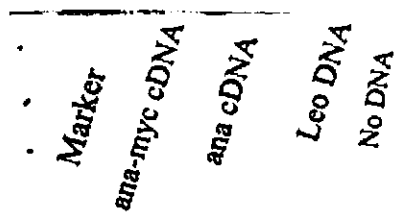Figure 10. Secondary negative results of Western blot of *ana*-myc

Marker ana-myc cDNA ana cDNA Leo DNA No DNA

Figure 11. Old membrane treated with Histidine antibody
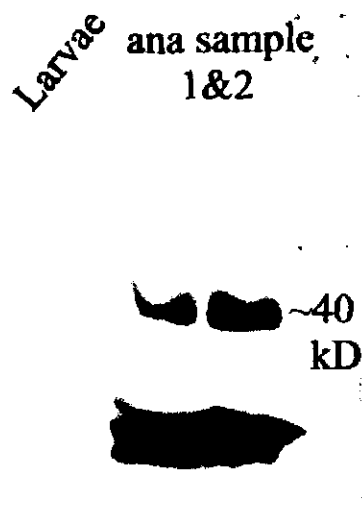
Larvae ana sample, 1&2

~40 kD

Figure 12. New membrane treated with Histidine antibody

The next system we utilized was a prokaryotic expression system called

pBAD/*Myc*-His. The results from the double restriction digests, using enzymes Eco R1 and

Hind III, are presented below. The pUC *ana* SB 39 and pBAD/*Myc*-His A digestion

products are shown in Figure 13. This picture shows the two DNA fragments: *ana* cDNA

fragment (1.6 Kb) and pBAD/*Myc*-His A fragment (~4 Kb). This experiment confirmed

that we excised the correct pieces of DNA and were then ready to ligate (Figure 14) the two

to produce a new vector. The new vector will be one in which the *ana* cDNA would be in

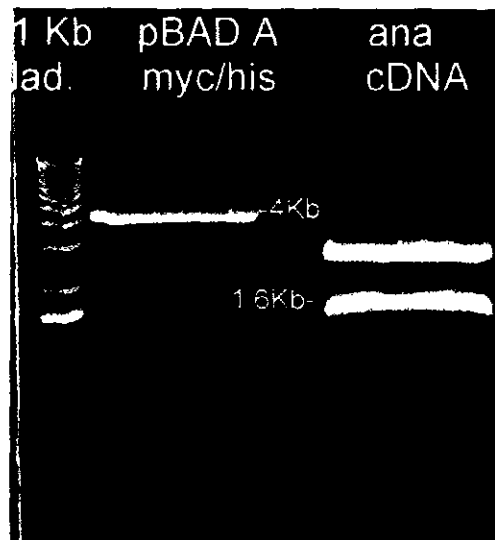front of and in frame with the myc epitope/polyhistidine region on the pBAD vector.

Figure 13. Double Digest of pUC *ana* SB 39 and pBAD/*Myc*-His A

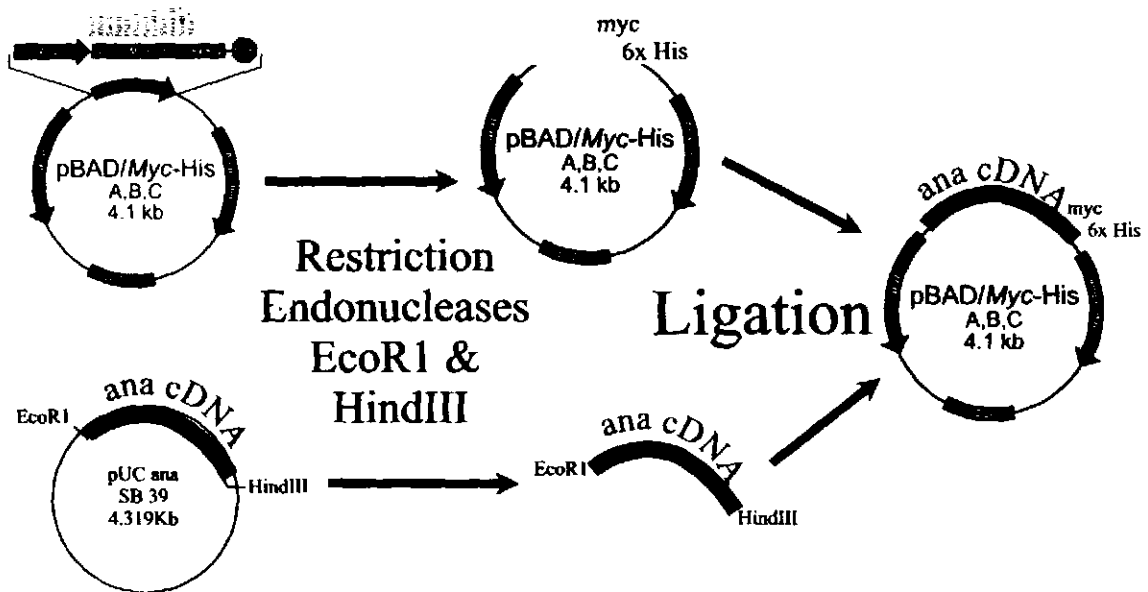Figure14. Ligation of *ana* cDNA fragment and pBAD/*Myc*-His A fragment

Once the ligation of the two DNA pieces was complete, confirmation of the two

fragments was done using enzymes Eco R1 and Hind III. The results of that double digest

are shown in Figure 15. These results show that we have ligated the correct pieces of DNA

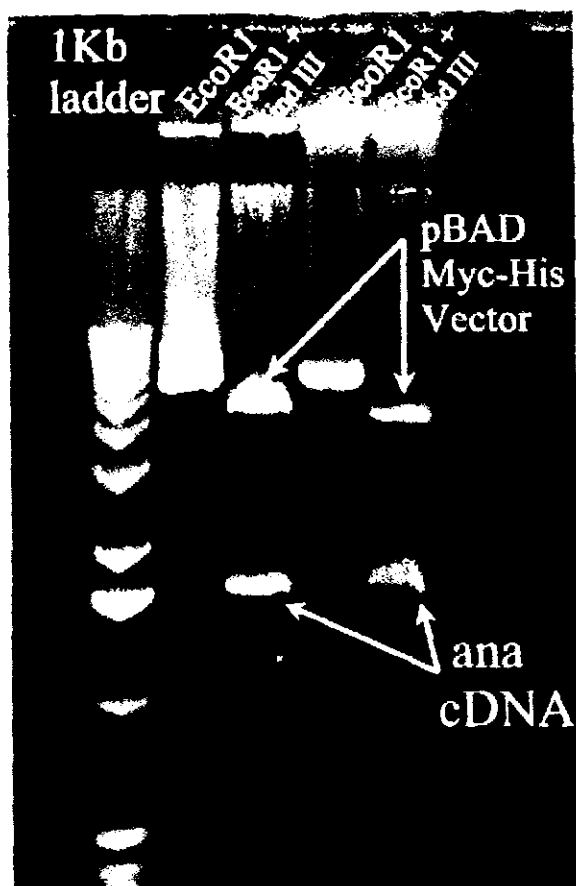together, and are now ready to induce production of *ana*-myc-his protein.

Figure 15. Double Digest of pBAD *Myc*-His/*ana* cDNA

Preliminary expression results using SDS-PAGE analysis and Coomassie blue staining show three possible *ana*-myc proteins (Figure 16). The positive control, *lacZ*, showed a stronger band after induction, signifying that the expression is working.
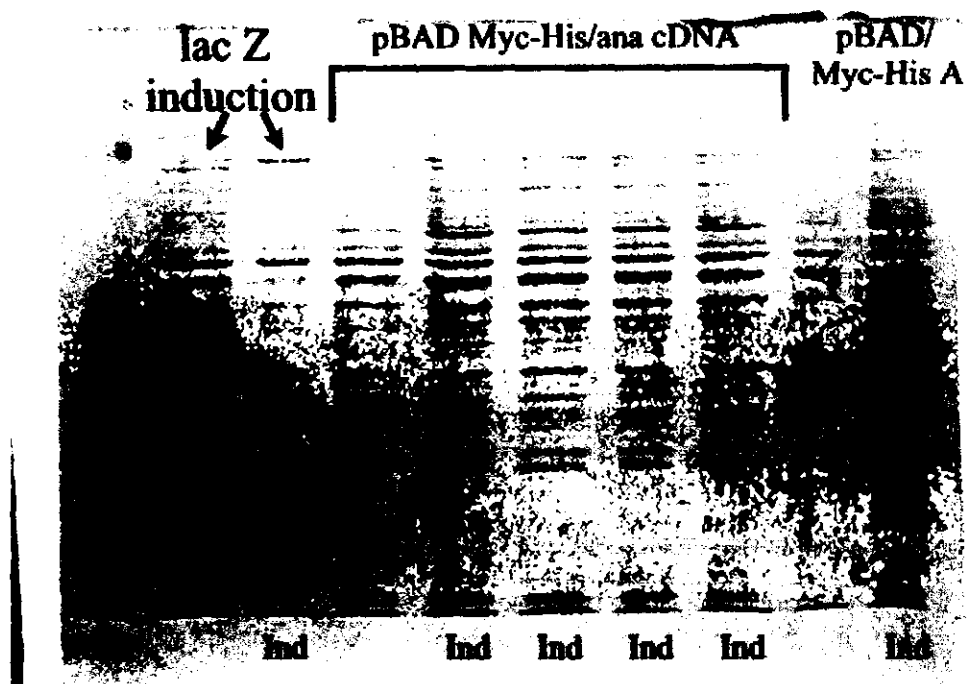
Figure 16. Preliminary expression results

CONCLUSIONS

General Considerations

Several key elements must be considered and maintained in order to synthesize a tagged protein. First, the correct cDNA sequence must be obtained for both the protein and the tag. Next, the two cDNA's must be joined in frame, so that the translation of the entire sequence results in the protein and tag being synthesized. Then, the question of whether the protein is fully functional arises. And if the protein is made and is functional, then the process must be reproducible.

Reticulocyte Lysate System

The Reticulocyte Lysate System would have been the most straightforward and simplest approach to making our labeled protein. However, the positive results (Figure 8) were unable to be confirmed. It turned out that the irreproducible results were for good reasons. Upon close analysis, the construct of *ana* cDNA and myc-epitope sequence (pBS11 SK *ana* myc) was not properly created. The source of myc epitope sequence used in this construct is uncertain. Also, the PCR technique used to change the stop codon may not have yielded the insert expected. As the consequence of Colleen Hough's research, the correct myc epitope sequence was not added in frame with the *ana* cDNA, thus making it impossible to synthesize an *ana* protein with a myc epitope.

pBAD/*Myc*-His System

The results of the double digest of pUC *ana* SB 39 and pBAD/*Myc*-His A (Figure 14) show that the correct fragments have been excised. We have then taken the *ana* cDNA fragment and cloned it into the pBAD/*Myc*-His vector. Analysis of the double digest of pBAD *Myc*-His/*ana* cDNA (Figure 15) shows that we have successfully ligated the correct two pieces of DNA together. This new vector contains the correct myc epitope sequence and is in frame with the *ana* protein. Preliminary expression results using SDS-PAGE analysis and Coomassie blue staining show three possible *ana*-myc proteins (Figure 16).

Future Research

Now that we have the proper construct of *ana* cDNA and myc-epitope sequence (correct sequence and in frame), we can transform this vector back in to *E.coli*. Once transformed, we will induce *E.coli* to transcribe and translate our protein using its own cellular machinery. The bacteria will transcribe and translate from the beginning of the *ana* cDNA sequence through the myc-epitope and poly-Histidine sequences, producing an *ana*-myc-polyHis protein. This protein should react with myc antibodies as well as poly-Histidine antibodies. Once this protein is made and easily identified, we will verify that it is fully functional.

This project set out to find the protein-protein interactions of *ana*. To do that, we had to devise a method for synthesizing a readily identifiable protein. That method proved to be somewhat more difficult than expected. At the moment we are not able to answer the questions set forth, but our results show we have made great progress in the direction of

those answers. Currently, Western blot analysis is underway to confirm that *ana* is in frame

with the myc epitope using the pBAD system. With continued research of this project, we

will soon be able identify *ana*'s interactions.

REFERENCES

Becker, T. and Technau, G. (1990). Single cell transplantation reveals interspecific cell communication in *Drosophila* chimeras. Development 109, 821-832.

Copp, A. J., Brook, F. A. and Roberts, H. J. (1988). A cell-type-specific abnormality of cell proliferation in mutant (curly tailed) mouse embryos developing spinal neural tube defects. Development 104, 285-295.

Datta, S. (1995). Control of proliferation activation in quiescent neuroblasts of the *Drosophila* central nervous system. Development 121, 1173-1182.

Datta, S. and Kankel, D. R. (1992). *l*(1)*trol* and *l*(1)*devl*, loci affecting the development of the adult central nervous system in *Drosophila melanogaster*. Genetics 130, 523-537.

Ebens, A. J., Garren, H., Cheyette, B. N. R. and Zipursky, S. L. (1993). The *Drosophila anachronism* locus: A glycoprotein secreted by glia inhibits neuroblast proliferation. Cell 74, 15-28.

Ikuta, K., Uchida, N., Friedman, J. and Weissman, I. L. (1992). Lymphocyte development from stem cells. Annual review of Immunology 10, 759-783.

Miller, R. W. (1988). Intrauterine radiation exposures and mental retardation. Health Physics 55, 295-298.

Park, Y., Caldwell, M. C. and Datta, S. (1997). Mutation of the central nervous system neuroblast proliferation repression *ana* leads to defects in larval olfactory behavior. J. Neurobiology

Ross, R., Spengler, B., Domenech, C., Porubcin, M., Rettig, W. and Biedler, J. (1995).
Human neuroblastoma I-type cells are malignant neural crest stem cells. Cell
Growth & Diff. 6, 449-456.

Truman, J. W. and Bate, M. (1988). Spatial and temporal patterns of neurogenesis in the
central nervous system of *Drosophila melanogaster*. Dev. Biol. 125, 145-157.

Vita

Name: Joseph Edward Amos

Permanent Address: 114 Stephanie Dr.
Palestine, TX 75801

Educational Background: Texas A & M University
Bachelor of Science Biochemistry, May 2001
Fall 1998-Spring 2001

Kilgore College
Summer 1997-Fall 1998

Honors: Phi Theta Kappa International Honor Society (1997); Golden Key National Honor Society (1999); Nominated for Pfizer Undergraduate Fellowship (2000); Texas A&M University Undergraduate Research Fellow (2000-01); Biochemistry Research Scholar (2000-01); Cherokee County Aggie Moms Scholarship (2000); TAMU Summer University Undergraduate Research Fellow (2000); Howard Hughes Honors Scholarship (2000); Julia Ball Lee Memorial Scholarship (2000); Houston Rodeo Endowed Scholarship (2000); Texas A&M Club of Albuquerque Scholarship (2000); Dean's Honor Roll (2000); Phi Kappa Phi Honor Society (2001); Nominated for Senior Merit Award (2001)

Research Experience: Texas A&M University Biochemistry/Biophysics Laboratory, Texas A&M University Undergraduate Research Fellows Program, Biochemistry Research Scholars Program