8-1999

Production of an Improved Vector for Synthesis and Purification of Eukaryotic Transcription Factors (GAL4-VP16, GAL4-AD and GAL4-D)

Pedro Oyarbide Valencia
Southern Illinois University Carbondale

Production of an Improved Vector for Synthesis and Purification of Eukaryotic Transcription Factors (GAL4-VP16, GAL4-AD and GAL4-D)

Pedro Oyarbide Valencia

August 4, 1999

Undergraduate Honors Thesis, Department of Microbiology

Laboratory of Medical Biochemistry

Southern Illinois University Carbondale

Carbondale, Illinois 62901 U.S.A.

Table of Contents

Introduction

In recent years, the study of eukaryotic gene expression has evolved due to technological developments that have greatly increased knowledge and understanding of the subject. There are many different types of molecules involved in the process of gene expression. Some of them have become critically important due to their enhanced expression activation properties. These are called transcription factors (TF), and the following study is focused on the sub-cloning and expression of the GAL4 TF series for the later study of their particular activities and the nature of their interactions in the chromatin remodeling process.

In general TF as well as other regulatory proteins, possess the following two abilities: (1) the ability to recognize specific target sequences located in the DNA template that belong to enhancers, promoters and other regulatory elements that affect a particular gene, and; (2) the ability, once bound to DNA, to exercise their function by binding to other components of the transcription apparatus. These TFs can be divided into two different groups according to their aminoacid residue composition. The first group includes general transcription factors, such as TFIIIA, TFIIIB, while the second group is made up of specific or acidic transcription factors such as GAL4 series. The latter possesses a higher percentage of acidic aminoacid residues in its primary structure.

Regarding this particular series of acidic transcriptional activators, a wide array of features have been attributed to them. This thesis mainly focuses on the GAL4-AD, GAL4-D and GAL4-VP16 transcriptional activators. The GAL4 series members are transcriptional activators (TA) of fungal origin, composed of several functional domains that include the following:

> a characteristic cysteine rich DNA binding domain common to all members of the family, a dimerization domain, various transactivation domains generally exhibiting a high acidic content and a highly variable central region, that is supposed to be involved in regulation and in effector recognition. (Poch 1997: 229)

One hypothesis gives credence to the existence of eight highly conserved functional motifs. These motifs are located in a large domain of approximately 225-405 residues. One researcher found that "its inhibitory activity may be mediated by hydrophobic interactions linked to the presence of amphipathic alpha-helices" (Poch 1997: 229). Therefore, the function of a transcription factor depends directly on the type of interaction of its different motifs. Differences in functional motifs constitute the differences between transcription factors. There is a molecular basis for the differentiation between TFs in terms of their ability to bind sites within chromatin. GAL4 is known to interact with its cognate site when it is assembled into nucleosome core particles, meaning that its affinity for nucleosomally organized target sites is lower than for the same sites in free DNA (Beato 1997: 3561). It has also been demonstrated that GAL4 derivatives can potentiate replication-independent transcription in vitro from preassembled chromatin templates. For instance, the GAL4-VP16 acidic activator is able to enhance transcription of a template assembled into chromatin in vitro if added only during, not after, DNA replication (Xu 1998: 1201). As a general trait, most eukaryotic gene promoters contain multiple binding sites for one or more transcriptional activators that interact in a synergistic manner. This is the manifestation of the need of many contacts between activators and the general transcription machinery. While normally most of the transcription activators, (TA), need several contacts to elicit the maximum level of

possible transcriptional activity, GAL4 can produce the same activation pattern with only one contact under physiological conditions (Xu 1995: 7677).

These features are most useful to researchers in understanding eukaryotic gene regulation. It is necessary to determine sequences within these TF that will interact with DNA and which are the target sequences that they bind preferentially. For this reason, this study investigates segments of TA. The GAL4-AD, a 34-amino acid region at the carboxy terminus end of the GAL4 TF (Melcher 1995: 2839) has been proved to specifically and directly bind to the TATA-binding protein (TBP). Other studies have shown that for GAL4 TF the palindromic CGG triplets at the ends of the 17-bp recognition site in DNA are essential for tight binding, whereas the identity of the remaining internal 11-bp are of much less importance (Liang 1996: 3773).

The study of both primary sequences and more complex sturctures in GAL4 derivatives has led to suppositions of several other hypotheses regarding their stability and the driving force for their interactions with DNA. For example, the binuclear metal ion structure present in these TF, is shown to have a net negative charge of -2. This fact has allowed for hypothesizing that this metal ion is the recipient of several hydrogen bonds, notably from the main chain amide protons of the ligating cysteine residues, indicating the charge is stabilized in this manner (Gadhavi 1997: 145).

Abstract

The main goal in conducting this series of experiments is to accomplish the construction of an expression vector that contains the coding regions for the GAL4 series transcription factors. This expression vector should possess special characteristics including a histidine tag that helps to easily locate the vector. The other main characteristic is a T7 promoter that allows for a high induction level upon addition of IPTG. The original pET21d expression vector was used in these experiments because it possesses these two intrinsic features. A sub-cloning process of the coding regions of the GAL4 series transcription factors (GAL4-AD, GAL4-D and GAL4-VP16) into the original pET21d expression vector resulted in the pET21d-GAL4 expression vector. The coding regions for the GAL4-AD and GAL4-D transcription factors were extracted from a pGEX-cs vector by Polymerase Chain Reaction (PCR) amplification with specially designed primers. The GAL4-VP16 transcription factor was extracted from the pJL2 plasmid. In the case of the GAL4-VP16 transcription factor, the sequence was not initially available and experimental sequencing of the transcription factor was necessary to acquire enough information about the sequence to design the primers for the PCR amplification[1].

[1] See Appendix A for sequences and genetic maps for the three different transcription factors, primers and expression vectors.

Materials and Methods[2]

A detailed description of the experimental procedures, the conditions in which the experiments were performed and the pertinent information about the materials used will clarify the specifics for conducting these experiments.

BMH Competent Cell Preparation (for Electroporation)[3]

1. Inoculate a single colony of BMH into 5 mL LB + 40 ul of tetracycline at 37° C overnight.
2. Inoculate 5 mL culture in 1 L LB grow at 37° C until the OD at 600 nm. Reaches the interval [0.5 - 0.7].
3. Chill on ice for 10 minutes. Pour into a 1 L centrifuge bottle that has been pre-chilled.
4. Pellet 1 L of culture at 4000 g at 4° C for 15 minutes.
5. Resuspend in 1 L of sterile, cold H_2O. Pellet at 4000 g at 4° C for 15 minutes.
6. Resuspend in 500 mL of sterile, cold H_2O. Pellet at 4000 g at 4° C for 15 minutes.
7. Resuspend in 500 mL of sterile, cold H_2O. Pellet at 4000 g at 4° C for 15 minutes.
8. Resuspend in 20 mL of sterile, cold H_2O. Pellet at 4000 g at 4° C for 15 minutes.
9. Resuspend in 2 mL sterile, cold 10% glycerol (filter sterilized).
10. Aliquots may be frozen at -70° C.

---

[2] A detailed listing of the experimental procedures used, as well as a listing of the main materials, are available in appendix B.

[3] Medical Biochemistry, Lab 207, June 14, 1998.

Dephosphorilation of the 5' End of the pET21d Vector[4]

1. Take the sample out of the -20° C storage and thaw.

Sample	1
pET21d vector, digested with NcoI and XhoI	40 ul
10X SAP buffer	4.5 ul
Shrimp Alkaline Phosphatase (1U/ul)	1 ul
Total volume	**45.5 ul**

2. Incubate the mixture for 60 minutes at 37° C.
3. Reload the same amount of enzyme after the first 60 minutes, and expose the mixture for another 60 minutes in the same conditions.
4. After 120 minutes of incubation, heat deactivate the shrimp alkaline phosphatase by introducing the mixture into a 65° C bath for 15 minutes.
5. Store the product at -20° C.

[4] Medical Biochemistry, Lab 207, July 1, 1998.

Preparation of the pET21d Vector for the Later Sub-Cloning of the GAL4 Transcription Factors[5]

1. Digest the purified sample with NcoI and XhoI.

SAMPLE	1
pET21d (1.183 mg/ul)	3 ul
2.5 X 11.14 U of NcoI	3 ul
2.5 X 44.5 U of XhoI	7 ul
100 ug/ul of ABS	0.5 ul
10% total volume of buffer #2	5 ul
DD H_2O	31.5 ul
Total volume	**50 ul**

2. Incubate the mixture for 120 minutes at 37°C.
3. Purify the product by spin column.
4. 100 ul S-200 resin equilibrated with TE (centrifuge 5000-6000 rpm).
5. Check on 1.5% agarose + 1X TBE gel.
   a) Uncut pET21d vector
   b) Digested pET21d vector

[5] Medical Biochemistry, Lab 207, May 24, 1998.

Transformation of Constructs into DH5alfa Competent Cells[6]

1. Set up the electro cell manipulator 600 as follows:

Choose mode: T	2.5KV/Resistance High Voltage (HV)
Set capacitance: C	Not used in HV mode
Set resistance: R	R5 (129 ohm)
	BTX Disposable cuvette P/N 620 (2mm gap), place on ice for five minutes
Set charging voltage: S	2.45 KV
Desired field strength:	12.25 KV/cm
Desired pulse length: I	5-6 mseconds

2. Remove 10 Eppendorf tubes of DH5alfa competent cells, prepared for electroporation and placed on ice.
3. Once the competent cells have thawed out, add the following DNA, to the competent cells:
    a) 1 ul of GAL4-AD ligation 1' product
    b) 1 ul of GAL4-D ligation 2' product
    c) 1 ul of GAL4-AD ligation 1 product
    d) 1 ul of GAL4-D ligation 2 product
    e) 1 ul of vector pET21d ligation product
    f) 1 ul of GAL4-AD insert ligation product
    g) 1 ul of unligated pET21d vector
    h) 1 ul of unligated GAL4-AD insert

[6] Medical Biochemistry, Lab 207, July 17, 1998.

i) 1 ul of T.E. pH 8.0

4. Flick the Eppendorf tube to mix and settle the cell mixture by tapping the cuvette on the table.
5. Transfer the competent cells into the BTX cuvette, right after mixing with DNA, and put the cap on.
6. Press A to activate automatic charge and pulse sequence.
7. Immediately after green charging light goes out, add 1000 ul of 2xYT into the cuvette.
8. Briefly and gently pipette out the content from the cuvette into a sterile, small-size Falcon tube and vortex.
9. Incubate the Falcon tubes at 37° C for 60 minutes. (Shaking at 225 rpm will improve recovery.)
10. Plate 100 ul from each of the Falcon tubes onto LB agar plates containing 40 ug/ml of ampicilin.
11. Allow the plates to dry.
12. Incubate the plates at 37° C overnight.
13. Check for growth of individual colonies.

Preparation of the GAL4-AD and GAL4-D Inserts[7]

1. PCR amplify the GAL4-AD and GAL-D fragments from the pGEX-cs vector.

SAMPLE	1	2
pGEX-cs vector (50ng/ul)	1 ul	1 ul
GAL4-AD 5' primer (50 pmol/ul)	2 ul	2 ul
GAL4-AD 3' primer (50pmol/ul)	2 ul	
GAL4-D 3' primer. (50pmol/ul.)		2 ul
DD H2O	76 ul	76 ul
10X Pfu buffer	10 ul	10 ul
2.5 mM dNTP	8 ul	8 ul
VORTEX SAMPLES		
2.5 U/ul Pfu DNA Polymerase.	1 ul	1 ul
Total volume	**100 ul**	**100 ul**

2. Program for PCR reaction:

   25 cycles: 1 minute at 94° C, 1 minute at 50° C, 4 minutes at 72° C

   1 cycle: 1 minute at 94° C, 1 minute at 50° C, 10 minutes at 72° C

3. Check results on a 0.5% agarose + 1X TBE gel.

[7] Medical Biochemistry, Lab 207, July 7, 1998.

GAL4-VP16 Sequencing Reaction[8]

1. Preparation of the sequencing gel:

   a) 8% Acrylamide

      10 ml of 40% acrylamide

      15 ml of E-pure H_2O

      5 ml of 5 x TBE

      25 g of urea

   b) Mix gently until the urea is dissolved and filter through a 0.45 micron filter.

      100 ul of TEMED

      100 ul of APS

   c) Pour the gel.

2. Sequencing reaction:

   a) Denaturing reaction.

Sample.	1	2
GAL4-VP16 (bigger prep.)	8 ul	8 ul
1.0 M NaOH	2 ul	2 ul
PGEX-cs 3' (5pmol/ul)		1 ul
PGEX-cs down (5pmol/ul)	1 ul	

   b) Mix thoroughly and incubate for 10' at 37 °C. Then place the mixture on ice and add to each of the tubes:

[8] Medical Biochemistry, Lab 207, July 16, 1998.

Sample.	1	2
1.0 M HCL	2 ul	2 ul
Plasmid reaction buffer	2 ul	2 ul
Total volume	15 ul	15 ul

c) Annealing reaction: Incubate the template/primer/buffer mixture at 37° C for 10 minutes, then chill on ice.

d) While annealing, label, fill and cap tubes with 2.5 ul of each of the termination mixtures (G, A,T and C, red capped vials). Keep covered on ice for steps 5-7.

e) If needed, dilute labeling mix 15-fold to working concentration. Retain for use in step 6.

Dilution	
Labeling mix	4 ul
DD H_2O	16 ul

f) Pre-warm the termination tubes from step 2(d) in a water bath at 37° C.

g) Labeling reaction: Add to ice-cold DNA from step 2(c). Mix and incubate at room temperature for 2-5 minutes.

Labeling	
DTT (0.1 M)	1 ul
Diluted labeling mix	2 ul
5 uCi{s35} dATP	0.5 ul
Sequenase polymerase	2 ul

3. Termination reaction: Transfer 4.5 ul of the labeling reaction into each of the termination tubes (G, A,T and C), mix and continue the incubation of the termination tubes at room temperature for 5 minutes.
   a) Stop the reaction by adding 4 ul of stop solution.
   b) Heat the samples to 75° C for 2 minutes immediately before loading onto the sequencing gel. Load 2-3 ul in each lane.

GAL4-VP16 Transcription Factor DNA Sequencing Reaction[9]

1. Preparation of the sequencing gel:

   a) 8% Acrylamide

      10 ml of 40% acrylamide

      15 ml of E-pure H_2O

      5 ml of 5 x TBE

      25 g of urea

   b) Mix gently until the urea is dissolved and filter through a 0.45 micron filter.

      100 ul of TEMED

      100 ul of APS

   c) Pour the gel.

2. Sequencing reaction: Sequencing for GAL4-VP16 by Stratagene Cyclist Exo-Pfu DNA sequencing kit.

   a) Add 3 ul of the appropriate ddNTPs to 4 termination tubes chilled on ice.

Sample	1
GAL4-VP16 (0.6 ug/ul)	2 ul
PGEX-cs down(5pmol/ul)	1 ul
10 X sequencing buffer	4 ul
S35-dATP (12.5uCi/ul)	1 ul
Exo-Pfu enzyme	1 ul
DD H_2O	17 ul
DMSO	4 ul
Total volume	**30 ul**

   b) Aliquote 7 ul of the sequencing reaction mixture into each of 4 the termination tubes.

[9] Medical Biochemistry Lab 207, May 19, 1998.

3. Program for PCR reaction:

   1 cycle: 95° C for 30 minutes.

   30 cycles: 95° C for 30 seconds. 60° C for 30 seconds. 72° C for 30 seconds.

   a) Stop the reactions by adding 5 ul of the stop solution.

   b) Heat samples at >80° C for 2-5 minutes. Then immediately load 2-4 ul onto sequencing gel.

BIGGERprep Plasmid DNA Preparation Procedure

1. Culture plasmid containing bacterial cells by inoculating 1 L of media of TB broth media with the appropriate concentration of antibiotics for 16-24 hours.
2. Pellet bacteria by centrifuging at 10,000 g for 3 minutes. Decant the supernatant.
3. Completely resuspend the bacterial pellet in 30 mL of solution I. Vortex to ensure pellet resuspension.
4. Lyse the bacteria by adding 30 mL of solution II to the resuspended bacteria. Mix very gently by rotating the centrifuge tube. Do not allow the solution to be active for more than 5 minutes. Lysate should be relatively clear.
5. Neutralize bacterial lysate by adding 30 mL of solution III. Immediately mix by extremely gentle inversion of the bottle for 5-10 seconds. A white material will appear. Increase the strength of the inversions, but do not vortex.
6. Pellet the bacterial cell wall by centrifuging at 16,000 g for 8 minutes. Transfer all the supernatant to a fresh centrifuge tube.
7. Bind the plasmid DNA by adding 80 mL of well-mixed BIGGERprep DNA Binding Matrix Suspension to supernatant and mix by vigorous inversion.
8. Pellet BIGGERprep DNA Binding Matrix and bound plasmid DNA at 16,000 g for 8 minutes. Pour off the supernatant making sure to keep the pellet.
9. Add 20 mL of diluted Purification Solution and mix by vigorous shaking. Purification solution is a 1:1 mixture of the reagent and 95% ethanol. Pour mixture into a BIGGERprep spin column.
10. Add 20 mL of the same mixture to rinse any possible residues from the centrifuge tube.

11. Purify the plasmid DNA by centrifuging at 1500 g for 7 minutes on a swinging-bucket centrifuge at room temperature. Pour off the filtrated liquid.
12. Repeat the same step for greater purity.
13. Transfer the BIGGERprep spin column to a fresh tube and elute the plasmid with 3 mL of 8.0 pH TE, preheated at 65-70° C.
14. Spin the column at 1500 g for 7 minutes. Transfer the elute plasmid DNA to a precipitation tube.
15. Precipitate the plasmid DNA by adding 140 ul of 5M NaCl, mix and then add 12 mL of 95%-100% ethanol at room temperature.
16. Mix well and pellet plasmid DNA at 16,000 g for 5 minutes at room temperature. Pour off supernatant.
17. Briefly wash with >70% ethanol. Wash twice and then let air dry.
18. Once the plasmid DNA is dry, resuspend in 500 ul of TE.
19. Store at -20° C.

GAL4-VP16 Transcription Factor Amplification (Taq Polymerase Protocol)

1. PCR amplify the GAL4-VP16 fragment from the pJL2 expression vector vector.

SAMPLE	1 and 2	3 and 4
pJL2 vector (50ng/ul)	1 ul	1 ul
50 mM $MgCL_2$	1.5 ul	1.5 ul
GAL4-VP16 5' primer (50pmol/ul)	1 ul	1 ul
GAL4-VP16 3' primer. (50pmol/ul.)	1 ul	1 ul
DD H2O	34.5 ul	34.5 ul
10X Taq extender buffer	5 ul	5 ul
10 mM dNTP	4 ul	4 ul
VORTEX SAMPLES		
2.5 U/ul Taq DNA Polymerase Exterder	1 ul	1 ul
2.5 U/ul Taq DNA Polymerase	1 ul	1 ul
Total volume	**50 ul**	**50 ul**

2. Program for PCR reaction:

   1 cycle: 30 seconds at 94° C

   35 cycles: 30 seconds at 94° C, 30 seconds at 55° C, 30 seconds at 72° C

   1 cycle: 7 minutes at 72° C

3. Check results on a 1.5% agarose + 1X TBE gel.

GAL4-VP16 Transcription Factor Amplification (Pfu Turbo Protocol)

1. PCR amplify the GAL4-VP16 fragment from the pJL2 expression vector vector.

SAMPLE	1 and 2	3 and 4
pJL2 vector (50ng/ul)	1 ul	1 ul
GAL4-VP16 5' primer (50pmol/ul)	1 ul	1 ul
GAL4-VP16 3' primer. (50pmol/ul.)	1 ul	1 ul
DD H2O	31.5 ul	31.5 ul
10X Cloned Pfu buffer	5 ul	5 ul
2.5 mM dNTP	9 ul	9 ul
VORTEX SAMPLES		
2.5 U/ul Pfu Turbo DNA Polymerase	1.5 ul	1.5 ul
Total volume	**50 ul**	**ul**

2. Program for PCR reaction:

   1 cycle: 30 seconds at 94° C

   35 cycles: 30 seconds at 94° C, 30 seconds at 55° C, 30 seconds at 72° C

   1 cycle: 7 minutes at 72° C

3. Check results on a 1.5% agarose + 1X TBE gel.

Ligation of the Dephophorilated pET21d Vector and the GAL4-AD and GAL4-D Transcription Factors.

1. Preparation of the ligation reactions:

Sample	1 and 2	3 and 4	5	6
pET21d vector (100ng/µl)	2 µl	2 µl	2 µl	2 µl
GAL4-AD insert	2 µl			2 µl
GAL4-D insert		2 µl		
10 X T4 DNA ligase buffer	1µl	1µl	1µl	1µl
T4 DNA Ligase	1µl	1µl	1µlv	1µlv
DD H_2O	4 µl	4 µl	4 µl	4 µl
TOTAL VOL.	**10 µl**	**10 µl**	**10 µl**	**10 µl**

2. Place the reaction mixture at a constant 16° C for an incubation period of 4 hours.

3. Store the ligation results at – 20° C.

4. Check the results on a 1.5% Agarose gel using a 1 Kb DNA ladder.

Digestion of the GAL4-AD and GAL4-D Inserts with NcoI and XhoI Restriction Enzymes.

1. Preparation of the digestion reactions of the GAL4-AD and GAL4-D inserts.

2. Extract the bands from the gel utilizing a gel extraction kit.

Sample	1	2
GAL4-AD insert	20 µl	
GAL4-D insert		20 µl
Buffer # 2	3.5 µl	3.5 µl
NcoI Restriction Enzyme	2.5 µl	2.5 µl
XhoI Restriction Enzyme	2.5 µl	2.5 µl
BSA (100mM)	3.5 µl	3.5 µl
DD H_2O	3 µl	3 µl
TOTAL VOL.	35 µl	35 µl

3. Incubate the reaction mixture at 37° C for four hours.

4. Inactivate the restriction enzymes by heating the reaction mixture at 65° C for 15 minutes

5. Check the results on a 1.5% Agarose gel.

TB Broth Media Preparation

1. Ingredients for preparation of 1 L of media:

Tryptophan	10 g
Yeast extract	15 g
NaCl	5 g
d H_2O	700 mL

2. Adjust to a final pH of 7.0 and equal volume to 1 L.
3. Sterilize by autoclaving for 20 minutes.
4. Add 100 ng/ul of ampiciline when the temperature is about 55° C.

2 x YT Media Preparation

1. Ingredients for preparation of 1 L of media:

Bacto-tryptone	16 g
Bacto-yeast extract	10 g
NaCL	5 g
dH_2O	900 mL

2. Adjust to a final pH of 7.0 and equal volume to 1 L.
3. Sterilize by autoclaving for 20 minutes.
4. Add 100 ng/ul of ampiciline when the temperature is about 55° C.

Experimental Results[10]

The restriction digestion of the expression vector pET21d, in which the desired constructs were inserted, showed a unique band on a 1.5% agarose gel, corresponding to a fragment of approximately 5350 ± 10 base pairs resulting from the digestion with the restriction enzymes: NcoI and XhoI. These restriction enzymes have specific cut sites at 1 base pair and 5365 base pairs respectively. The resulting fragment was then purified by spincolumn. As a comparison, a separate sample was run as a control, containing an undigested version of the vector. The results from this purification process can be seen in Figure #1, together with that of the pET21d expression vector digestion. To avoid further complications with the expression vector, this was dephosphorilated at its 5' end.

The preparation of the inserts by means of PCR amplification of the coding region for the transcription factors GAL4-AD and GAL4-D, utilizing the GAL4-AD 5', GAL4-AD 3' and GAL4-D 3' primers resulted in fragments of approximately 390 base pairs and 330 base pairs respectively, after digestion with restriction enzymes NcoI and XhoI, as shown in Figure #2. After the amplification, the working concentrations of both of the inserts are 0.97ug/ul for GAL4-AD and 0.93ug/gl for GAL4-D. 20 µl from the amplification products were loaded into a 1.5% agarose gel from which they were later extracted and utilized for the digestion step. Figure #2 shows the gel from which they were extracted. The results from the ligation of the dephosphorilated pET21d expression vector and the GAL4-AD and Gal4-D constructs and the later electroporation are satisfactory. The growth of the trasformants indicates a successful insertion of the plasmid. The controls used indicate growth of the transformant with the unligated plasmid but there is no growth of the bacteria transformed with the GAL4-

[10] The results will be noted in chronological order of experiments.

AD insert alone, therefore resulting in a satisfactory ligation and trasformation. The results from the ligation experiment can be seen in Figure #3. The restriction digestion analysis from several of the transformed colonies indicates the presence of a band of approximately 5700 base pairs.

The results from the GAL4-VP16 sequencing have rendered an incomplete sequence for this transcription factor. The experimental sequence can be seen in appendix A. There has been no possibility of designing a DNA primer corresponding to the 5' end of the coding region of this transcription factor for the amplification procedure.

Using an alternative source of DNA containing the GAL4-VP16 insert, it has been possible to design the primers to be used in the amplification of this transcription factor. Also, the sequence for the GAL4-VP16 is available in Appendix A.

The bacterial culture containing the GAL4-VP16 insert has been grown and DNA extracted. The concentration of the working samples are 146 ng/µl and 176 ng/µl. The amplification of this transcription factor has been attempted using two different DNA polymerases. The first amplification has been performed using Taq polymerase and Taq extender. This amplification has resulted in good concentrations of amplified products, see Figure #4. The second amplification has been performed using Pfu Turbo as the DNA polymerase, the results from this latter amplification are more satisfactory in terms of yield and can be seen in Figure #5.

For a figure containing all the transcription factors together in an uncut GAL4-VP16 and digested versions of GAL4-AD and GAL4-D, refer to Figure #6.

Discussion

As a result of the satisfactory outcome obtained from the several experiments performed, a pET21d-GAL4 expression vector was obtained for each of the transcription factors (GAL4-AD and GAL4-D). The ultimate products from the transformation of the ligated plasmid-insert complex into *Escherichia coli* were kept as DNA stocks for future transformations. This method is preferred for DNA storage due to the relatively short life and much faster degradation rate of the bacterial cells. When transformed again into competent cells, they will enable an easy over-expression of the transcription factors in the future for their use in chromatin remodeling studies.

First of all, the resulting fragment from the pET21d expression vector digestion with NcoI and XhoI restriction enzymes, resulted in a unique band corresponding to a fragment of approximately 5400 base pairs approximately. This fragment is the expected outcome. The restriction enzyme NcoI cuts the expression vector at position 5365 and the restriction enzyme XhoI cuts at the expression vector at position 1. After digestion with both of these restriction enzymes, the expression vector is linearized in a fragment of 5364 base pairs, as indicated by experimental results, and there is a smaller fragment of approximately 75 base pairs. This fragment is too small to be detected with the size of the ladder used for the detection of our fragment of interest. The purification process as seen in Figure #1 maintains the size of the fragment. The fact that no other bands corresponding to fragments of similar sizes are apparent is another good indicator of the success of the digestion and purification processes. As a comparison tool, the sample has been run together with the undigested version of the vector that comprises 5440 base pairs in its entirety. A dephosphorilation step was performed on the purified sample of the pET21d expression vector with the sole intention of preventing

the vector from recircularizing, before the ligation experiment with the inserts was completed. The subsequent satisfactory ligation indicates that the concentration of the dephosphorilated vector was high enough to allow for a good ligation.

On the other hand, the amplification experiments have proved successful as well. The results from the amplification can be seen in Figure #2. We can see how the bands corresponding to the GAL4-AD transcription factor run about 390 base pairs, whereas those corresponding to the GAL4-D transcription factor are an average of 60 to 70 base pairs shorter, running about 320 base pairs. The reason for this is that the GAL4-D (domain) transcription factor lacks the last 60 base pairs that encode for the last 20 amino acids present in the GAL4-AD transcription factor sequence. As can be seen in Figure #3 the ligation products revealed the presence on a 1.5% agarose gel of fragments of two sizes. One is of approximately 5750 base pairs corresponding to the ligation of the pET21d expression vector and the GAL4-AD transcription factor insert. The other fragment is of approximately 5680 to 5690 base pairs corresponding to the ligation product of the expression vector and the GAL4-D transcription factor insert. The ligation experiment yielded a high enough concentration of ligated plasmid-insert complex to allow for a successful electroporation of the products into DH5 alpha competent cells. The controls used in the ligation experiment were also used in the transformation step. As expected, the negative control containing the products from the ligation of the GAL4-AD insert alone resulted in no growth on a LB broth plate containing ampicilin. The positive control consisted of the pET21d expression vector alone. After transformation the cells grew normally on the LB broth plates. The transformants resulting from the ligation of pET21d expression vector and the insert corresponding to the GAL4 series transcription factors all resulted in growth on the LB broth plates. These results are

satisfactory and the products have been kept as DNA stocks that can be used for later transformations into DH5 alpha competent cells for GAL4 proteins over-expression purposes. The purified version of these over-expressed proteins can be integrated into the study of chromatin remodeling.

The results from the sequencing of the GAL4-VP16 acidic transcription factor have provided only part of the sequence corresponding to the 3' end of the open reading frame containing this transcription factor. Therefore, no design of a primer covering the 5' region of the open reading frame was possible. The reason for this seems to be an advance stage of degradation of the DNA that has been used as a template for the sequencing experiments. The obtained partial sequence can be seen in Appendix A.

Utilizing a new *Escherichia coli* strain that contains the pJL2 expression vector, DNA has been isolated and the design of the necessary primers for the PCR amplification has been performed. To ensure a positive amplification of this transcription factor, two separate PCR reactions have been run. The first reaction consists of a combination of Taq polymerase and Taq polymerase extender. As mentioned in the results, the amlification obtained in this series of reactions is satisfactory. The second method is based on a PCR amplification using Pfu Turbo as the DNA polymerase and the results from this amplification are much more successful. The figures corresponding to these amplifications are Figure #4 and Figure #5 containing the Taq polymerase and the Pfu Turbo amplifications respectively.

In Figure #6 we can see the comparison of the different transcription factors in their digested version, with the exception of the GAL4-VP16 which is only present in its undigested version. As can be seen in the case of the GAL4-VP16 transcription factor there is more than one amplification product.

The project has rendered satisfactory results with respect to the cloning of the GAL4-AD and GAL4-D transcription factors. The obtained fragments matched the expected sizes after the PCR amplification step at approximately 390 base pairs for the GAL4-AD and 330 base pairs for the GAL4-D. Also, after the cloning step, the resulting fragments run about 5750 base pairs and 5680 base pairs respectively. These approximations calculated from the gels indicate that the cloning is acceptable. On the other hand, the GAL4-VP16 transcription factor amplification has shown positive results in both protocols that have been used for its amplification. The Taq polymerase amplified inserts are of the expected size of approximately 230 base pairs and the yield is considerably good, making it easier to digest and use in the cloning step. The second amplification performed with Pfu Turbo has resulted in extremely good amplification and there are different fragments. One of the fragments corresponds to the desired size while the others are slightly larger in size than expected as calculated from the genetic map of the pJL2 plasmid expression vector. These amplification products are excellent for the digestion protocol, and this step would allow for the isolation of the fragment that could be purified by spin column for a subsequent cloning step.

Appendix A

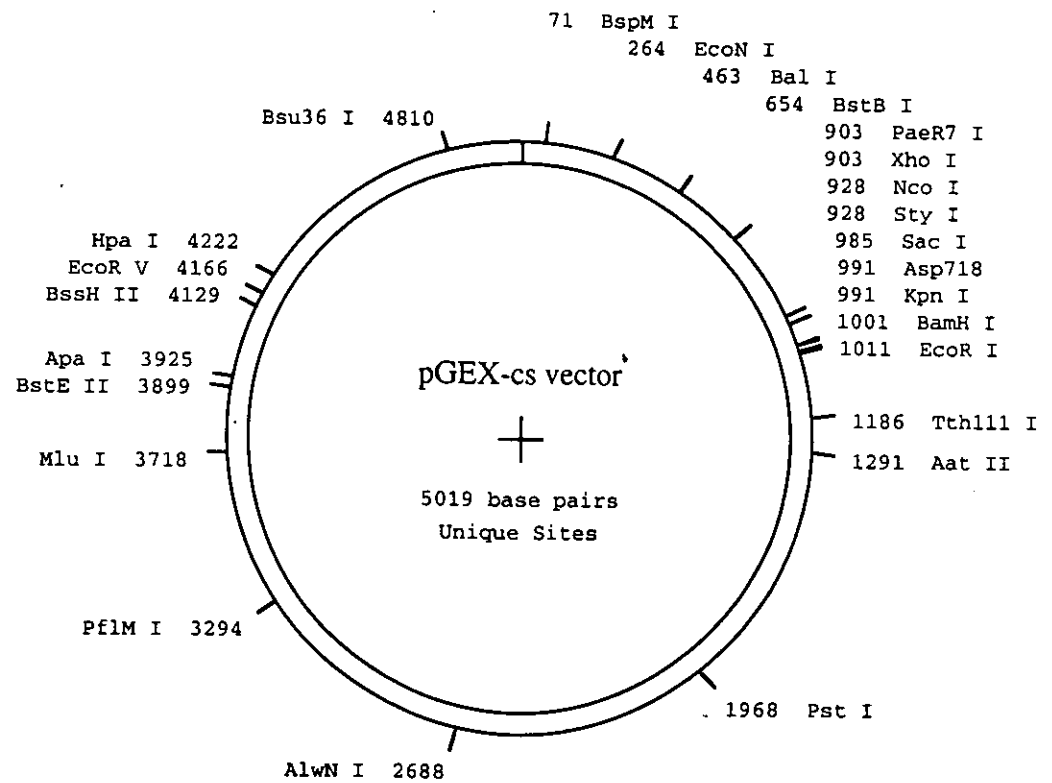
pGEX-cs Vector Linear Sequence

DNA sequence 5019 b.p. ACGTTATCGACT ... GGCGTTGGAATT circular

```
        |  10      |  20      |  30      |  40      |  50      |  60
   1 ACGTTATCGA CTGCACGGTG CACCAATGCT  TCTGGCGTCA GGCAGCCATC GGAAGCTGTG 60
  61 GTATGGCTGT GCAGGTCGTA AATCACTGCA  TAATTCGTGT CGCTCAAGGC GCACTCCCGT 120
 121 TCTGGATAAT GTTTTTTGCG CCGACATCAT  AACGGTTCTG GCAAATATTC TGAAATGAGC 180
 181 TGTTGACAAT TAATCATCGG CTCGTATAAT  GTGTGGAATT GTGAGCGGAT AACAATTTCA 240
 241 CACAGGAAAC AGTATTCATG TCCCCTATAC  TAGGTTATTG GAAAATTAAG GGCCTTGTGC 300
 301 AACCCACTCG ACTTCTTTTG GAATATCTTG  AAGAAAAATA TGAAGAGCAT TTGTATGAGC 360
 361 GCGATGAAGG TGATAAATGG CGAAACAAAA  AGTTTGAATT GGGTTTGGAG TTTCCCAATC 420
 421 TTCCTTATTA TATTGATGGT GATGTTAAAT  TAACACAGTC TATGGCCATC ATACGTTATA 480
 481 TAGCTGACAA GCACAACATG TTGGGTGGTT  GTCCAAAAGA GCGTGCAGAG ATTTCAATGC 540
 541 TTGAAGGAGC GGTTTTGGAT ATTAGATACG  GTGTTTCGAG AATTGCATAT AGTAAAGACT 600
 601 TTGAAACTCT CAAAGTTGAT TTTCTTAGCA  AGCTACCTGA AATGCTGAAA ATGTTCGAAG 660
 661 ATCGTTTATG TCATAAAACA TATTTAAATG  GTGATCATGT AACCCATCCT GACTTCATGT 720
 721 TGTATGACGC TCTTGATGTT GTTTTATACA  TGGACCCAAT GTGCCTGGAT GCGTTCCCAA 780
 781 AATTAGTTTG TTTTAAAAAA CGTATTGAAG  CTATCCCACA AATTGATAAG TACTTGAAAT 840
 841 CCAGCAAGTA TATAGCATGG CCTTTGCAGG  GCTGGCAAGC CACGTTTGGT GGTGGCgacc 900
 901 atctcgagaa tctttatttt cagggcgcca  tggCACTCAT CTTTGGCACA GTCAACGCTA 960
 961 ACATCCTGAA GGAAGTGTTC GGTGGAGCTC  GGTACCCCGG GGatcccggg gaattcATCG 1020
1021 TGACTGACTG ACGATCTGCC TCGCGCGTTT  CGGTGATGAC GGTGAAAACC TCTGACACAT 1080
1081 GCAGCTCCCG GAGACGGTCA CAGCTTGTCT  GTAAGCGGAT GCCGGGAGCA GACAAGCCCG 1140
1141 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG  TCGGGGCGCA GCCATGACCC AGTCACGTAG 1200
1201 CGATAGCGGA GTGTATAATT CTTGAAGACG  AAAGGGCCTC GTGATACGCC TATTTTTATA 1260
1261 GGTTAATGTC ATGATAATAA TGGTTTCTTA  GACGTCAGGT GGCACTTTTC GGGGAAATGT 1320
1321 GCGCGGAACC CCTATTTGTT TATTTTTCTA  AATACATTCA AATATGTATC CGCTCATGAG 1380
1381 ACAATAACCC TGATAAATGC TTCAATAATA  TTGAAAAAGG AAGAGTATGA GTATTCAACA 1440
1441 TTTCCGTGTC GCCCTTATTC CCTTTTTTGC  GGCATTTTGC CTTCCTGTTT TTGCTCACCC 1500
1501 AGAAACGCTG GTGAAAGTAA AAGATGCTGA  AGATCAGTTG GGTGCACGAG TGGGTTACAT 1560
1561 CGAACTGGAT CTCAACAGCG GTAAGATCCT  TGAGAGTTTT CGCCCCGAAG AACGTTTTCC 1620
1621 AATGATGAGC ACTTTTAAAG TTCTGCTATG  TGGCGCGGTA TTATCCCGTG TTGACGCCGG 1680
1681 GCAAGAGCAA CTCGGTCGCC GCATACACTA  TTCTCAGAAT GACTTGGTTG AGTACTCACC 1740
1741 AGTCACAGAA AAGCATCTTA CGGATGGCAT  GACAGTAAGA GAATTATGCA GTGCTGCCAT 1800
1801 AACCATGAGT GATAACACTG CGGCCAACTT  ACTTCTGACA ACGATCGGAG GACCGAAGGA 1860
1861 GCTAACCGCT TTTTTGCACA ACATGGGGGA  TCATGTAACT CGCCTTGATC GTTGGGAACC 1920
1921 GGAGCTGAAT GAAGCCATAC CAAACGACGA  GCGTGACACC ACGATGCCTG CAGCAATGGC 1980
1981 AACAACGTTG CGCAAACTAT TAACTGGCGA  ACTACTTACT CTAGCTTCCC GGCAACAATT 2040
2041 AATAGACTGG ATGGAGGCGG ATAAAGTTGC  AGGACCACTT CTGCGCTCGG CCCTTCCGGC 2100
2101 TGGCTGGTTT ATTGCTGATA AATCTGGAGC  CGGTGAGCGT GGGTCTCGCG GTATCATTGC 2160
2161 AGCACTGGGG CCAGATGGTA AGCCCTCCCG  TATCGTAGTT ATCTACACGA CGGGGAGTCA 2220
2221 GGCAACTATG GATGAACGAA ATAGACAGAT  CGCTGAGATA GGTGCCTCAC TGATTAAGCA 2280
2281 TTGGTAACTG TCAGACCAAG TTTACTCATA  TATACTTTAG ATTGATTTAA AACTTCATTT 2340
2341 TTAATTTAAA AGGATCTAGG TGAAGATCCT  TTTTGATAAT CTCATGACCA AAATCCCTTA 2400
2401 ACGTGAGTTT TCGTTCCACT GAGCGTCAGA  CCCCGTAGAA AAGATCAAAG GATCTTCTTG 2460
2461 AGATCCTTTT TTTCTGCGCG TAATCTGCTG  CTTGCAAACA AAAAAACCAC CGCTACCAGC 2520
2521 GGTGGTTTGT TTGCCGGATC AAGAGCTACC  AACTCTTTTT CCGAAGGTAA CTGGCTTCAG 2580
2581 CAGAGCGCAG ATACCAAATA CTGTCCTTCT  AGTGTAGCCG TAGTTAGGCC ACCACTTCAA 2640
2641 GAACTCTGTA GCACCGCCTA CATACCTCGC  TCTGCTAATC CTGTTACCAG TGGCTGCTGC 2700
2701 CAGTGGCGAT AAGTCGTGTC TTACCGGGTT  GGACTCAAGA CGATAGTTAC CGGATAAGGC 2760
2761 GCAGCGGTCG GGCTGAACGG GGGGTTCGTG  CACACAGCCC AGCTTGGAGC GAACGACCTA 2820
2821 CACCGAACTG AGATACCTAC AGCGTGAGCT  ATGAGAAAGC GCCACGCTTC CCGAAGGGAG 2880
2881 AAAGGCGGAC AGGTATCCGG TAAGCGGCAG  GGTCGGAACA GGAGAGCGCA CGAGGGAGCT 2940
2941 TCCAGGGGGA AACGCCTGGT ATCTTTATAG  TCCTGTCGGG TTTCGCCACC TCTGACTTGA 3000
3001 GCGTCGATTT TTGTGATGCT CGTCAGGGGG  GCGGAGCCTA TGGAAAAACG CCAGCAACGC 3060
3061 GGCCTTTTTA CGGTTCCTGG CCTTTTGCTG  GCCTTTTGCT CACATGTTCT TTCCTGCGTT 3120
3121 ATCCCCTGAT TCTGTGGATA ACCGTATTAC  CGCCTTTGAG TGAGCTGATA CCGCTCGCCG 3180
3181 CAGCCGAACG ACCGAGCGCA GCGAGTCAGT  GAGCGAGGAA GCGGAAGAGC GCCTGATGCG 3240
3241 GTATTTTCTC CTTACGCATC TGTGCGGTAT  TTCACACCGC ATAAATTCCG ACACCATCGA 3300
3301 ATGGTGCAAA ACCTTTCGCG GTATGGCATG  ATAGCGCCCG GAAGAGAGTC AATTCAGGGT 3360
3361 GGTGAATGTG AAACCAGTAA CGTTATACGA  TGTCGCAGAG TATGCCGGTG TCTCTTATCA 3420
3421 GACCGTTTCC CGCGTGGTGA ACCAGGCCAG  CCACGTTTCT GCGAAAACGC GGGAAAAAGT 3480
3481 GGAAGCGGCG ATGGCGGAGC TGAATTACAT  TCCCAACCGC GTGGCACAAC AACTGGCGGG 3540
3541 CAAACAGTCG TTGCTGATTG GCGTTGCCAC  CTCCAGTCTG GCCCTGCACG CGCCGTCGCA 3600
3601 AATTGTCGCG GCGATTAAAT CTCGCGCCGA  TCAACTGGGT GCCAGCGTGG TGGTGTCGAT 3660
3661 GGTAGAACGA AGCGGCGTCG AAGCCTGTAA  AGCGGCGGTG CACAATCTTC TCGCGCAACG 3720
3721 CGTCAGTGGG CTGATCATTA ACTATCCGCT  GGATGACCAG GATGCCATTG CTGTGGAAGC 3780
3781 TGCCTGCACT AATGTTCCGG CGTTATTTCT  TGATGTCTCT GACCAGACAC CCATCAACAG 3840
3841 TATTATTTTC TCCCATGAAG ACGGTACGCG  ACTGGGCGTG GAGCATCTGG TCGCATTGGG 3900
3901 TCACCAGCAA ATCGCGCTGT TAGCGGGCCC  ATTAAGTTCT GTCTCGGCGC GTCTGCGTCT 3960
3961 GGCTGGCTGG CATAAATATC TCACTCGCAA  TCAAATTCAG CCGATAGCGG AACGGGAAGG 4020
4021 CGACTGGAGT GCCATGTCCG GTTTTCAACA  AACCATGCAA ATGCTGAATG AGGGCATCGT 4080
4081 TCCCACTGCG ATGCTGGTTG CCAACGATCA  GATGGCGCTG GGCGCAATGC GCGCCATTAC 4140
4141 CGAGTCCGGG CTGCGCGTTG GTGCGGATAT  CTCGGTAGTG GGATACGACG ATACCGAAGA 4200
4201 CAGCTCATGT TATATCCCGC CGTTAACCAC  CATCAAACAG GATTTTCGCC TGCTGGGGCA 4260
4261 AACCAGCGTG GACCGCTTGC TGCAACTCTC  TCAGGGCCAG GCGGTGAAGG GCAATCAGCT 4320
4321 GTTGCCCGTC TCACTGGTGA AAAGAAAAAC  CACCCTGGCG CCCAATACGC AAACCGCCTC 4380
4381 TCCCCGCGCG TTGGCCGATT CATTAATGCA  GCTGGCACGA CAGGTTTCCC GACTGGAAAG 4440
4441 CGGGCAGTGA GCGCAACGCA ATTAATGTGA  GTTAGCTCAC TCATTAGGCA CCCCAGGCTT 4500
4501 TACACTTTAT GCTTCCGGCT CGTATGTTGT  GTGGAATTGT GAGCGGATAA CAATTTCACA 4560
4561 CAGGAAACAG CTATGACCAT GATTACGGAT  TCACTGGCCG TCGTTTTACA ACGTCGTGAC 4620
4621 TGGGAAAACC CTGGCGTTAC CCAACTTAAT  CGCCTTGCAG CACATCCCCC TTTCGCCAGC 4680
4681 TGGCGTAATA GCGAAGAGGC CCGCACCGAT  CGCCCTTCCC AACAGTTGCG CAGCCTGAAT 4740
4741 GGCGAATGGC GCTTTGCCTG GTTTCCGGCA  CCAGAAGCGG TGCCGGAAAG CTGGCTGGAG 4800
4801 TGCGATCTTC CTGAGGCCGA TACTGTCGTC  GTCCCCTCAA ACTGGCAGAT GCACGGTTAC 4860
4861 GATGCGCCCA TCTACACCAA CGTAACCTAT  CCCATTACGG TCAATCCGCC GTTTGTTCCC 4920
4921 ACGGAGAATC CGACGGGTTG TTACTCGCTC  ACATTTAATG TTGATGAAAG CTGGCTACAG 4980
4981 GAAGGCCAGA CGCGAATTAT TTTTGATGGC  GTTGGAATT                        5019
        |  10      |  20      |  30      |  40      |  50      |  60
```

pGEX-cs Vector Genetic Map

DNA sequence 5019 b.p. ACGTTATCGACT ... GGCGTTGGAATT circular



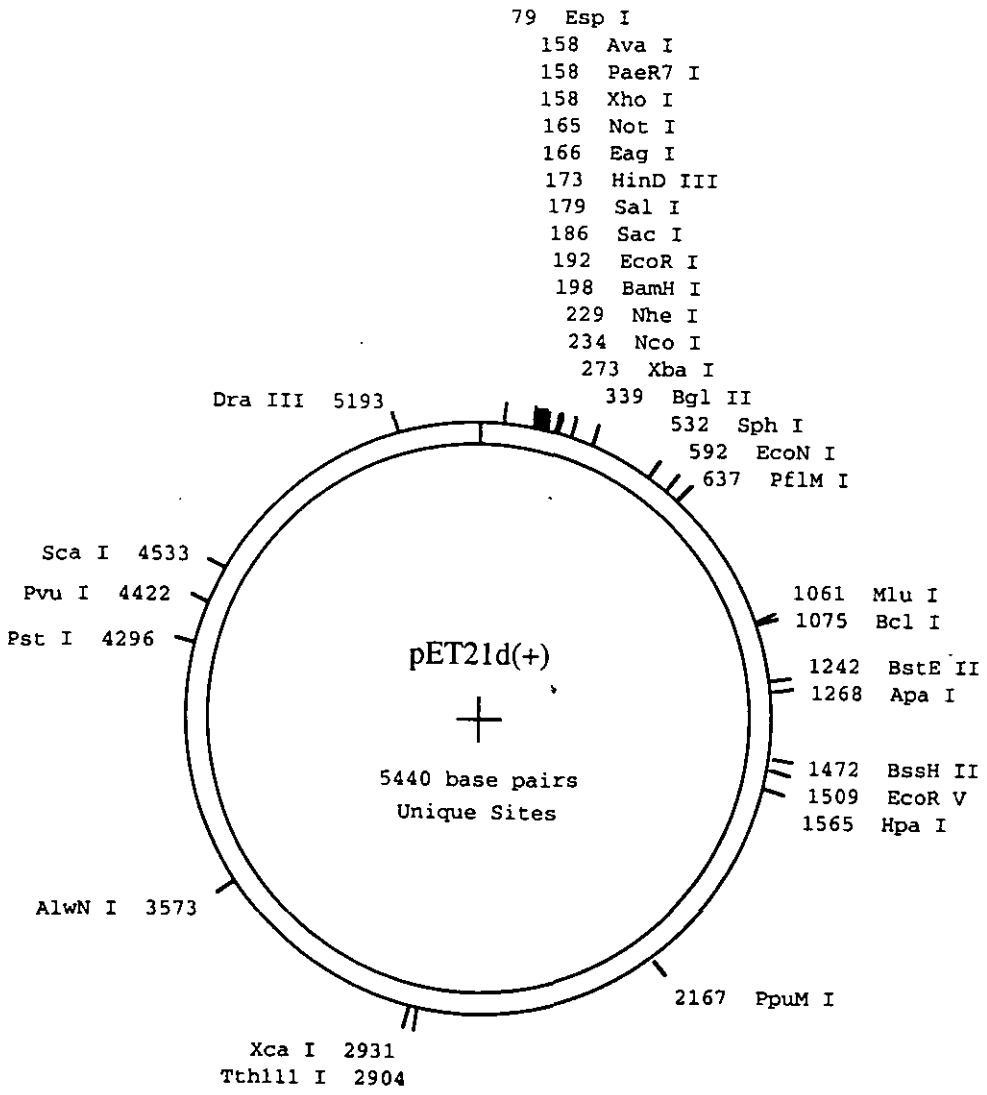
pET21d Vector Linear Sequence

DNA sequence 5446 b.p. ATCCGGATATAG . TCCCATTCGCCA linear

```
        |  10      |  20      |  30      |  40      |  50      |  60
   1 ATCCGGATAT AGTTCCTCCT TTCAGCAAAA AACCCCTCAA GACCCGTTTA GAGGCCCCAA 60
  61 GGGGTTATGC TAGTTATTGC TCAGCGGTGG CAGCAGCCAA CTCAGCTTCC TTTCGGGCTT 120
 121 TGTTAGCAGC CGGATCTCAG TGGTGGTGGT GGTGGTGCTC GAGTGCGGCC GCAAGCTTGT 180
 181 CGACGGAGCT CGAATTCGGA TCCGACCCAT TTGCTGTCCA CCAGTCATGC TAGCCATGGT 240
 241 ATATCTCCTT CTTAAAGTTA AACAAAATTA TTTCTAGAGG GGAATTGTTA TCCGCTCACA 300
 301 ATTCCCCTAT AGTGAGTCGT ATTAATTTCG CGGGATCGAG ATCTCGATCC TCTACGCCGG 360
 361 ACGCATCGTG GCCGGCATCA CCGGCGCCAC AGGTGCGGTT GCTGGCGCCT ATATCGCCGA 420
 421 CATCACCGAT GGGGAAGATC GGGCTCGCCA CTTCGGGCTC ATGAGCGCTT GTTTCGGCGT 480
 481 GGGTATGGTG GCAGGCCCCG TGGCCGGGGG ACTGTTGGGC GCCATCTCCT TGCATGCACC 540
 541 ATTCCTTGCG GCGGCGGTGC TCAACGGCCT CAACCTACTA CTGGGCTGCT TCCTAATGCA 600
 601 GGAGTCGCAT AAGGGAGAGC GTCGAGATCC CGGACACCAT CGAATGGCGC AAAACCTTTC 660
 661 GCGGTATGGC ATGATAGCGC CCGGAAGAGA GTCAATTCAG GGTGGTGAAT GTGAAACCAG 720
 721 TAACGTTATA CGATGTCGCA GAGTATGCCG GTGTCTCTTA TCAGACCGTT TCCCGCGTGG 780
 781 TGAACCAGGC CAGCCACGTT TCTGCGAAAA CGCGGGAAAA AGTGGAAGCG GCGATGGCGG 840
 841 AGCTGAATTA CATTCCCAAC CGCGTGGCAC AACAACTGGC GGGCAAACAG TCGTTGCTGA 900
 901 TTGGCGTTGC CACCTCCAGT CTGGCCCTGC ACGCGCCGTC GCAAATTGTC GCGGCGATTA 960
 961 AATCTCGCGC CGATCAACTG GGTGCCAGCG TGGTGGTGTC GATGGTAGAA CGAAGCGGCG 1020
1021 TCGAAGCCTG TAAAGCGGCG GTGCACAATC TTCTCGCGCA ACGCGTCAGT GGGCTGATCA 1080
1081 TTAACTATCC GCTGGATGAC CAGGATGCCA TTGCTGTGGA AGCTGCCTGC ACTAATGTTC 1140
1141 CGGCGTTATT TCTTGATGTC TCTGACCAGA CACCCATCAA CAGTATTATT TTCTCCCATG 1200
1201 AAGACGGTAC GCGACTGGGC GTGGAGCATC TGGTCGCATT GGGTCACCAG CAAATCGCGC 1260
1261 TGTTAGCGGG CCCATTAAGT TCTGTCTCGG CGCGTCTGCG TCTGGCTGGC TGGCATAAAT 1320
1321 ATCTCACTCG CAATCAAATT CAGCCGATAG CGGAACGGGA AGGCGACTGG AGTGCCATGT 1380
1381 CCGGTTTTCA ACAAACCATG CAAATGCTGA ATGAGGGCAT CGTTCCCACT GCGATGCTGG 1440
1441 TTGCCAACGA TCAGATGGCG CTGGGCGCAA TGCGCGCCAT TACCGAGTCC GGGCTGCGCG 1500
1501 TTGGTGCGGA TATCTCGGTA GTGGGATACG ACGATACCGA AGACAGCTCA TGTTATATCC 1560
1561 CGCCGTTAAC CACCATCAAA CAGGATTTTC GCCTGCTGGG GCAAACCAGC GTGGACCGCT 1620
1621 TGCTGCAACT CTCTCAGGGC CAGGCGGTGA AGGGCAATCA GCTGTTGCCC GTCTCACTGG 1680
1681 TGAAAAGAAA AACCACCCTG GCGCCCAATA CGCAAACCGC CTCTCCCCGC GCGTTGGCCG 1740
1741 ATTCATTAAT GCAGCTGGCA CGACAGGTTT CCCGACTGGA AAGCGGGCAG TGAGCGCAAC 1800
1801 GCAATTAATG TAAGTTAGCT CACTCATTAG GCACCGGGAT CTCGACCGAT GCCCTTGAGA 1860
1861 GCCTTCAACC CAGTCAGCTC CTTCCGGTGG GCGCGGGGCA TGACTATCGT CGCCGCACTT 1920
1921 ATGACTGTCT TCTTTATCAT GCAACTCGTA GGACAGGTGC CGGCAGCGCT CTGGGTCATT 1980
1981 TTCGGCGAGG ACCGCTTTCG CTGGAGCGCG ACGATGATCG GCCTGTCGCT TGCGGTATTC 2040
2041 GGAATCTTGC ACGCCCTCGC TCAAGCCTTC GTCACTGGTC CCGCCACCAA ACGTTTCGGC 2100
2101 GAGAAGCAGG CCATTATCGC CGGCATGGCG GCCCCACGGG TGCGCATGAT CGTGCTCCTG 2160
2161 TCGTTGAGGA CCCGGCTAGG CTGGCGGGGT TGCCTTACTG GTTAGCAGAA TGAATCACCG 2220
2221 ATACGCGAGC GAACGTGAAG CGACTGCTGC TGCAAAACGT CTGCGACCTG AGCAACAACA 2280
2281 TGAATGGTCT TCGGTTTCCG TGTTTCGTAA AGTCTGGAAA CGCGGAAGTC AGCGCCCTGC 2340
2341 ACCATTATGT TCCGGATCTG CATCGCAGGA TGCTGCTGGC TACCCTGTGG AACACCTACA 2400
2401 TCTGTATTAA CGAAGCGCTG GCATTGACCC TGAGTGATTT TTCTCTGGTC CCGCCGCATC 2460
2461 CATACCGCCA GTTGTTTACC CTCACAACGT TCCAGTAACC GGGCATGTTC ATCATCAGTA 2520
2521 ACCCGTATCG TGAGCATCCT CTCTCGTTTC ATCGGTATCA TTACCCCCAT GAACAGAAAT 2580
2581 CCCCCTTACA CGGAGGCATC AGTGACCAAA CAGGAAAAAA CCGCCCTTAA CATGGCCCGC 2640
2641 TTTATCAGAA GCCAGACATT AACGCTTCTG GAGAAACTCA ACGAGCTGGA CGCGGATGAA 2700
2701 CAGGCAGACA TCTGTGAATC GCTTCACGAC CACGCTGATG AGCTTTACCG CAGCTGCCTC 2760
2761 GCGCGTTTCG GTGATGACGG TGAAAACCTC TGACACATGC AGCTCCCGGA GACGGTCACA 2820
2821 GCTTGTCTGT AAGCGGATGC CGGGAGCAGA CAAGCCCGTC AGGGCGCGTC AGCGGGTGTT 2880
2881 GGCGGGTGTC GGGGCGCAGC CATGACCCAG TCACGTAGCG ATAGCGGAGT GTATACTGGC 2940
2941 TTAACTATGC GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATATGC GGTGTGAAAT 3000
3001 ACCGCACAGA TGCGTAAGGA GAAAATACCG CATCAGGCGC TCTTCCGCTT CCTCGCTCAC 3060
3061 TGACTCGCTG CGCTCGGTCG TTCGGCTGCG GCGAGCGGTA TCAGCTCACT CAAAGGCGGT 3120
3121 AATACGGTTA TCCACAGAAT CAGGGGATAA CGCAGGAAAG AACATGTGAG CAAAAGGCCA 3180
3181 GCAAAAGGCC AGGAACCGTA AAAAGGCCGC GTTGCTGGCG TTTTTCCATA GGCTCCGCCC 3240
3241 CCCTGACGAG CATCACAAAA ATCGACGCTC AAGTCAGAGG TGGCGAAACC CGACAGGACT 3300
3301 ATAAAGATAC CAGGCGTTTC CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG TTCCGACCCT 3360
3361 GCCGCTTACC GGATACCTGT CCGCCTTTCT CCCTTCGGGA AGCGTGGCGC TTTCTCATAG 3420
3421 CTCACGCTGT AGGTATCTCA GTTCGGTGTA GGTCGTTCGC TCCAAGCTGG GCTGTGTGCA 3480
3481 CGAACCCCCC GTTCAGCCCG ACCGCTGCGC CTTATCCGGT AACTATCGTC TTGAGTCCAA 3540
3541 CCCGGTAAGA CACGACTTAT CGCCACTGGC AGCAGCCACT GGTAACAGGA TTAGCAGAGC 3600
3601 GAGGTATGTA GGCGGTGCTA CAGAGTTCTT GAAGTGGTGG CCTAACTACG GCTACACTAG 3660
3661 AAGGACAGTA TTTGGTATCT GCGCTCTGCT GAAGCCAGTT ACCTTCGGAA AAAGAGTTGG 3720
3721 TAGCTCTTGA TCCGGCAAAC AAACCACCGC TGGTAGCGGT GGTTTTTTTG TTTGCAAGCA 3780
3781 GCAGATTACG CGCAGAAAAA AAGGATCTCA AGAAGATCCT TTGATCTTTT CTACGGGGTC 3840
3841 TGACGCTCAG TGGAACGAAA ACTCACGTTA AGGGATTTTG GTCATGAGAT TATCAAAAAG 3900
3901 GATCTTCACC TAGATCCTTT TAAATTAAAA ATGAAGTTTT AAATCAATCT AAAGTATATA 3960
3961 TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA TCTCAGCGAT 4020
4021 CTGTCTATTT CGTTCATCCA TAGTTGCCTG ACTCCCCGTC GTGTAGATAA CTACGATACG 4080
4081 GGAGGGCTTA CCATCTGGCC CCAGTGCTGC AATGATACCG CGAGACCCAC GCTCACCGGC 4140
4141 TCCAGATTTA TCAGCAATAA ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA GTGGTCCTGC 4200
4201 AACTTTATCC GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG TAAGTAGTTC 4260
4261 GCCAGTTAAT AGTTTGCGCA ACGTTGTTGC CATTGCTGCA GGCATCGTGG TGTCACGCTC 4320
4321 GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG TTCCCAACGA TCAAGGCGAG TTACATGATC 4380
4381 CCCCATGTTG TGCAAAAAAG CGGTTAGCTC CTTCGGTCCT CCGATCGTTG TCAGAAGTAA 4440
4441 GTTGGCCGCA GTGTTATCAC TCATGGTTAT GGCAGCACTG CATAATTCTC TTACTGTCAT 4500
4501 GCCATCCGTA AGATGCTTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT TCTGAGAATA 4560
4561 GTGTATGCGG CGACCGAGTT GCTCTTGCCC GGCGTCAATA CGGGATAATA CCGCGCCACA 4620
4621 TAGCAGAACT TTAAAAGTGC TCATCATTGG AAAACGTTCT TCGGGGCGAA AACTCTCAAG 4680
4681 GATCTTACCG CTGTTGAGAT CCAGTTCGAT GTAACCCACT CGTGCACCCA ACTGATCTTC 4740
4741 AGCATCTTTT ACTTTCACCA GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC AAAATGCCGC 4800
4801 AAAAAAGGGA ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC TTTTTCAATA 4860
4861 TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG AATGTATTTA 4920
4921 GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA AAAGTGCCAC CTGAAATTGT 4980
4981 AAACGTTAAT ATTTTGTTAA AATTCGCGTT AAATTTTTGT TAAATCAGCT CATTTTTTAA 5040
5041 CCAATAGGCC GAAATCGGCA AAATCCCTTA TAAATCAAAA GAATAGACCG AGATAGGGTT 5100
5101 GAGTGTTGTT CCAGTTTGGA ACAAGAGTCC ACTATTAAAG AACGTGGACT CCAACGTCAA 5160
5161 AGGGCGAAAA ACCGTCTATC AGGGCGATGG CCCACTACGT GAACCATCAC CCTAATCAAG 5220
5221 TTTTTTGGGG TCGAGGTGCC GTAAAGCACT AAATCGGAAC CCTAAAGGGA GCCCCCGATT 5280
5281 TAGAGCTTGA CGGGGAAAGC CGGCGAACGT GGCGAGAAAG GAAGGGAAGA AAGCGAAAGG 5340
5341 AGCGGGCGCT AGGGCGCTGG CAAGTGTAGC GGTCACGCTG CGCGTAACCA CCACACCCGC 5400
5401 CGCGCTTAAT GCGCCGCTAC AGGGCGCGTC CCATTCGCCA                        5446
        |  10      |  20      |  30      |  40      |  50      |  60
```

pET21d Vector Genetic Map

DNA sequence 5440 b.p. ATCCGGATATAG ... TCCCATTCGCCA circular



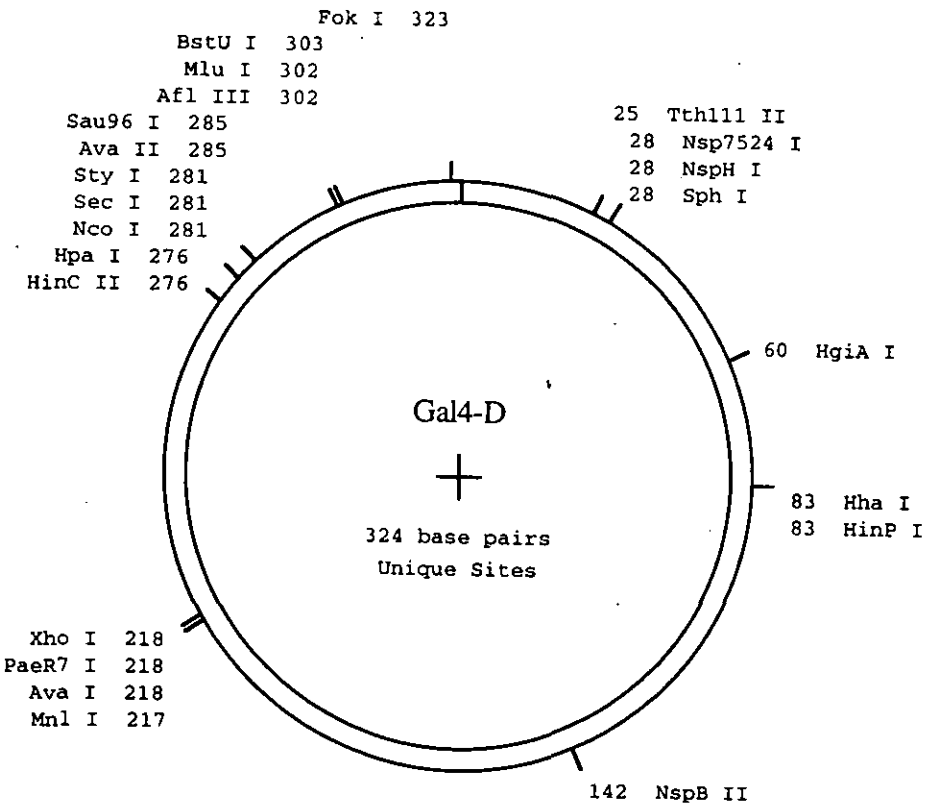
GAL4-D 1-phase Translation

```
DNA  sequence    324 b.p.   atgaagctactg ... atcactacaggg   linear

1    /    1                                31   /   11
atg aag cta ctg tct tct atc gaa caa gca tgc gat att tgc cga ctt aaa aag ctc aag
met lys leu leu ser ser ile glu gln ala cys asp ile cys arg leu lys lys leu lys
61   /   21                               91   /   31
tgc tcc aaa gaa aaa ccg aag tgc gcc aag tgt ctg aag aac aac tgg gag tgt cgc tac
cys ser lys glu lys pro lys cys ala lys cys leu lys asn asn trp glu cys arg tyr
121  /   41                               151  /   51
tct ccc aaa acc aaa agg tct ccg ctg act agg gca cat ctg aca gaa gtg gaa tca agg
ser pro lys thr lys arg ser pro leu thr arg ala his leu thr glu val glu ser arg
181  /   61                               211  /   71
cta gaa aga ctg gaa cag cta ttt cta ctg att ttt cct cga gaa gac ctt gac atg att
leu glu arg leu glu gln leu phe leu leu ile phe pro arg glu asp leu asp met ile
241  /   81                               271  /   91
ttg aaa atg gat tct tta cag gat ata aaa gca ttg tta acc atg gac caa act gcg tat
leu lys met asp ser leu gln asp ile lys ala leu leu thr met asp gln thr ala tyr
301  /  101
aac gcg ttt gga atc act aca ggg
asn ala phe gly ile thr thr gly
```

GAL4-D Genetic Map

DNA sequence 324 b.p. atgaagctactg ... atcactacaggg circular



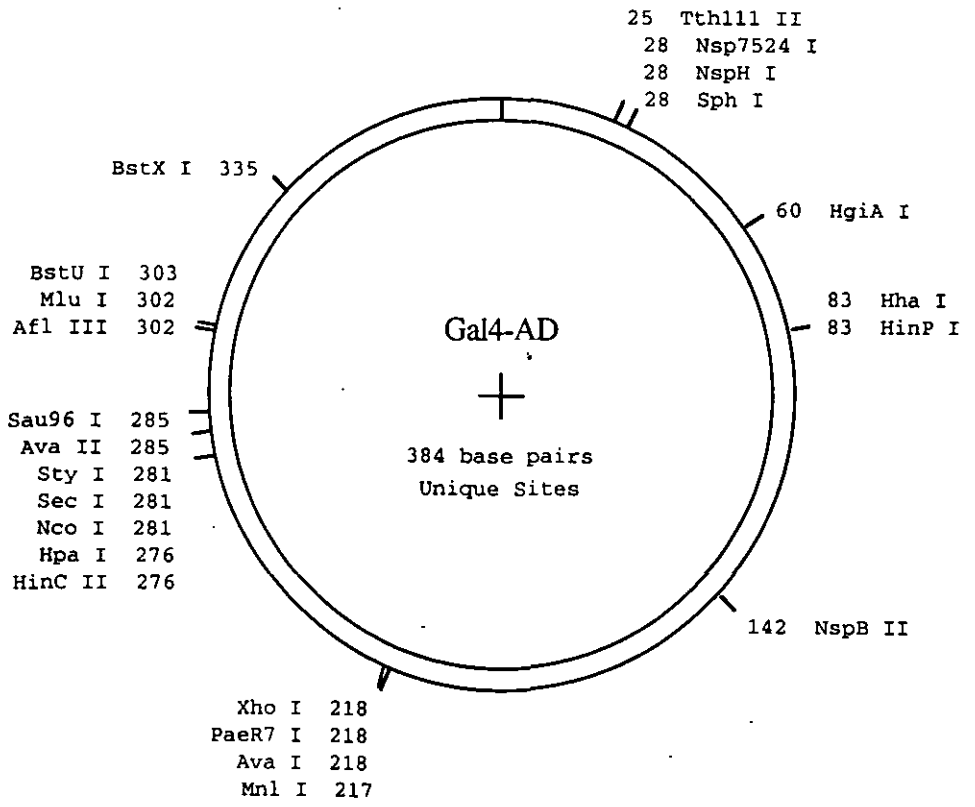
GAL4-AD 1-phase Translation

```
DNA  sequence    384 b.p.   atgaagctactg ... gatgaagatacc   linear

1    /    1                              31   /   11
atg aag cta ctg tct tct atc gaa caa gca tgc gat att tgc cga ctt aaa aag ctc aag
met lys leu leu ser ser ile glu gln ala cys asp ile cys arg leu lys lys leu lys
61   /   21                             91   /   31
tgc tcc aaa gaa aaa ccg aag tgc gcc aag tgt ctg aag aac aac tgg gag tgt cgc tac
cys ser lys glu lys pro lys cys ala lys cys leu lys asn asn trp glu cys arg tyr
121  /   41                             151  /   51
tct ccc aaa acc aaa agg tct ccg ctg act agg gca cat ctg aca gaa gtg gaa tca agg
ser pro lys thr lys arg ser pro leu thr arg ala his leu thr glu val glu ser arg
181  /   61                             211  /   71
cta gaa aga ctg gaa cag cta ttt cta ctg att ttt cct cga gaa gac ctt gac atg att
leu glu arg leu glu gln leu phe leu leu ile phe pro arg glu asp leu asp met ile
241  /   81                             271  /   91
ttg aaa atg gat tct tta cag gat ata aaa gca ttg tta acc atg gac caa act gcg tat
leu lys met asp ser leu gln asp ile lys ala leu leu thr met asp gln thr ala tyr
301  /  101                            331  /  111
aac gcg ttt gga atc act aca ggg atg ttt aat acc act aca atg gat gat gta tat aac
asn ala phe gly ile thr thr gly met phe asn thr thr thr met asp asp val tyr asn
361  /  121
tat cta ttc gat gat gaa gat acc
tyr leu phe asp asp glu asp thr
```

GAL4-AD Genetic Map

DNA sequence 384 b.p. atgaagctactg ... gatgaagatacc circular



GAL4-VP16 Experimental Linear Sequence

5' end – accatctgctaaactcgaaggttcggcccaattattcccataggaaggcgccgtaggggcgagtcgtgggggcgtaatcc ggacccggggaatccccgtcccccaacatgtcagatcgaatcgtctagcgctcgcatgccatcgtctgcgtctagctggcttccagtga ctctgcatgggcattcctactcactttcttact - 3' end.

GAL4-VP16 Linear Sequence[11]

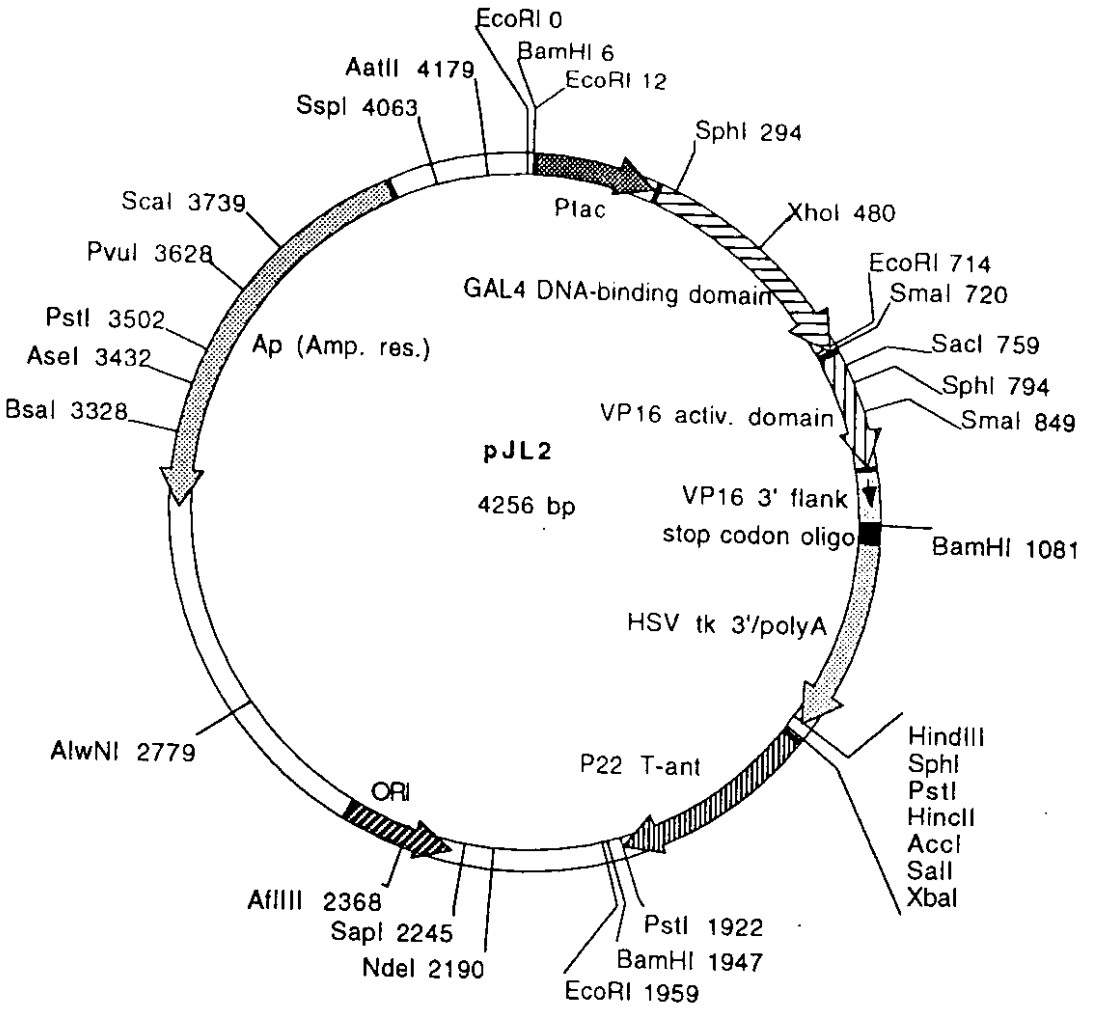
ATGAAGCTACTGTCTTCTATCGAACAAGCATGCGATATTTGCCGACTTAA
AAAGCTCAAGTGCTCCAAAGAAAAACCGAAGTGCGCCAAGTGTCTGAAGA
ACAACTGGGAGTGTCGCTACTCTCCCAAAACCAAAAGGTCTCCGCTGACT
AGGGCACATCTGACAGAAGTGGAATCAAGGCTAGAAAGACTGGAACAGCT
ATTTCTACTGATTTTTCCTCGAGAAGACCTTGACATGATTTTGAAAATGG
ATTCTTTACAGGATATAAAAGCATTGTTAACAGGATTATTTGTACAAGAT
AATGTGAATAAAGATGCCGTCACAGATAGATTGGCTTCAGTGGAGACTGA
TATGCCTCTAACATTGAGACAGCATAGAATAAGTGCGACATCATCATCGG
AAGAGAGTAGTAACAAAGGTCAAAGACAGTTGACTGTATCG

CCGGAATTCCCGGGGATCTGGG

GCCCCCCCGACCGATGTCAGCCTGGGGGACGAGCTCCACTTAGAC
GGCGAGGACGTGGCGATGGCGCATGCCGACGCGCTAGACGATTTC
GATCTGGACATGTTGGGGGACGGGGATTCCCCGGGTCCGGGATTTACC
CCCCACGACTCCGCCCCCTACGGCGCTCTGGATATGGCCGACTTCGAG
TTTGAGCAGATGTTTACCGATGCCCTTGGAATTGACGAGTACGGTGGGTAG

[11] Contributed by the Department of Biochemistry, Michigan State University.

pJL2 Bacterial Expression Vector



Plasmid name: pJL2
Plasmid size: 4256 bp
Constructed by: Janet Leatherwood
Construction date: 1987?
Comments/References: Expression of GAL4-VP16 fusion protein from ptac promoter. Map constructed in 1993 by Steve Triezenberg. Some map positions, especially in P22-ant and ptac regions, are only approximate.

Appendix B

Materials:

1. pET21d plasmid (1997 Novagen Catalog, p. 122)
2. pGEX-cs plasmid (Novagen catalog)
3. GAL4-AD DNA, GAL4-D DNA
4. GAL4-VP16 DNA (Michigan State University)
5. PCR materials (Pfu, dNTP, PCR buffer, etc)
6. Shrimp alkaline phophatase
7. T4 DNA ligase
8. Media components
9. DNA extraction kit (BIGGERprep Plasmid Preparation Kit)
10. Agarose (type)
11. Acrylamide gels
12. NcoI restriction enzyme (New England Biolabs Catalog, p. 46)
13. XhoI restriction enzyme (New England Biolabs Catalog, p. 60)
14. Resin for the micro column purification method

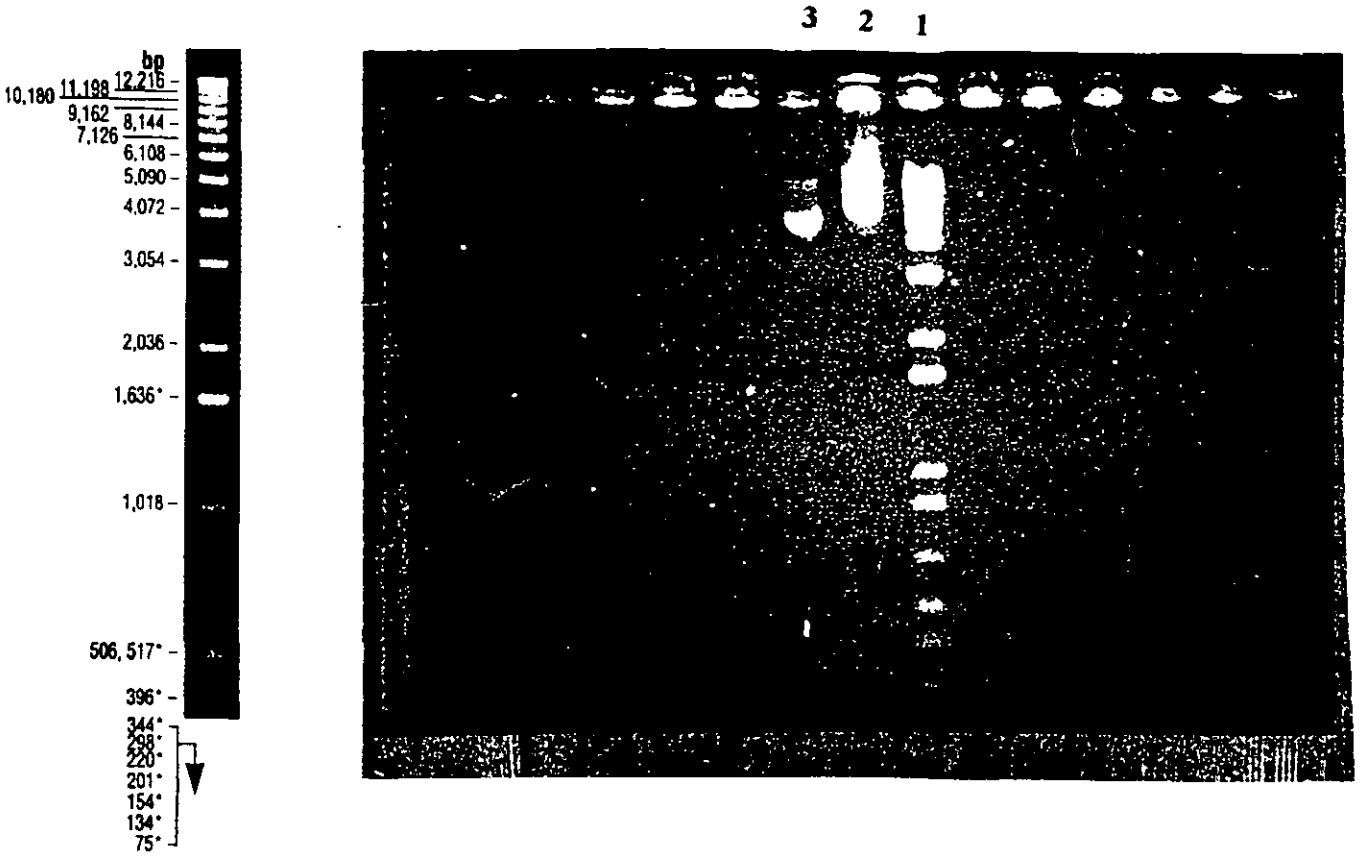
Methods:

1. Cell transformation by electroporation
2. DNA isolation kit. (Kiagen Biggerprep)
3. PCR amplification, insert construction.
4. Plasmid dephosphorilation
5. Plasmid-PCR constructs ligation

6. Competent cell production
7. pET21d plasmid digestion
8. pET21d plasmid purification by spincolumn procedure
9. Quantification of isolated DNA concentration by UV scan in spectrophotometer
10. GAL4-VP16 TF sequencing, method 1
11. GAL4-VP16 TF sequencing, method 2
12. Preparation of the media (recipes)
13. Agarose gels check for experimental results
14. Sequencing acrylamide gels

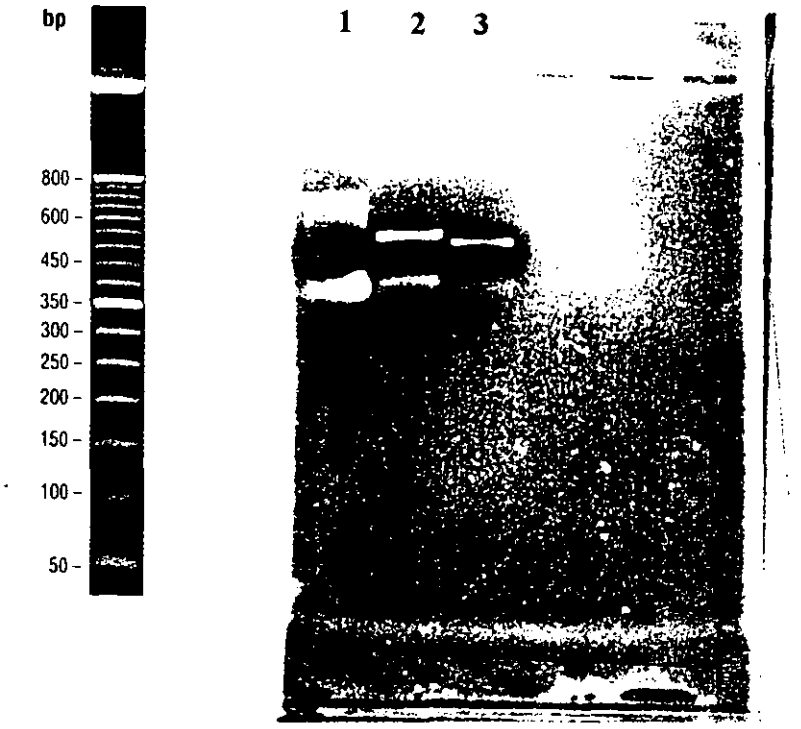
Appendix C

Figure 1: pET21d Digestion by XhoI and NcoI



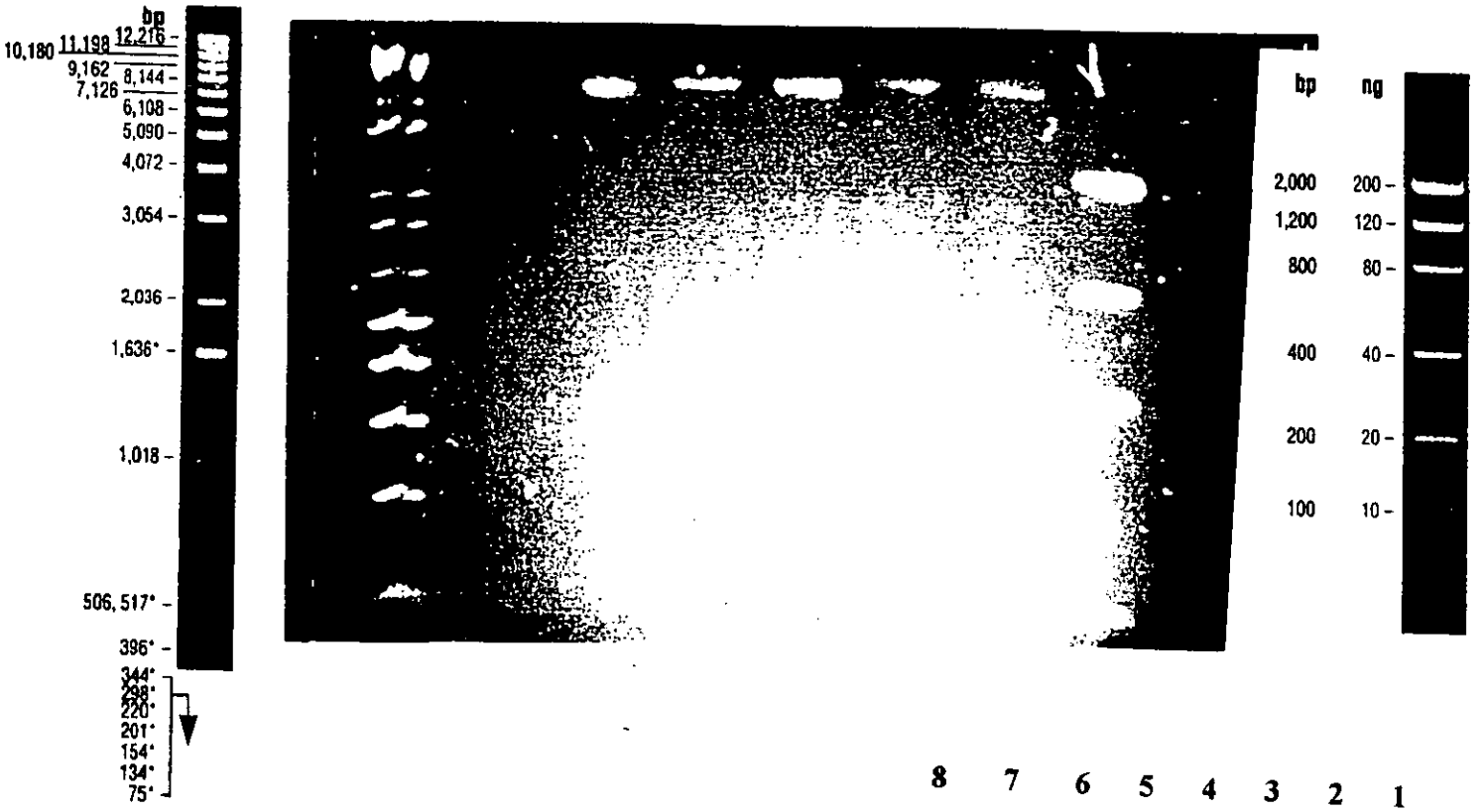
1. 1 Kb DNA Ladder (loaded 0.5µl).
2. Undigested pET21d Vector (50ng/µl, loaded 3µl).
3. Digested pET21d Vector (50ng/µl, loaded 3µl).

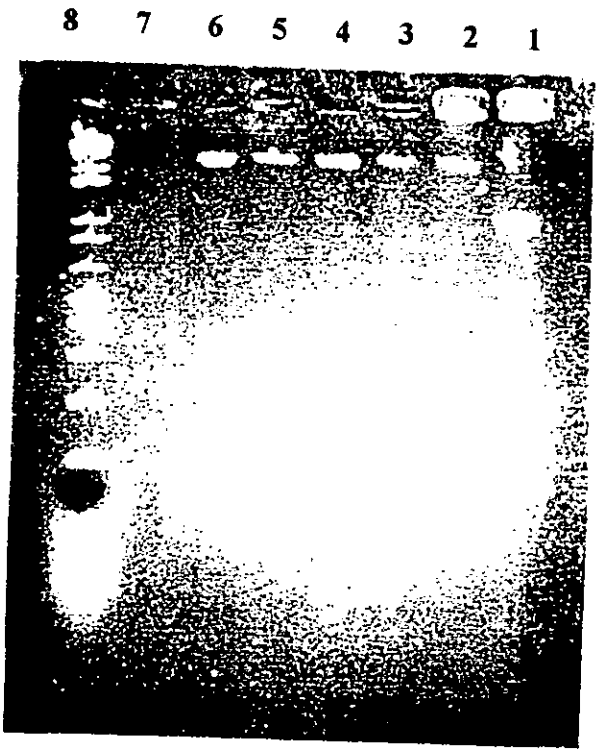
Figure 2: GAL4-AD and GAL4-D Digested PCR Inserts



1. 50 bp DNA Ladder (loaded 0.5µl).
2. GAL4-AD Insert (loaded 20µl of the PCR reaction).
3. GAL4-D Insert (loaded 20µl of the PCR reaction).

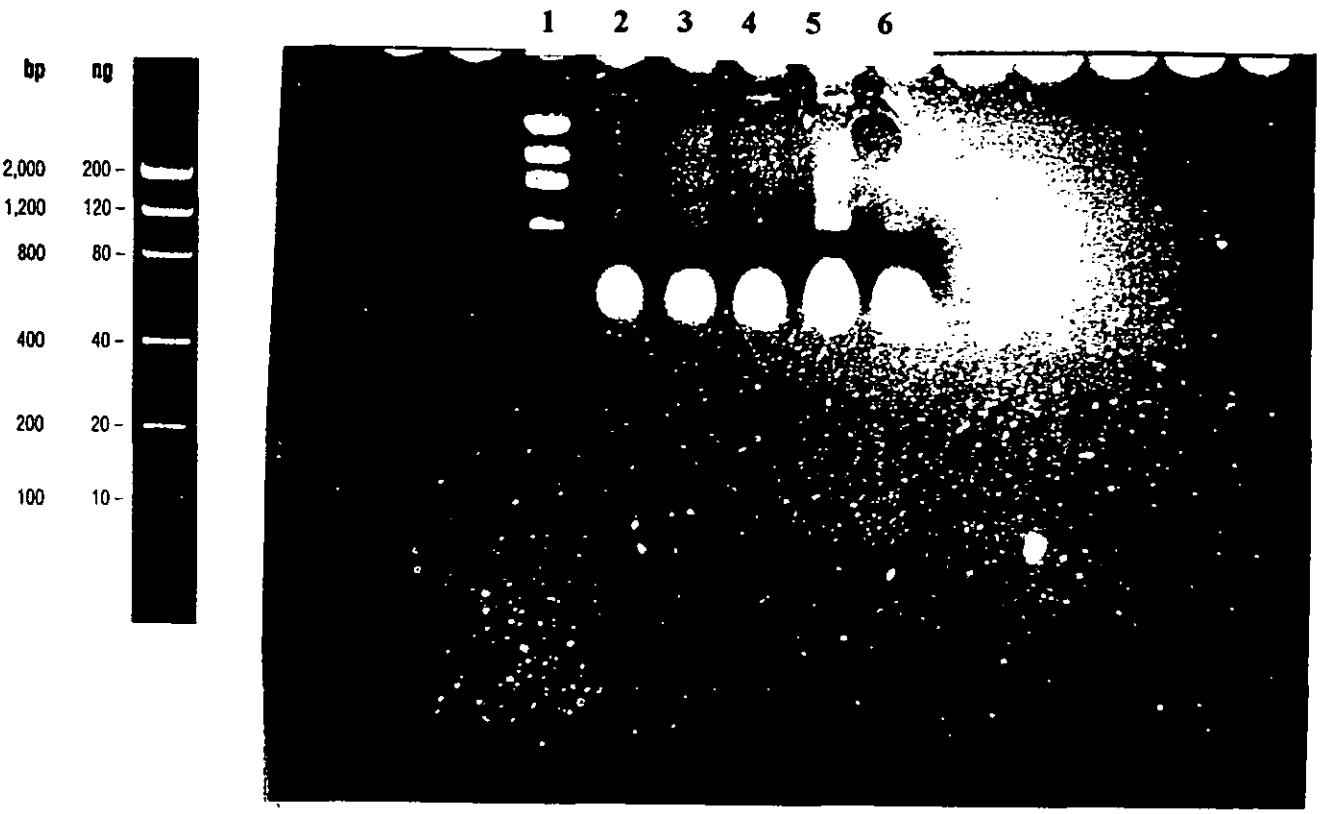
Figure 3: GAL4-AD and GAL4-D Ligation Products





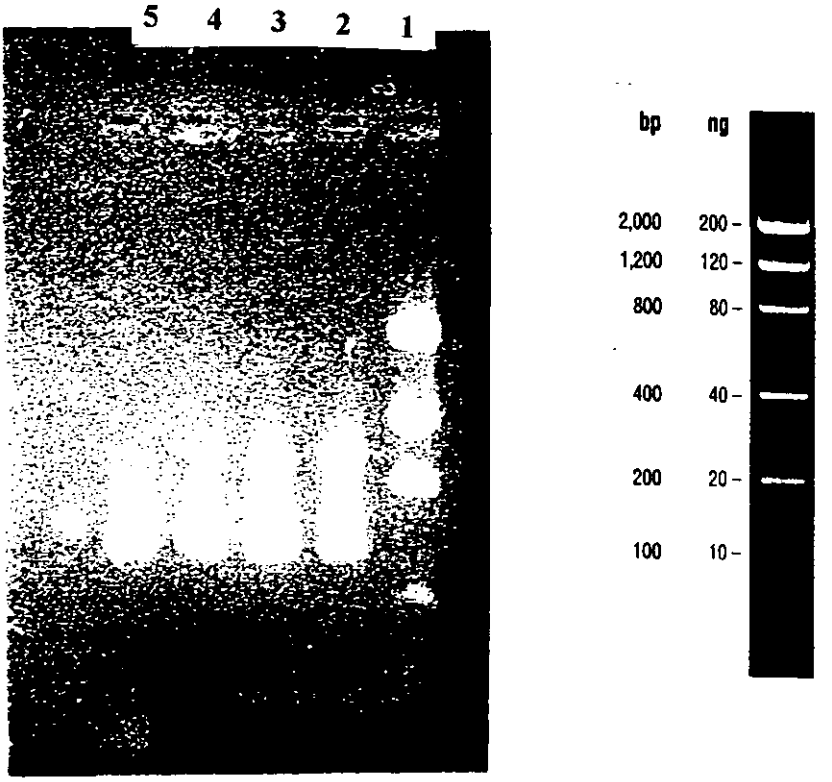
1. Low Mass DNA Ladder (loaded 2µl).
2. GAL4-AD Ligation Reaction (loaded 2µl).
3. GAL4-AD Ligation Reaction (loaded 2µl).
4. GAL4-D Ligation Reaction (loaded 2µl).
5. GAL4-D Ligation Reaction (loaded 2µl).
6. pET21d Ligation Reaction (loaded 2µl).
7. GAL4-AD Insert Ligation Reaction (loaded 2µl).
8. 1 Kb DNA Ladder (loaded 0.5µl).

Figure 4: Taq Polymerase Protocol Amplifications of GAL4-VP16



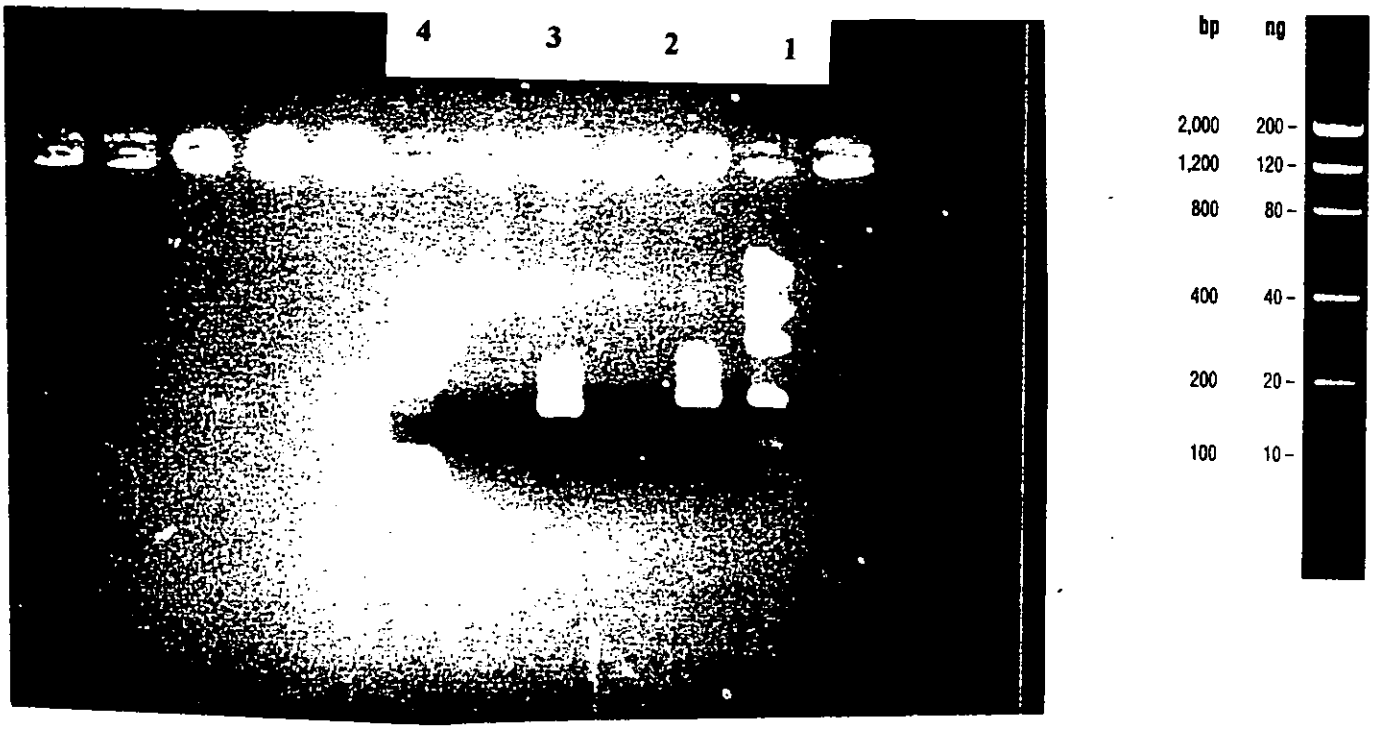
1. Low Mass DNA Ladder (loaded 2µl).
2. Taq Polymerase Amplified GAL4-VP16 Inserts (loaded 6µl from the PCR reaction).
3. Taq Polymerase Amplified GAL4-VP16 Inserts (loaded 6µl from the PCR reaction).
4. Taq Polymerase Amplified GAL4-VP16 Inserts (loaded 6µl from the PCR reaction).Taq Polymerase Amplified GAL4-VP16 Inserts (loaded 6µl from the PCR reaction).
5. Taq Polymerase Amplified GAL4-VP16 Inserts (loaded 6µl from the PCR reaction).
6. Taq Polymerase Amplified GAL4-VP16 Inserts (loaded 6µl from the PCR reaction).

Figure 5: Pfu Turbo Protocol Amplifications of GAL4-VP16



1. Low Mass DNA Ladder (loaded 2µl).
2. Pfu Turbo Amplified GAL4-VP16 Inserts (loaded 6µl from the PCR reaction).
3. Pfu Turbo Amplified GAL4-VP16 Inserts (loaded 6µl from the PCR reaction).
4. Pfu Turbo Amplified GAL4-VP16 Inserts (loaded 6µl from the PCR reaction).
5. Pfu Turbo Amplified GAL4-VP16 Inserts (loaded 6µl from the PCR reaction).

Figure 6: GAL4 Series Amplification Products



1. Low Mass DNA Ladder (loaded 2µl).
2. GAL4-AD Insert (loaded 6µl from digested insert).
3. GAL4-D Insert (loaded 6µl from digested insert).
4. GAL4-VP16 Insert (loaded 6µl from the PCR reaction).

Acknowledgements

I would like to thank my advisor on this project, Dr. Blaine Bartholomew, for this research opportunity and for his support and guidance. Also, I would like to thank Jim Persinger for all of his practical research advice.

I would like to express my gratitude to Søren Jaglo-Ottosen of Michigan State University's Department of Biochemistry, for providing with a vector containing GAL4-VP16.

Finally, I would like to thank my wife for her unconditional support and her assistance in the editing process of this thesis.

References

1. Beato, M. & K. Eisfeld. 1997. Transcription factor access to chromatin. *Nucleic Acids Res*, 25: 3559-3563.
2. Gadhavi P.L. 1997. An electrospray ionisation mass spectrometry (ESI-MS) study to probe the metal ion binding site in the DNA binding domain of the yeast transcriptional activator GAL4. *FEBS Lett*, 417(1): 145-149.
3. Lewin, B. 1997. Genes VI. Oxford Univ, New York. 1260 pp.
4. Liang S.D., R. Marmorstein, S.C. Harrison & M. Ptashine. 1996. DNA sequence preferences of GAL4 and PPR1: how a subset of Zn2 Cys6 binuclear cluster proteins recognize DNA. *Mol Cell Biol*, 16(7): 3773-3780.
5. Melcher K. & S.A. Johnston. 1995. GAL4 interacts with TATA-binding protein and coactivators. *Mol Cell Biol*, 15(5): 2839-2848.
6. Poch O. 1997. Conservation of a putative inhibitory domain in the GAL4 family members. *Gene*, 184(2), 229-235.
7. Xu H.E., T. Kodadek & S.A. Johnston. 1995. A single GAL4 dimer can maximally activate transcription under physiological conditions. *Proc Natl Acad Sci U S A*, 92(17): 7677-7680.
8. Xu M., R.T. Simpson & M.P. Kladde. 1998. GAL4p-mediated chromatin remodeling depends on binding site position in nucleosome but does not require DNA replication. *Mol Cel Biol*, 18(3): 1201-1212.