Possible Identification of Schizosaccharomyces pombe gene(s) Encoding POU Domain Transcription Factors

Charles E. Foulds

University Undergraduate Fellow, 1990-1991

Texas A&M University

Department of Biochemistry & Biophysics

Approved:

Fellows Advisor: David Peterson

Honors Program Director: D. V. Fowler

Acknowledgements

I would like to thank Kathy Beifuss, Chris Bral, Mike Huang for technical assistance. I thank Dr. Ryland Young for S.pombe strains and complementary DNA. Most of all, I thank Dr. David O. Peterson for letting me pursue individual research in his lab and for being a good mentor by showing interest in my future career.

Table of Contents

17. Figure 10

18. Table 1

Abstract

Eukaryotic RNA polymerase II transcription is characterized by a complex array of various transcription factors which recognize and bind to specific DNA sequence elements in promoters for accurate initiation of transcription and in enhancers for increased transcriptional rates. A novel family of transcription factors which share binding site specificities and a conserved domain is called the POU domain family. The goal of this study was to attempt to identify gene(s) in the fission yeast Schizosaccharomyces pombe that encode POU domain transcription factors. Two methodologies were employed. Southern blotting of yeast chromosomal DNA with probes derived from the cloned mammalian gene encoding Oct-1(which is a POU domain-containing factor) were done in an attempt to find homologous yeast nucleotide sequences. Results from these Southern blots suggested a weak cross hybridization of yeast DNA when the mammalian Oct-1 gene was used as a probe. Another attempt to find homologous yeast POU domain encoding gene(s) was based on polymerase chain reaction (PCR) of S.pombe cDNAs using a set of synthesized degenerate oligonucleotides corresponding to seven highly conserved amino acids in the POU domain. The other primer was the universal (-20) primer. The PCR products were cloned into a plasmid vector, and double stranded dideoxy DNA sequencing has revealed the orientation that a possible POU domain encoding gene has (3' to the insertion site). Currently, sequencing reactions using the forward (-40) primer are being done to sequence through this stretch of DNA to determine if it encodes a POU domain (by looking at the predicted amino acid sequences).

Introduction

Transcription in eukaryotic systems is not a simple process, but a very complex and dynamic one. Eukaryotic transcription differs from prokaryotic in a number of ways (such as 5' capping, 3' poly A tailing, and intron splicing). In addition, eukaryotes, unlike prokaryotes, have three types of RNA polymerases, RNA polymerases I, II, and III. RNA polymerase II(RNA pol II) is a very important enzyme due to its catalysis of the formation of mRNA precursors, which contain the genetic code for translation of functional proteins. This mRNA synthesis represents a major control point in the regulation of gene expression.

Two DNA sequence elements, promoters and enhancers, are necessary for the regulation of RNA pol II gene expression. Eukaryotic promoters are control regions of about 100 base pairs upstream from the transcriptional start site, which are responsible for accurate initiation of transcription (Maniatis et al., 1987). Promoters of most eukaryotic genes transcribed by RNA polymerase II contain an AT- rich sequence element termed the TATA box at about -25 with respect to the transcriptional start site (ie. +1). RNA pol II by itself cannot recognize the TATA box and initiate transcription. It needs five general proteins (TFIIA,B,D,E,and F) before accurate initiation of transcription can occur (for review, see (Mermelstein et al., 1989)). The TATA box is not always present such as in the case of the mammalian terminal deoxynucleotidyl transferase promoter and the promoter for SV40 late genes; another element overlying the start site is important for initiation in these cases (Dynan, 1989). Besides the TATA box, one or more 8-12 base pair specific protein binding sequences (called cis elements) can be found in eukaryotic promoters (e.g. GC box, CCAAT box, ATGCAAAT octamers) (Maniatis et al., 1987). These

additional promoter elements can increase the rate of transcription initiation either by working cooperatively with each other or independently in a given promoter (Dynan, 1989). All of the sequence elements together form the modular promoter characteristic of eukaryotic genes.

Eukaryotic transcriptional regulation is accomplished through the action of DNA binding proteins, some of which bind "upstream" from the promoter within enhancer sequences (discussed later) and some which bind to the promoter itself. These proteins, called transcription factors or trans-acting factors, modulate transcriptional rates through either enhancement or repression mechanisms. A fairly large number of such proteins have been partially to fully purified. Their sites of binding to DNA have been analyzed by such methods as DNA footprinting, gel retardation assays, and promoter mutagenesis (Maniatis et al., 1987). The mechanism of action of these transcription factors is currently being investigated, with the hope of understanding how these factors modulate transcriptional rates.

Eukaryotic transcription can further be stimulated by enhancer sequences, which can be very distant (sometimes thousands of base pairs) from the transcriptional start site on either the 5' or 3' side or in the middle of the gene. Enhancers lack any promoter activity of thier own but act to increase transcriptional rates from promoters (Maniatis et al., 1987). Recent work has suggested that enhancers are organized in much the same way as the promoters discussed earlier. That is, they both contain an assortment of DNA sequence elements to which transcription factors bind. It is interesting to note that some of the sequence elements in promoters and enhancers are interchangeable (e.g. the "octamer" in the

immunoglobin enhancer is found in a number of different promoters), suggesting that promoters and enhancers appear to be functionally related (Maniatis et al., 1987).

To summarize, the promoters and enhancers, which control transcription of genes in eukaryotes, are organized into multiple genetic elements. The transcriptional machinery is then able to integrate the regulatory information passed on by each element, allowing different genes to have specific, complex regulation (Dynan, 1989).

Schizosaccharomyces pombe is a type of yeast in which cell division occurs by fission. This yeast was investigated in this project due to its relative similarity to mammalian cells with regard to gene organization and structure. At least three S.pombe transcription factors have been identified to date (ie. PGA4, a factor homologous to mammalian AP-1, and TFIID) (Fikes et al., 1990; Ruden, 1990). Thus, S.pombe is a choice organism for investigating the conservation of basic elements in the RNA pol II transcription apparatus.

The goal of this study was to find S.pombe gene(s) encoding POU domain transcription factors. The reasoning that S.pombe might have gene(s) encoding such proteins was based on the two observations. First, a number of transcription factors have been shown to be highly conserved among eukaryotes. For example, TFIID is highly conserved from humans to Drosophila to Arabidopsis to yeasts (see Figure 1, (Hoffman et al., 1990)). Also, the POU domain (discussed later) is highly conserved from mammalian proteins(Pit-1, Oct-1, Oct-2, Brn-1, etc.) to a C.elegans protein unc-86 (see Figure 2, (He et al., 1989)). Secondly, it has been shown that S.pombe nuclear extracts contain a protein that specifically recognizes the octameric sequence 5'-ATGCAAAT-3' in a band shift assay

(Fee, 1989). Thus, a protein which recognizes the octamer exists in the fission yeast, and it is reasonable to assume that it contains a POU domain.

The POU domain is a 150-160 amino acid region which has a 75-82 amino acid POU-specifc domain, a short variable linker region, and a 60 amino acid homeodomain (see Figure 2). This domain was first found in three mammalian transcription factors(*P*it-1, *O*ct-1, Oct-2) and a C.elegans developmental gene product (*U*nc-86), and the domain has been identified in a rapidly growing number of proteins(eg. ceh-6, Tst-1, SCIP, Oct-3/4,6, Cf1a, Brn1,2,3) (He et al., 1989; Ruvkun and Finney, 1991; Monuki et al., 1990; Scholer et al., 1990; Johnson and Hirsh, 1990). Research on the POU proteins has been concerned with their regulatory roles in development, as well as to determine which domain of these proteins confers DNA binding specificity. Also, studies on how POU domain transcription factors activate transcription are currently underway.

The Pit-1 and octamer-binding proteins (Oct-1, 2, 3/4, 6) recognize distinct high affinity DNA sequences (Ruvkun and Finney, 1991). The Pit-1 protein binds to sites upstream of the growth hormone and prolactin genes (specifically a 5'-$(A/T)_4$TNCAT-3' sequence). Oct-1 and Oct-2 bind to a 5'-ATGCAAAT-3' sequence, which has been found adjacent to many genes (eg. in the immunoglobin heavy-chain promoter) (Lebowitz et al., 1988). The entire POU domain has been shown to be involved in DNA binding, but the POU-specific domain carries much of the binding specificity (Ruvkun and Finney, 1991).

POU domain proteins are important factors in the regulation of some genes. When Pit-1 or Oct-2 were expressed in HeLa cells (where these factors are not normally expressed), their expression was sufficient

to activate expression of prolactin, growth hormone, and octamer-containing promoters. Oct-1 is ubiquitously expressed in mammals and appears in many cell types (which correlates with the transcription of many generally expressed genes). The DNA-binding and transcriptional activation functions are separable in the Oct-1, Oct-2, and Pit-1 proteins. A glutamine-rich transcriptional activation domain exists in Oct-1 and 2, and Oct-2 and Pit-1 have a Ser/Thr-rich activation domain near to the POU domain (Ruvkun and Finney, 1991). Oct-1 and Oct-2 bind the same consensus sequence and have POU domains that are 87% identical, but they activate different sets of genes (Herr et al., 1988). Oct-1 activates transcription from ubiquitously expressed histone and snRNA genes, whereas Oct-2 activates transcription of the B cell-specific immunoglobin genes. Oct-1, in a novel case, can activate transcription of a model mRNA gene containing an octamer enhancer element by forming a complex with the herpes simplex virus (HSV) VP16 gene product. VP16 cannot bind DNA by itself, but it has an acidic transcriptional activation domain. VP16 recruits Oct-1 and other host proteins to form a complex capable of activating transcription of immediate early viral genes. Thus, Oct-1 can activate transcription at particular promoters by first binding to VP16 (which modifies the DNA binding and transcriptional activation abililites of Oct-1) (Ruvkun and Finney, 1991).

The POU gene unc-86 is required in several neuroblast lineages in nematodes for daughter cells to differentiate from their mothers. The Pit-1 gene is necessary for the growth of pituitary blast cells and for the differentiation of three pituitary cell types. Expression of Oct-3/4 (which are the same protein) is tied to development in the mouse. Expression is limited to large regions of the early embryo and adult germline.

Expression of Oct-3/4 drops off considerably when embryonic cell lines differentiate (Ruvkun and Finney, 1991). Four other POU genes (Tst-1, Brn-1, 2, 3) have been shown to be expressed during rat brain development (He et al., 1989).

As stated earlier, the goal of this study was to identify S.pombe gene(s) encoding POU domain transcription factors. Two methodologies were employed to try to answer this question. Southern blotting was done to try to find nucleotide sequences in S.pombe related to the mammalian gene encoding Oct-1 (a POU domain protein), and a polymerase chain reaction (PCR) approach was attempted in which one primer was a set of degenerate oligonucleotides encoding seven highly conserved amino acids (N-RVWFCNR-C) in the POU-homeodomain (see Figure 2). Details of the two methods used to search for a S.pombe POU domain encoding gene are discussed below.

Southern Blotting Approach

S.pombe Chromosomal DNA Isolation and Restriction Digestions

Two hundred ml of wild-type S.pombe (from Dr. Ryland Young's lab at TAMU) were grown in YE media (DIFCO yeast extract & glucose) to an OD_{595} of 0.5 with shaking at 32°C. The cells were harvested in 250 ml centrifuge bottles at 10,000 rpm for 3 minutes and were resuspended in 5ml of 20mM citrate/phosphate buffer, pH 5.6, 40mM EDTA, and 1.2M Sorbitol and transfered to a 50ml tube. Then, 15mg (Miles Laboratories) lysing enzyme were added to the tube and incubated at 37°C for 90 minutes. Cell lysis was verified by placing a 10µl sample with 1µl of 10%(w/v) SDS on a slide and viewing under a light microscope. The solution was then harvested at 5000 rpm for 5 minutes. The pellet was

resuspended in 5ml TE, pH7.6 (10mM TRIS-Cl, pH 7.6 & 1mM EDTA, pH 8.0). 250μL of 20% SDS and a spatula tip of Proteinase K was added to the solution. The solution was incubated at 56°C overnight with occasional agitation. Ten μl of RNase A (10mg/ml) were added to this solution and allowed to incubate at 37°C for 4 hours. Five ml (roughly one volume) of TE-saturated phenol were added to the yeast solution and then vortexed briefly. This solution was then centrifuged at 3000 rpm for 10 minutes. The aqueous (top layer) phase was drawn off and put into another tube. Five ml of saturated phenol and 2.5ml $CHCl_3$ (roughly .5 volume) were added to the solution. After vortexing briefly, the solution was centrifuged at 3000 rpm for 10 minutes. Again, the aqueous phase was drawn off and put in another tube. Five ml of $CHCl_3$ were added to remove excess phenol. After brief vortexing, the solution was centrifuged at 3000 rpm for 10 minutes. To the final aqueous phase (about 4-5ml), 1/10 volume of 2.5M NaOAc, pH 5.5 and 2.5 volumes of cold 95% ethanol were added. This solution was then centrifuged for 15 minutes at 3000 rpm at 4°C. After decanting off the 95% ethanol, the yeast DNA was washed in 2.5 volumes of cold 70% ethanol, vortexed, and centrifuged for 15 minutes at 3000 rpm at 4°C. The ethanol was decanted, and the yeast DNA was resuspended in an Eppendorf tube containing 1ml TE, pH 7.6. The tube was placed at 65°C for 10 minutes and then placed at -20°C overnight. The DNA in the solution was quanitated by making a 1/20 dilution and measuring the absorbance at 260nm. The absorbance reading then equates to a mg/ml concentration of ds DNA (this calculation was based on the fact that an OD_{260} of 1 equates to 50μg/ml ds DNA concentration (Sambrook et al., 1989)).

This S.pombe ds DNA (2.22µg/µl) was then digested separately with the following restriction enzymes, EcoRI, HindIII, and BamHI (from BRL Laboratories) , as described below. 30µl of S.pombe DNA (ie. 66.6µg) was digested at 37°C for one hour in a total volume of 35µl, with 3.5µl of 10X restriction enzyme buffer and 1.5 µl of the restriction enzyme. For the EcoRI digestion, REact 10x buffer 3 was used, and the BRL EcoRI enzyme had a concentration of 5U/µl. For the HindIII digestion, REact 10x buffer 2 was used, and the BRL HindIII enzyme's concentration was 5U/µl. BamHI had a concentration of 5U/µl, and REact 10x buffer 3 was used in the BamHI digestion. To check that the restriction digests were complete, a control was used. After 30 minutes incubation, 3µl of the three restriction digest reactions were added to .25µg of bacteriophage lambda DNA and allowed to incubate for one hour. Then, a 1%(x/v) agarose gel was run at 75V for several hours and stained in ethidium bromide. If the number of bands appeared in the gel as predicted, then the digestions with yeast DNA should have been complete. After the yeast DNA digests incubated for one hour, the DNA was ethanol precipitated (ie. two vol. of 100% ethanol, .25M NaCl final conc.) and placed at -20°C overnight. After centrifugation for 15 minutes in a microcentrifuge, the ethanol was decanted, and the pelleted DNA was resuspended in 30µl TE. The three digested DNA solutions of 30µl were then loaded into different wells of a 1%(w/v) agarose gel and were electrophoresed at 20V overnight (along with 270ng of lambda-HindIII size markers). The gel was then stained in ethidium bromide and a photo was taken (see Figure 3) . The gel was then denatured and neutralized by standard procedures (Sambrook et al., 1989). A Zetabind™ nylon filter was used in the DNA transfer scheme, instead of nitrocellulose (for more details on the Southern transfer, see (Sambrook

et al., 1989)). After the digested yeast DNA was transfered by capillary action to the nylon filter, the filter was removed from the gel, rinsed in 2x SSC, and baked for two hours in a vacuum oven at 80°C.

Random Primed Labeled pBS Oct-1 as a Probe

pBS Oct-1 is a pUC 18 derivative that contains the entire mammalian gene encoding the transcription factor Oct-1 discussed earlier. It was used as a probe in a Southern blot with the filter described above. 25ng of this plasmid were $\alpha[^{32}P]$dCTP labeled according to Boeringer Mannheim's random primed DNA labeling kit's protocol.

The filter described above was placed in a plastic bag with 20ml of annealing solution (Cocktails for formamide blots, 1991). This bag was placed on a shaker platform overnight at 42°C for prehybridization. Then, the nylon filter was taken out of the prehybridization bag and placed in another bag containing $5x10^6$ counts of the labeled pBS Oct-1 probe and 20ml of hybridization solution (Cocktails for formamide blots, 1991). The hybridization occurred at 42°C with shaking overnight. After the hybridization had occurred, the filter was washed (with shaking) two times for 15 minutes in a 1%SDS, 1X SSC solution at room temperature. Then, the filter was washed in a .1%SDS, .1X SSC solution twice for 30 minutes at 60-65°C with shaking. After the two washings were complete, the blot with pBS Oct-1 as a probe was placed on XAR-5 film (Eastman Kodak) for two weeks (see Figure 4).

PCR Amplification of a 300bp Fragment Encoding only the POU domain of Oct-1

A subclone called pJ-1/100 is a pUC-18 derivative which contains a 300bp fragment which encodes only the POU domain of Oct-1. E.coli cells containing this plasmid were grown on LB- carbenicillin plates overnight at 37°C. Then, the plasmids were amplified in rich medium with chloramphenicol (Sambrook et al., 1989) and then were harvested by centrifugation at 5000 rpm for 20 minutes at room temperature. Plasmid DNA was isolated by a alkaline lysis protocol, followed by CsCl-ethidium bromide density gradient centrifugation (Sambrook et al., 1989). After the plasmid DNA was isolated in a 15ml tube, two volumes of TE, pH7.6 were added to dilute the CsCl. The plasmid DNA was then ethanol-precipitated as described earlier and centrifuged for 30 minutes at 4°C at 8000 rpm. The ethanol was decanted, and the pelleted DNA was resuspended in 1ml TE, pH 7.6. Absorbance readings were taken at 260nm as described earlier, yielding a ds DNA concentration of .643mg/ml. Also, the absorbance at 280nm was .336; thus, the OD_{260}/OD_{280} ratio was 1.91, indicating good purity (Sambrook et al., 1989). Finally, the plasmid DNA was further purified from RNA by running it over a Biogel A-5m Econocolumn equilibrated in 1X TE, .2M NaCl. One ml fractions were collected, and absorbance readings at 260nm were taken to monitor the plasmid DNA. The plasmid DNA came out in fractions #3-6. These fractions were pooled, ethanol-precipitated, and centrifuged at 8000 rpm for 30 minutes at 4°C. The DNA was resuspended in 1ml TE, and absorbance measurements at 260nm of a 1:20 dilution yielded a [DNA] 0f .282mg/ml. This purified plasmid DNA was then used in a PCR reaction to amplify a possible POU domain gene. The PCR reaction had the following parameters: 59.7µl dd H_20, 10µl of 10X PCR buffer (Perkin Elmer Cetus), 1µl 100mM $MgCl_2$, 16µl of dNTPs mix (1.25mM in each dNTP), 1ng of pJ-1/100 plasmid DNA, 100

pmoles of the universal primer (5'-GTAAAACGACGGCCAGTGCC-3') and the reverse primer (5'-CAGGAAACAGCTATGACCATG-3'), and 2.5U of Taq polymerase (Perkin Elmer Cetus) for a total volume of 100µl. Three drops of paraffin oil were added on top of the solution. Thirty-five cycles were performed at 91°C for 1 min., 30 sec., 40°C for 2 min., 55°C for 1 min., and 72°C for 1 min. Ten µl of this reaction mix were run on a 1%(w/v) agarose gel to see if the reaction worked (data not shown). The PCR product was ~394 base pairs in size compaired to ϕx174 HaeIII size markers, which is roughly the size predicted for the insert.

Nick Translated 300bp Fragment as a Probe

The PCR product had a concentration of 1.20 µg/µl (as determined by OD_{260} readings). 2.1µl of a 1:10 dilution of this 300bp fragment (ie. .25µg) were used in the nick translation reaction, which included 2.5µl of a 10X nick translation buffer, 12.5µl of $\alpha[^{32}P]$dCTP, 1.5µl of .5mM dTTP, dGTP, and dATP, and 3.4µl of ddH_20 to make the final volume 25µl. The rest of the reaction details are given in Molecular Cloning: A Laboratory Manual (Sambrook et al., 1989).

More S.pombe DNA was isolated, digested with the same three restriction enzymes as described earlier, and run on an agarose gel in the same manner. Transfer of yeast DNA to a Zetabind™ filter proceeded in the same manner as described earlier. Prehybridization and hybridization (using the nick translated labeled 300bp fragment) conditions were similar to those described earlier, except that no formamide was used in either of the hybridization solutions and that the overnight shaking of the blot was done at room temperature (not 42°C) (Cocktails for aqueous blots, 1991). The washing conditions were different than described

earlier- 2x SSC, .5% SDS for 5 min. at room temperature, followed by 2x SSC, .1% SDS for 15 min. at room temperature. The blot was placed on XAR-5 film (Eastman Kodak) and placed at -70°C for four days. (The blot was not shown due to its low signal/noise ratio).

Controls on the Blots

Two main controls were implemented during the Southern blotting experiments. One control was to check that the Southern transfer was efficient and that indeed S.pombe DNA was on the Zetabind™ filters. This check consisted of nick translating pDB248New (a plasmid containing S.pombe DNA) to use as a probe with one of the Zetabind™ filters. The same hybridization and wash conditions used for the 300bp POU domain encoding fragment blot were used in this control. This control was developed after four days exposure to film; it showed strong hybridizations in the three lanes containing S.pombe DNA.

The second control was to check that the PCR produced probe described earlier did indeed encode a POU domain. This control consisted of digesting pBS Oct-1 with EcoRI and HincII, which will create four fragments. The smallest band is the 300 base pair POU domain encoding fragment. Southern transfer of these four bands onto a Zetabind™ filter was done as described earlier. When the nick translated PCR probe was used in a Southern blot with this new control filter, the bottom band was seen after film development.

PCR Approach

Amplification of S.pombe cDNA

Xi He, et al, described a PCR method (based upon two sets of degenerate oligonucleotide primers corresponding to nine highly conserved amino acid residues in both the POU-specific and POU-homeodomains) which identified four new members of the POU domain encoding gene family by using DNA complementary to human brain, rat brain and testes (see Figure 2 and (He et al., 1989)).

Dr. Ryland Young provided the S.pombe cDNAs used in this study. The cDNAs were cloned into pUC18 and theoretically could be in different orientations. Unlike the He experiment mentioned before, this study employed the use of only one set of degenerate oligonucleotides as one primer (called the POU primer) in the PCR reaction. The degenerate oligonucleotides were made with all possible codons in mind; they correspond to seven highly conserved amino acid residues in the POU homeodomain, which are N-RVWFCNR-C (see Figure 5). The other primer used in this study was the universal primer (already mentioned). The specific parameters used in the PCR reactions are as follows:

component	#1	#2	#3	#4	#5	#6
ddH_20(μl)	51.9	50.9	41.9	41.9	51.9	50.9
TMAC(M)	- -	10^{-6}	10^{-5}	10^{-4}	--	10^{-6}
10x buffer(μl)	10	10	10	10	10	10
dNTPs (μl of 1.25mM)	16	16	16	16	16	16
UNIV primer (μl of 40ng/μl)	13	13	13	13	13	13
POU primer (μl of $4x10^{-5}$)	2.5	2.5	2.5	2.5	2.5	2.5
pombe cDNA	10	10	10	10	10	10

(ng)						
Taq pol	1	1	1	1	1	1
(µl of2.5U/µl)						
Total Vol. (µl)	100	100	100	100	100	100

TMAC is an abbreviation for tetramethylammonium chloride. It was included in the PCR reactions because it has been shown to be a specificity enhancer for PCR reactions which reduces nonspecific priming by degenerate primers (Mody and Paul, 1990). 10x buffer and Taq polymerase were supplied from Perkin Elmer Cetus. Three drops of paraffin oil were added to these tubes. Thirty-five cycles were performed for reactions in tubes #1-4 at 91°C for 3.5 min., 50°C for 2 min., and 72°C for 3min. Reactions in tubes #5 and 6 had a different annealing temperature of 59°C and were performed for 40 cycles. Ten µl of each reaction were run out on a 4% polyacrylamide gel (PAGE) to see if the reactions had worked (see Figure 6).

Controls on the PCR Experiment

To check that the PCR amplification was specific due to the POU primer, the following controls were done. All PCR control reactions were similar to the ones described earlier, except that one component was omitted. Tube #1 did not have the POU primer in it; more ddH_2O was added to keep the total volume at 100µl. Tube #2 did not have the S.pombe cDNA template added to it. Tube #3 lacked the universal primer. Tube #4 did not lack any elements; the only difference it had with the reactions described earlier was the fact that only 3.6ng of S.pombe cDNA template

was used instead. A 1%(w/v) agarose gel was run on these four control reactions (data not shown).

Ligation of PCR Products and Transformation

Ligation of PCR products from the reactions #1-6 into pCR™1000 vector (TA Cloning™ System, Invitrogen Corp.) was performed as follows: 1.3µl of 10X ligation buffer (TA Cloning™ System), 1µl of pCR™1000 vector (25ng), 10µl of PCR reaction, and 1µl of T4 DNA Ligase (4U/µl) were added to six different Eppendorf tubes. The total volume of each tube was 13.3 µl, and these tubes were incubated at 15°C overnight. The pCR™1000 vector is presented in Figure 7; this vector has two 3' T overhangs which allow for more efficient cloning of PCR products (which contain 3' A residues due to a terminal deoxyadenosine transferase activity in Taq polymerase).

The transformation of the ligated vector into INVαF' E.coli cells (TA Cloning™ System) followed Invitrogen's TA Cloning™ Transformation protocol with the following exceptions: 13.3µl (not 1µl) of the six ligation reactions were pipeted into six different sets of INVαF' competent cells. In step #10, the vials were incubated for 2 min. at 43°C, not 1min. at 42°C. In step#13, vials were shaken at 37°C for 1.5 hours (not one hour) at 200 (not 225) rpm. White colonies were selected from the X-gal plates for further analysis.

Restriction Analysis of Putative Clones

Selected white colonies were grown in separate 5ml tubes in LB media (5g/L DIFCO yeast extract, 10g/L DIFCO tryptone, and 5g/L NaCl)

containing 50μg/ml kanamycin (the recombinant plasmids have kanamycin resistance). These tubes were incubated at 37°C overnight with agitation.

Boiling lysis mini-preps for isolation of plasmid DNA from each possible clone were then performed (Sambrook et al., 1989). After the plasmid DNA was isolated, it was digested with SacI (which will cut in the polylinker region of pCR™1000) for two hours at 37°C. Also, the pCR™1000 vector was digested with SacI at 37°C for two hours. A 1.5% (w/v) agarose gel was run at 20V overnight in which the size of the putative digested clones could be compared to the digested pCR™1000 vector (see Figure 8).

DNA Sequence Analysis of Putative Clones

The putative clones which showed an increased size compared to the pCR™1000 vector were further screened by double stranded dideoxy sequencing, with the hope that DNA sequences of these putative clones would show predicted amino acid sequences similar to the highly conserved stretch of amino acids in the POU homeodomain (see Figure 2).

Sequencing templates were prepared from earlier transformed bacteria. 1.5ml of an overnight culture was spun 3 minutes in a microfuge. After decanting the supernatant, the pellet was resuspended in 100μl cold STE (50mM glucose, 10mM EDTA, 25mM Tris-Cl, pH 8.0). 8 μl of a 10mg/ml lysozyme solution was added, and the solution was incubated for five minutes at room temperature. 200μl of a freshly-made .2N NaOH/1% SDS solution was added, and the mixture was incubated on ice for five minutes. 150μl of 3.0M KAc, pH 4.8 was added slowly to the mixture, which was incubated again on ice for 5 minutes. Cellular debris was removed by centrifugation for 5 minutes. The supernatant was removed to another

Eppendorf and incubated at 37°C with 2µl of a 10mg/ml RNase A solution for 20 minutes. Phenol-chloroform extractions followed. The plasmid DNA was then ethanol-precipitated. Then, the pellet was resuspended in 16µl of ddH_2O, and 3.2µl of 5M NaCl was added. 20µl of 13% polyethylene glycol-8000 was added to further purify the plasmid DNA. After centrifugation for 10 minutes, 30µl of the supernatant were removed. The remaining solution was ethanol-precipitated and spun for 10 minutes. The plasmid DNA pellet was the resuspended in 20µl ddH_2O. The double stranded plasmid DNA was then denatured by incubation with .2N NaOH/.2mM EDTA. The denaturing reaction was neutralized with NaAc, pH 5.5, and the plasmid DNA was recovered by ethanol precipitation. The DNA pellet was resuspended in 11µl ddH_2O. One µl of a 40ng/µl primer and 2µl of annealing buffer were added to the DNA solution. Pharmacia's ^{35}S $T7$Sequencing™ Kit was then followed exactly as written. The sequencing reactions were electrophoresed on an 8% PAGE sequencing gel containing 8.3M urea. The gel was given a two day exposure to XAR-5 film (Kodak). The primers used in different sequencing reactions were the universal primer, the M13 reverse primer (both were already mentioned), and finally the M13 forward (-40) primer (5'-GTTTTCCCAGTCACGAC-3') (see Figures 9& 10).

Results

Southern Blotting Approach. S.pombe chromosomal DNA was isolated and quantified as discussed earlier. The various restriction digests electrophoresed through a 1% agarose gel are shown in Figure 3. The Southern blot of yeast DNA using random primed labeled pBS Oct-1 as a probe is seen in Figure 4. Notice that several bands appeared in the

different lanes. This result suggested that S.pombe might have several genes homologous to the mammalian Oct-1 gene. However, the result was not conclusive in that other sequences in the pBS Oct-1 plasmid probe other than the Oct-1 gene could have possibly hybridized to sequences in yeast. In order to address this issue, the Southern blot of yeast DNA with the nick translated -labeled 300bp PCR amplified fragment encoding only the POU domain of Oct-1 was done. The blot was not shown because only one very faint band is visible in the yeast DNA/EcoRI lane. This result suggests that S.pombe might have DNA sequences homologous to the POU domain encoding probe. However, it is likely that the nucleotide sequences of POU domain genes are significantly different in S.pombe as compared to mammals accounting for the weak band intensity seen, but that the amino acid sequences may still be highly homologous (due to the degeneracy of the genetic code). This argument was the theoretical groundwork on which the PCR approach (described later) was based. The results from the two control blots were very conclusive. The Southern blot to see if yeast DNA was transfered to the nylon filter showed hybridization; thus, yeast DNA was attached to the filters in this study (blot not shown). The PCR amplified probe did indeed contain the gene encoding the POU domain of Oct-1 (ie. it did hybridize to the smallest fragment resulting from pBS Oct-1 digestion with EcoRI and HincII; blot not shown).

PCR Approach. Figure 6 shows that the PCR reactions described above did work; in fact, several products can be observed in each PCR reaction (lane). The fact that that the PCR products are of different sizes is not problematic because the cDNAs used in these reactions were also heterogeneous in size. The gel showing the electrophoresed control PCR

reactions is not shown but the results of the controls are summarized below. Without the cDNA template and the universal primer, no amplification occurred as expected. The gel also showed that when all components are present, the reaction works as expected. However, the interesting control reaction is Tube #1- the one without the POU primer. Amplification occurred in this tube. However, this finding is not discouraging, considering that PCR with one primer could result in single-stranded amplification products. Thus, the controls showed that the PCR reactions in Figure 6 amplified some template DNA with a degree of specificity, not randomness. After ligation of PCR products and then transformations as described above, ten white colonies (named α, β, γ, δ, θ, σ, ϕ, ε, μ, and κ) were seen on the X-gal plates. These ten white colonies were subjected to restriction analysis as mentioned earlier. Figure 8 shows that only four putative clones had inserts- β (originally from PCR reaction #1), γ (from reaction # 1), δ (from reaction #4), and θ (from reaction #3). Figures 9 and 10 show the DNA sequence gels with clone listed and the primer used in the reaction (e.g. R=reverse primer and F=universal primer). Table 1 lists the DNA sequences which were readible with the reverse primer. Note that the universal primer sequence shows up in all of the four clones sequence, indicating several things. First of all, this result conclusively shows that these clones have PCR products as inserts (because the universal primer was used in the PCR reactions). Secondly, this finding tells what orientation a possible POU domain encoding gene (that was amplified) has. If the reverse primer is used in sequencing reactions with the vector in Figure 7 and sequencing shows the universal primer's sequence near the 5' end of the insertion site, then a possible POU domain gene must be towards the 3' end of the insertion site.

Also, note that all four clones are different in sequence, which increases the likelihood of finding a POU domain gene.

The DNA sequencing reactions with the universal primer could not be interpreted because of many bands being in the same area. This multiple banding phenomenon was the result of having two universal primers in the same reaction. One primer came from the earlier described PCR experiment, and the second primer was mistakingly added to it. Since the universal primer sequencing reactions were not informative, DNA sequencing reactions of the four possible clones with the forward (-40) primer are currently underway to sequence the 3' end of the clones. The DNA sequences resulting from the forward primer reactions will immediately tell if any of the clones amplified in this study contain POU domain encoding gene(s) [by predicting amino acid sequences from those DNA sequences and looking for similarity with other published POU domain transcription factors].

Discussion

The results of this study show that nucleic acid hybridization experiments (such as Southern blotting) based upon using previously cloned mammalian genes encoding POU domains as probes with S.pombe DNA will not allow one to isolate a POU domain encoding gene in fission yeast (due to such a low signal/noise ratio as seen in Figure 4) by traditional screening of cDNA and genomic libraries. Instead, the better approach to attempt isolation of S.pombe gene(s) encoding POU domain factors (assuming the gene(s) exist) is with the tremendous specificity conferred by PCR and degenerate oligonucleotide primers synthesized from

conserved amino acids in the both the POU specific and POU homeodomains. The DNA sequencing reactions that are currently being performed with the forward (-40) primer will reveal whether the PCR approach worked, with the net result being at least one of the four putative clones mentioned earlier encoding a POU domain factor. If such a fission yeast POU domain gene is found by sequencing, then the next logical step to undertake would be to clone the entire gene (not just the fragment encoding the POU domain) by screening a S.pombe genomic or cDNA library with these PCR inserts as labeled probes. It would then be of interest to see if this fission yeast gene encodes a POU domain factor involved in yeast cell cycle regulation or other developmental processes (like other POU domain factors). If the forward (-40) primer sequencing reactions yield DNA sequences with predicted amino acids that are not similar to those found in the POU domain, then this would strongly suggest that S.pombe does not contain gene(s) encoding POU domain factors.

References

Cocktails for formamide blots. (1991). **Pre-hybridization** cocktail: 5ml 20x SSC, 2ml 100x Denhardt's, 2ml .5M $NaPO_4$ buffer, pH 6.4, 1g/20ml dextran sulfate, 10ml deionized formamide, .5ml 20%SDS, .25ml of 4.03mg/ml salmon sperm DNA, and .25ml ddH_2O. **Hybridization** cocktail: 5ml 20x SSC, .2ml 100x Denhardt's, .8ml .5M $NaPO_4$ buffer, pH 6.4, 2g/20ml dextran sulfate, 10ml deionized formamide, .5ml 20% SDS, .50ml of 4.03mg/ml salmon sperm DNA, and 3ml ddH_2O ((UnPub)

Cocktails for aqueous blots. (1991). **Pre-hybridization** cocktail: 3.75ml 20x SSC, 1.5ml 100x Denhardt's, 1.5ml .5M $NaPO_4$ buffer,pH 6.4, .75g/15ml dextran sulfate, .375ml 20%SDS, .188ml of 4.03mg/ml salmon sperm DNA, and 7.687ml ddH_2O. **Hybridization** cocktail: 3.75ml 20x SSC, .150ml 100x Denhardt's, .60ml .5M $NaPO_4$ buffer,pH 6.4, 1.5g/15ml dextran sulfate, .375ml 20%SDS, .375ml of 4.03mg/ml salmon sperm DNA, and 9.75ml ddH_2O. ((UnPub)

Dynan, W.S. (1989). Modularity in Promoters and Enhancers. Cell *58*, 1-4.

Fee, Brian (1989). Unpublished data on band shift assays of S.pombe nuclear extracts with the octamer 5'-ATGCAAAT-3'.

Fikes, John D., Becker, D.M., Winston, F., and Guarente, L. (1990). Striking conservation of TFIID in Schizosaccharomyces pombe and Saccharomyces cerevisiae. Nature *346*, 291-294.

He,Xi. and Rosenfeld, M.G. (1989). Expression of a large family of POU-domain regulatory genes in mammalian brain development. Nature *340*, 35-42.

Herr, W., Sturm, R.A., Clerc, R.G., Corcoran, L.M., Baltimore, D., Sharp, P.A., Ingraham, H.A., Rosenfeld, M.G., Finney, M., Ruvkun, G., and Horvitz, H.R. (1988). The POU domain: a large conserved region in the mammalian pit-1, oct-1, oct-2, and Caenorhabditis elegans unc-86 gene products. Genes Dev. *2*, 1513-1516.

Hoffman, A., Sinn, E., Yamamoto, T., Wang, J., Roy, A., Horikoshi, M., and Roeder, R.G. (1990). Highly conserved core domain and unique N terminus with presumptive regulatory motifs in a human TATA factor (TFIID). Nature *346*, 387-390.

Johnson, W.A. and Hirsh, J. (1990). Binding of a Drosophila POU-domain protein to a sequence element regulating gene expression in specific dopaminergic neurons. Nature *343*, 467-470.

Lebowitz, J.H., Kobayashi, T., Staudt, L., Baltimore, D., and Sharp, P.A. (1988). Octamer-binding proteins from B or HeLa cells stimulate transcription of the immunoglobulin heavy-chain promoter in vitro. Genes Dev. *2*, 1227-1237.

Maniatis, T., Goodbourn, S., and Fischer, J.A. (1987). Regulation of inducible and tissue-specific gene expression. Science *236*, 1237-1245.

Mermelstein, F.H., Flores, O., and Reinberg, D. (1989). Initiation of transcription by RNA polymerase II. Biochim. Biophys. Acta *1009*, 1-10.

Mody, R. and Paul, S. (1990). A Specificity Enhancer for Polymerase Chain Reactions that reduces nonspecific priming by degenerate primers. Clontech Laboratories,Inc. *V*, 1-2.

Monuki, E.S., Kuhn, R., Weinmaster, G., Trapp, B.D., and Lemke, G. (1990). Expression and Activity of the POU Transcription Factor SCIP. Science *249*, 1300-1303.

Ruden, D.M. (1990). Identification of Schizosaccharomyces pombe Transcription Factor PGA4, Which Binds Cooperatively to Saccharomyces cerevisiae GAL4-Binding Sites. Mol. Cell. Biol. *10*, 1432-1438.

Ruvkun, G. and Finney, M. (1991). Regulation of Transcription and Cell Identity by POU Domain Proteins. Cell *64*, 475-478.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, 2nd ed. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Scholer, H.R., Ruppert, S., Suzuki, N., Chowdhury, K., and Gruss, P. (1990). New type of POU domain in germ line-specific protein Oct-4. Nature *344*, 435-439.

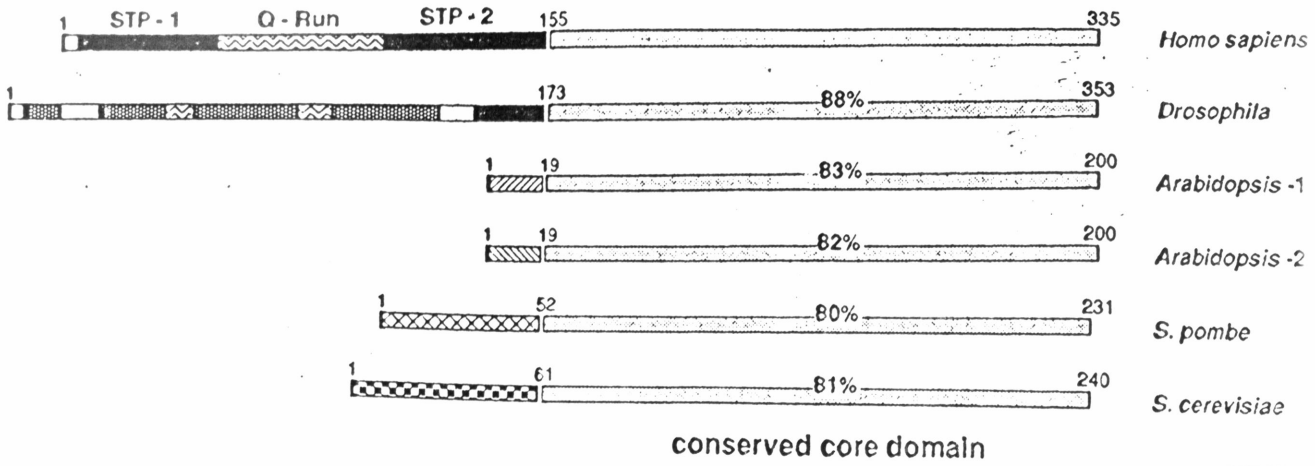


FIGURE 1. **The primary structure of the TATA factor (TFIID) is highly conserved among species.** Percentages reflect the degree of sequence identity (relative to human) within the C-terminal conserved core domains of TFIID from human to yeasts.(Taken from: Hoffman, et al, p.390).

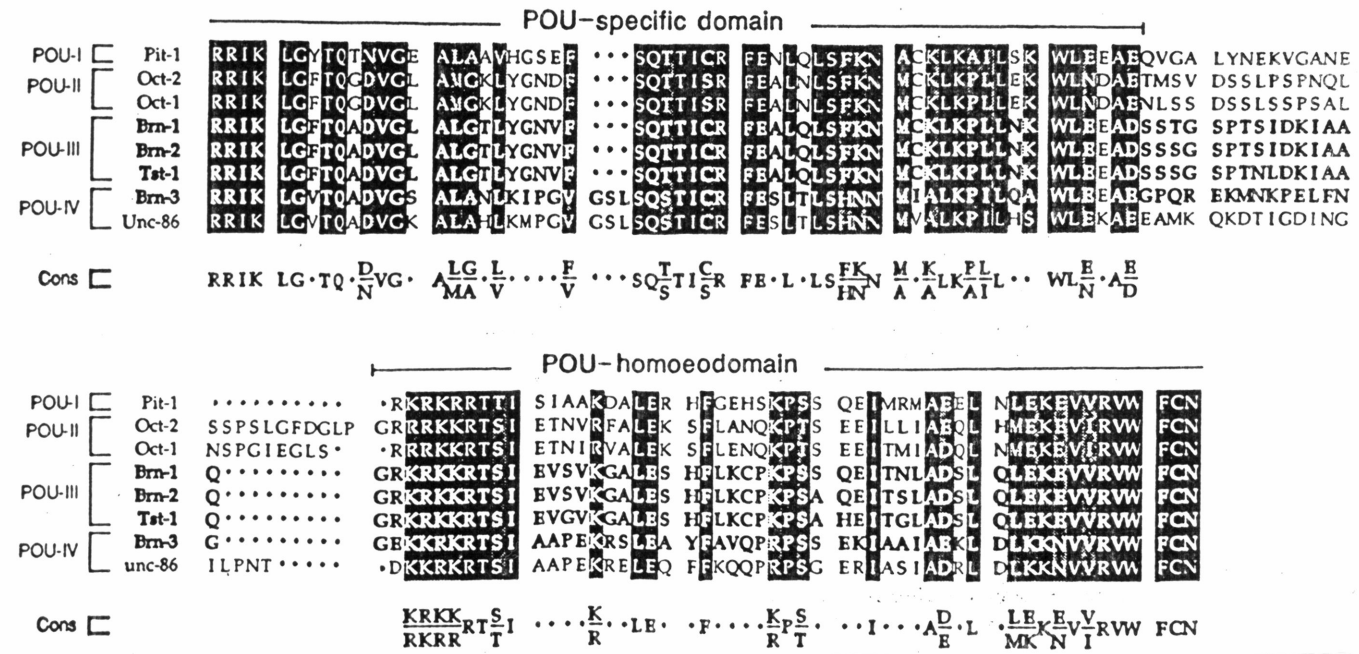


FIGURE 2. **Partial amino-acid sequences of eight characterized POU domain proteins showing two highly conserved subdomains.** The black shading indicates highly conserved residues. Note that the POU domain is subdivided into a POU-specific domain and a POU-homeodomain. Consensus sequences are given at the bottom of each subdomain. (Taken from He, et al, p.36).

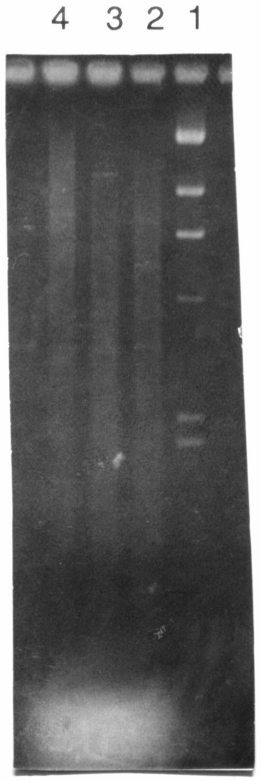
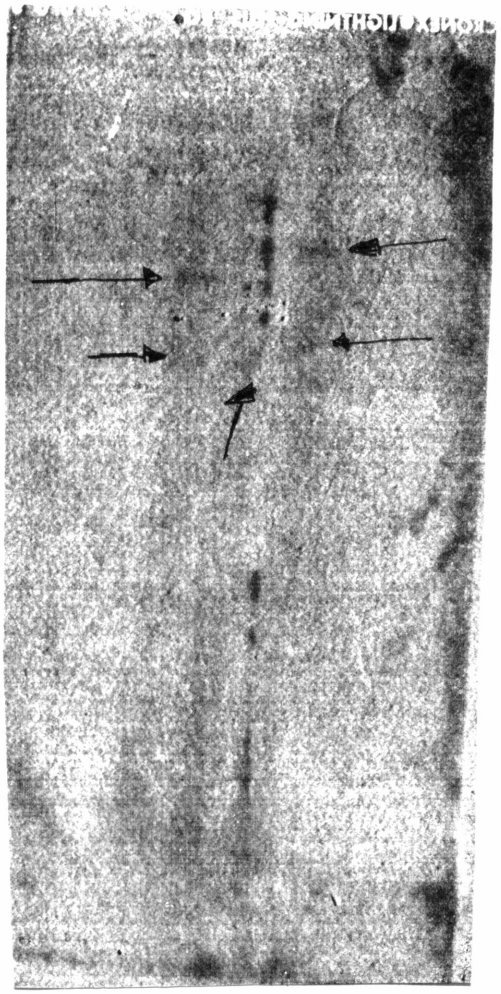


FIGURE 3. **Preparation for Southern Blotting: S.pombe chromosomal DNA isolation and restriction enzyme digestions.** Lanes #1-4 of this 1%(w/v) agarose gel represent the following: lane #1 contains 270ng bacteriophage lambda digested with HindIII, lane #2 contains 66.6µg of S.pombe chromosomal DNA digested with EcoRI, lane #3 contains 66.6µg of S.pombe chromosomal DNA digested with HindIII, lane #4 contains 66.6µg of S.pombe chromosomal DNA digested with BamHI. The gel was electrophoresed at 20V overnight.



4 3 2

FIGURE 4. **Yeast Southern blot using random primed pBS Oct-1.** Lanes#2-4 contain the following: lane #2, S.pombe chromosomal DNA digested by EcoRI;lane #3, yeast DNA digested by HindIII; lane #4, yeast DNA digested by BamHI. The blot was allowed to stay on film for two weeks. The results were the following: lane #4 had two hybridizations in it; lane #3 had only one; and lane #2 had two bands of hybridization. The three lower bands had the same intensity and can barely be seen (look at the arrows). Thus, only weak cross- hybridizations could be seen.

N----------->C

R	V	W	F	C	N	R
(ARG)	(VAL)	(TRP)	(PHE)	(CYS)	(ASN)	(ARG)
[C]GT	GTT	TGG	TTT	TGT	AAT	CG[T]
[C]GC	GTC		TTC	TGC	AAC	CG[C]
[C]GA	GTA					CG[A]
[C]GG	GTG					CG[G]
[A]GA						AG[A]
[A]GG						AG[G]
5'G4	GT4	TGG	TTT/C	TGT/C	AAT/C	C/AG3'
3'C4	CA4	ACC	AAA/G	ACA/G	TTA/G	G/TC5'

FIGURE 5. **The POU primer as a set of degenerate oligonucleotides corresponding to seven highly conserved amino acids in the POU homeodomain.** The amino acid residues listed above are in the N-terminus to C-terminus direction. These seven residues are invariant in every known POU domain protein. Note that brackets indicate respective nucleotides that were omitted during synthesis to reduce the amount of unwanted degenerancy without giving up specificity. A 4 indicates that all nucleotides (A,G,C,and T) are needed to satisfy the degeneracy inherent in the genetic code for that amino acid.

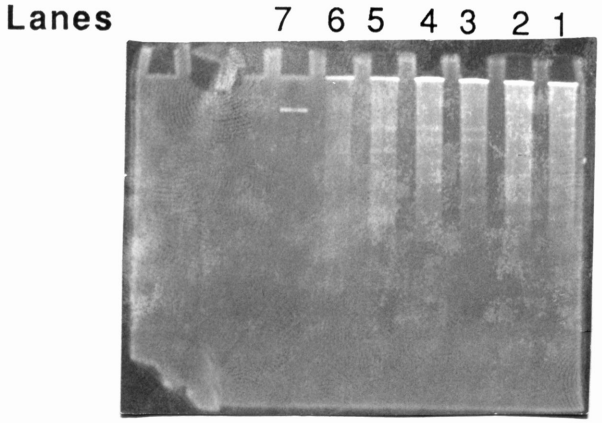


FIGURE 6. **4% Polyacrylamide gel to see if the PCR experiment resulted in amplification products.** Contents of lanes#1-7: lanes # 1-6, 10µl of PCR reactions described in text; lane #7, 2µl (which equals 50ng) of pCR™1000 vector. As the reactions amply show, amplification of products occurred in every PCR mix.

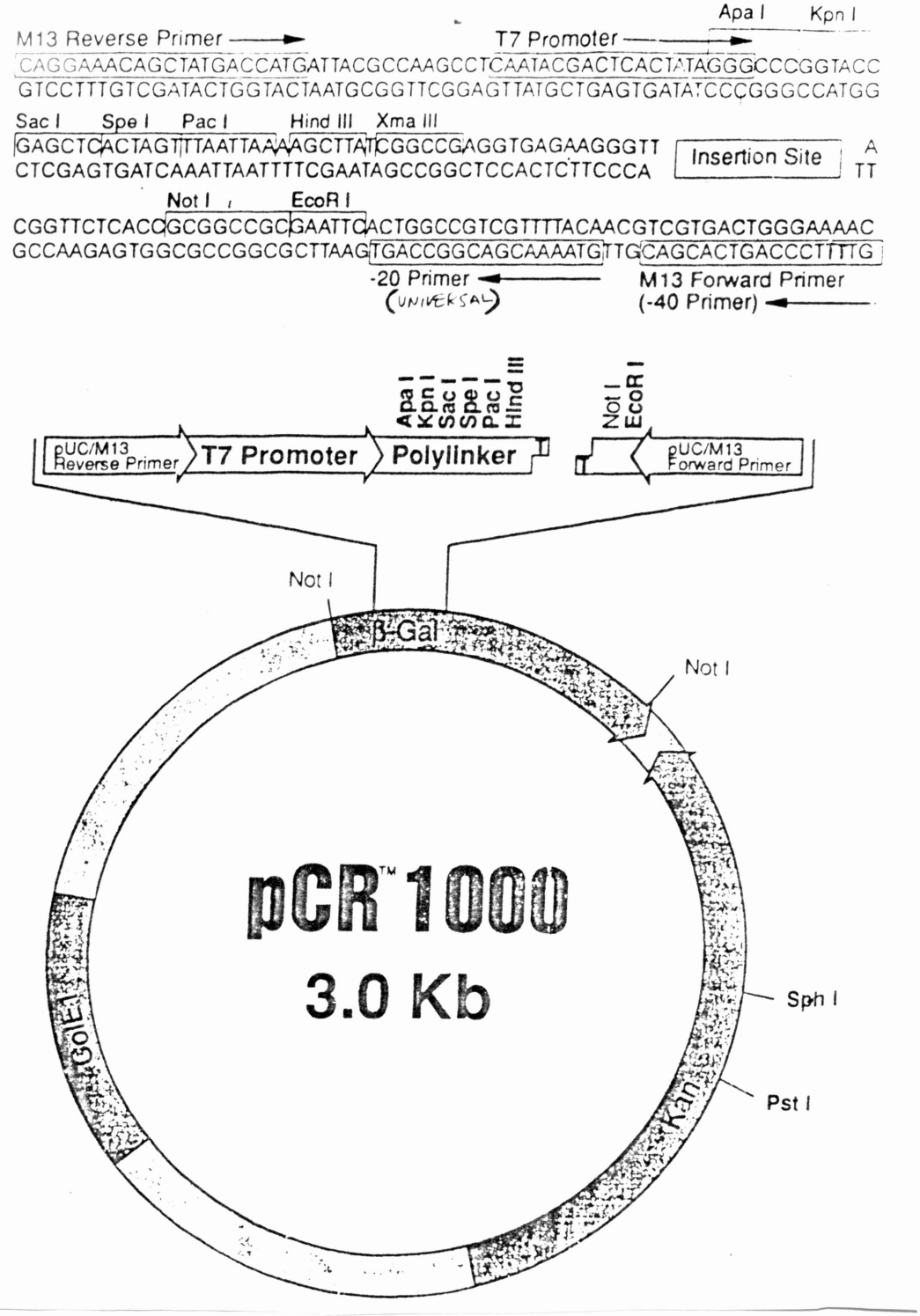


FIGURE 7. **TA Cloning™ vector pCR™1000.** This plasmid vector has special $^{3'}$T overhangs which allow for more efficient cloning of PCR products (which contain $^{3'}$ A residues). Note that this vector has a kanamycin resistance gene and a β-Gal gene, so that selective screening of recombinant clones is straightforward. Also, the vector has sites for three primers to hybridize, which was very important in this study.

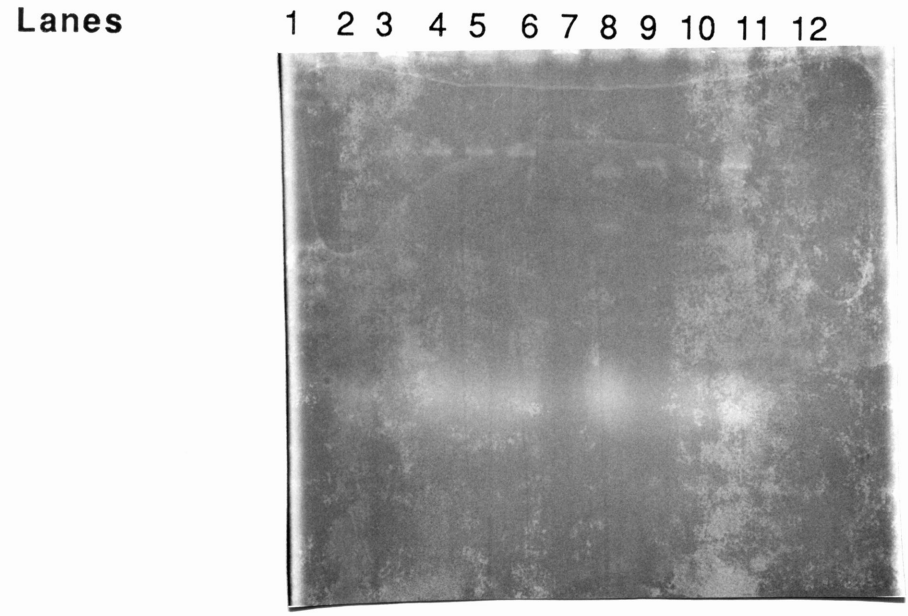


FIGURE 8. **Restriction analysis of putative clones**. Contents of the 1.5%(w/v) agarose gel: lane#1, lambda-HindIII + phi-x174-HaeIII size markers, lane#2-11, α,β,γ,δ,θ,σ,φ,ε,μ,and κ plasmid DNA digested with SacI (20μl), and lane#12, pCR™1000 (50ng) digested with SacI also. The gel shows clearly that β,γ,δ, and θ (lanes#3-6) are larger than the pCR™1000 vector (which suggests that they have inserts).

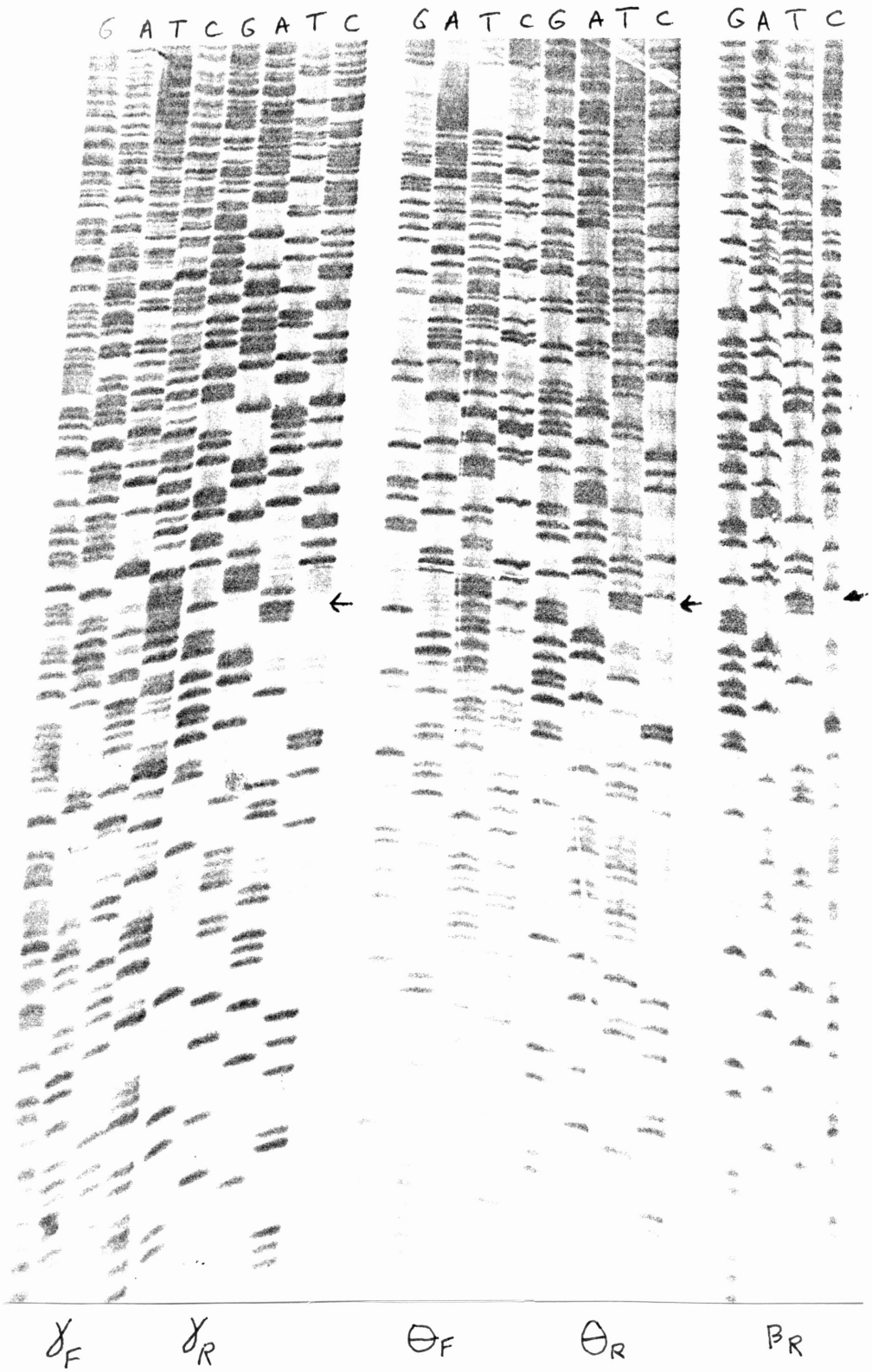


FIGURE 9. **DNA sequence analysis of putative clones.** Contents of the 8% PAGE, 8.3M urea sequencing reactions: γ^F,γ^R,θ^F,θ^R,and β^R from left to right. Note that the arrows indicate where the 5' insertion site of the clones begins. (F= universal primer reaction, R= reverse primer reaction).

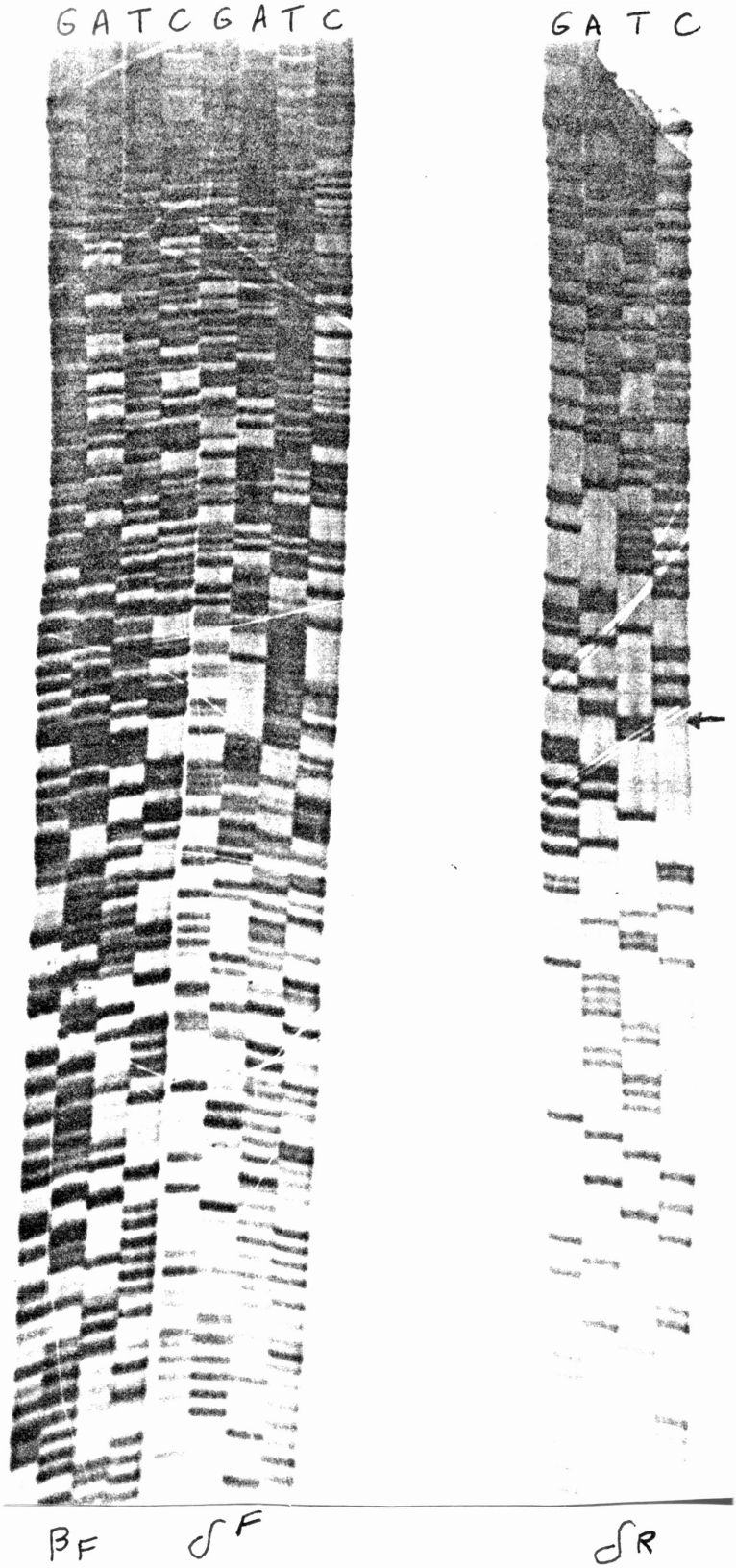


FIGURE 10. **DNA sequence analysis of putative clones.** Contents of the 8% PAGE, 8.3M urea sequencing reactions: β^F,δ^F, and δ^R from left to right. Note that the arrows indicate where the 5' insertion site of the clones begins. (F= universal primer reaction, R= reverse primer reaction).

β

5'TCGATATCGAG**[GTAAAACGACGGCCAGT]**GCAAGCTTGGCTGCA
GGTCGACGGATCCGGAATCTACTCTAAGTAACTAGACTCTATAGCTCA
CGTACTATATACACTATCAGTCG3'

δ

5'T**[AACGACGGCCAGT]**GGAAACCGCCTCTTCTGTCTCGGCATCTCTC
AGTATCAGCACGACTATACGCTCGATCACAGTGCTACAGTGCGTGTGT
CATGCGATCG3'

γ

5'T**[GTAAAACGACGGCCAGT]**GCGATAAGTCGAGTCTACGGTGACT
CAGACGATAGTACGATAGCGCACTCGCTGACGGTCGTGCACACAGCA
GCTGACGACGACTACGACTGATC3'

θ

5'TCGATATCGAGAG**[GTAAAACGACGGCCAGT]**GTGAAGAGTGTTC
GTCAGTATTCACGATGATATCAGTGTCAGTGTCGTGCTCACAGTATAC
GTAATGCTCACATGATCGATCGATCTCA3'

TABLE 1. **DNA sequence analysis of four putative clones using the M13 reverse primer.** Note that the following sequences are read in the 5' to 3' direction of the top strand of the pCR™1000 vector starting at the 5' insertion site. The sequence of the universal primer (5'-GTAAAACGACGGCCAGT-3') is seen in all four clones in bold case brackets and letters. This result conclusively shows that the cloned inserts are from the earlier described PCR reactions, as well as suggests that a possible POU encoding gene is at the opposite side of the insertion site (ie. the 3' end). Also, notice that all four clones have different sequences, which increase the likelihood of finding a POU domain gene.