Seton Hall University
eRepository @ Seton Hall

Seton Hall University Dissertations and Theses (ETDs)

Seton Hall University Dissertations and Theses

2006

Identification of Proteins Interacting with the Poly C Binding Protein 1 in the Central Nervous System Using Yeast Two Hybrid Screening System

Joanne Eghenayahiore Odiase
Seton Hall University

Follow this and additional works at: https://scholarship.shu.edu/dissertations

Part of the Biology Commons

Recommended Citation

Odiase, Joanne Eghenayahiore, "Identification of Proteins Interacting with the Poly C Binding Protein 1 in the Central Nervous System Using Yeast Two Hybrid Screening System" (2006). *Seton Hall University Dissertations and Theses (ETDs)*. 670.
https://scholarship.shu.edu/dissertations/670

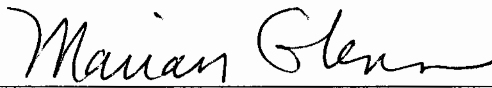
Identification of Proteins Interacting with the Poly C Binding Protein 1 in the Central Nervous System Using Yeast Two Hybrid Screening System.

By
Joanne Eghenayahiore Odiase

Submitted in partial fulfillment of the requirements of the degree of Master of Science
from the Department of Biology of Seton Hall University
December 2005

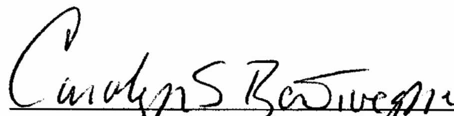
Approved By:

Dr. Jane Ko, PhD
Mentor



Dr. Marian Glenn, PhD
Committee Member

Dr. Allan Blake, PhD
Committee Member



Dr. Carolyn Bentivegna, PhD
Chair, Department of Biology

Acknowledgements

I would like to extend my highest appreciation to my mentor Dr. Jane L. Ko for her patience, support and guidance throughout this research. I would also like to thank the members of my thesis committee, for their time and recommendation regarding this thesis. Finally, I would like to especially thank my family and friends for their encouragement and support throughout this research.

Table of Contents

Section:	Page:
Abstract	6
Introduction	8
Methods and Materials	12
Results	17
Discussion	32
Conclusion	36
References	37

List of Figures

Figure:	Page:
Figure 1	18
Figure 2	20
Figure 3	24
Figure 4	26
Figure 5	29
Table 1	31

Abstract

Poly C binding proteins which belong to the KH domain family, are known to be involved in a wide range of biological processes such as mRNA stabilization, translational enhancement and inhibition. In addition, poly C binding protein 1 (PCBP) has been reported to act as a transcriptional regulator, which participates in the mouse mu-opioid receptor gene regulation. To further understand the physiological role of PCBP, we aimed to identify the proteins that interacted with PCBP using yeast two-hybrid method to screen a mouse cDNA library with the PCBP as a bait protein. The library was screened under low, medium and high stringent selections, and three candidate clones (named HY87, HY94 and HY18) survived the high stringent selection. Among these three clones, only one pACT2-cDNA plasmid (containing cDNA fragments) from HY87 clone was successfully isolated, and subjected to DNA sequencing.

Interestingly, this HY87 cDNA sequence matched the partial sequences of the mouse chromosome 5 published in NCBI database. However, the NCBI database search also indicated that this cDNA fragment shares no significant homology to any known proteins. It was then further subjected to the open reading frames (ORF) program. In total, 4 different predicted amino acid sequences were obtained and were further subjected to Blast Protein/Protein databases. Only one sequence, entitled Sequence 2, displayed a partial homology (46%) to both the arginine biosynthesis bifunctional protein (Arg protein) and the cell division ftsY homolog (CDF homolog). Furthermore, Sequence 2 was found to have a homology to the N-domain localized in the CDPF homolog,

suggesting that Sequence 2 potentially may be involved in protein-protein interaction and related to protein trafficking.

Introduction

There are numerous methods used to identify protein-protein interactions. One of these methods is the yeast two hybrid system, which is a popular functional assay based on protein-protein interaction. This will allow us to screen an unknown protein from a cDNA library by using a known protein as the bait. The two hybrid system has been successfully used to identify diverse sets of interacting proteins in yeast and mammalian cells. Additionally, it has been particularly successful in studies of oncogenes, such as the c-myc promoter binding protein 1 (Ghosh et al., 2001), tumor suppressors such as the neurofibromatosis type 2 gene (Jannatipour et al., 2001) and enzymes such as the phosphodiesterase 6 (Morin et al., 2003). In this study, we investigated the poly C binding protein 1 (PCBP), which belongs to a member of the KH domain family, using the yeast two hybrid screening method.

The KH domain has been identified in a wide spectrum of RNA-binding proteins (Gibson et al., 1993). This domain was identified as a repeated 45-amino-acid motif in hnRNP K initially (Siomi et al. 1993) subsequent alignments and structural studies have defined an extensive 68-72 residue KH "maxi domain" (Musco et al., 1996). NMR spectroscopic and X-ray crystallographic analyses of a number of KH domains (Musco et al., 1996, 1997; Baber et al., 2000; Lewis et al., 1999; Wimberly et al., 2000) have revealed that the initially described 45-residue core is configured as a $\beta\alpha\alpha\beta$ unit. The data further reveal two KH domain subtypes based on the more extensive structure: The type I KH domain (e.g., KH-3 of hnRNP K) includes a C-terminal $\beta\alpha$ extension and the type II KH domain (e.g., ribosomal protein S3) contains an N-terminal $\alpha\beta$ extension

(Grishin, 2001). Two unstructured surface loops extend from the type I KH domain, one containing an invariant GXXG and a second of variable length and sequence. These loops may be of central importance in determining RNA binding specificity (Adinolfi et al., 1999).

There exists evidence that proteins containing KH domains (belonging to the KH domain family) are involved in important regulatory mechanisms which in some cases are associated with human diseases such as the humans FMR1 (fragile X mental retardation) gene product which is associated with the Fragile–x-Syndrome (Siomi et al, 1994). A reported mutation in one of the KH domains of the FMR1-protein impairs its RNA binding activity and results in mental retardation (Siomi et al., 1993). Another example of KH proteins is NOVA, a neuronal protein, which is recognized by an autoantibody of breast and lung cancer patients who develop a disorder of motor control. A third example of KH proteins is gld-1 which is a tumor suppressor gene identified in *Caenorhabditis elegans.* A common feature of many KH proteins is their ability to bind RNA containing poly rC sequence *in vitro* (Buckanovich et al., 1993).

Among the KH domain family, the hnRNP K protein is a well-characterized member, which has been reported to possess the binding capability of poly rC RNA and for single-stranded DNA (Shnyreva et al., 2000). In addition to its diverse binding profile, using the yeast two hybrid screening system hnRNP K has been reported to interact with several protein kinases, such as Src (Van Seunigen et al., 1995), protein kinases C (Schullery et al, 1999), and the proto-oncogene *vav* (Bustelo et al, 1995), implicating the involvement of different cellular processes. Furthermore, hnRNP K protein can recruit other factors to stimulate or repress transcription, such a TATA-

binding protein (TBP) (Michelotti et al, 1996), Sp1 (Michelotti et al, 1996), and the transcriptional repressor, Zik1 (Denisenko et al, 1996).

Recently, hnRNPK has been shown to function as transcription regulator and is involved in transcriptional controls of the human *c-myc* gene (Michelotti et al., 1996). It has the ability to interact with a C-rich DNA sequence termed the CT element, which is upstream of the *c-myc* gene. (Michelotti et al., 1996; Tomonaga & Levens, 1996). Moreover, phosphorylation of hnRNP K by mitogen-activated protein kinase/extracellular-signal-regulated kinase (MARK/ERK) stimulates cytoplasmic accumulation of hnRNP K under physiological growth conditions (Habelhah et al., 2001). A shift of hnRNP K from nucleus to cytoplasm establishes its function as a silencer of LOX mRNA translation (Makeyev et al., 2002). Taken together, these results suggested that hnRNP K is a versatile protein and may act as a "docking platform" mediating cross-talk between various molecules

The notion that hnRNP K served as a docking platform is further supported by its structural study. For example, the interaction of hnRNP K with SH3 domains from Src, Fyn, Lyn, and Vav as well as transcriptional repressors and Y-box-binding protein is found to be mediated by the centrally located proline-rich domain, denoted KI (Denisenko et al., 1997), and TATA-binding protein associates with hnRNP K via a site adjacent to KI domain (Shnyreva et al., 2000). The interaction of hnRNP K with αCP-2, hnRNP L, hnRNP I, or the transcriptional factor CCAAT/ enhancer-binding protein β are mediated by the third KH domain located at the N-terminius of hnRNP K (Miau et al., 1998; Kim et al., 2000).

In this study, we are interested in examining the interacting protein of another multi-functional protein, poly C binding protein 1 (PCBP), which is also a member of the KH domain family. PCBP is a multi-function protein acting as an important regulator of mRNA stability and translational regulation. It recognizes C-rich RNA through its KH domain and is thought to carry out their function through its interaction with other RNA-binding proteins (Sidiqi et al., 2005).

The PCBP has recently been reported to interact with a 26-bp polypyrimidine sequence located in the mouse mu-opioid receptor (MOR) proximal promoter region and can serve as a transcription regulator of the MOR gene expression (Ko and Loh, 2005). The MOR is a member of the G protein coupled receptor which is mainly found in the central nervous system, mediating various important pharmacological effects such as analgesia and addiction (Ko et al., 1998). The PCBP shares many similar characteristics as hnRNP K, such as a similar overall topology of hnRNP K, consisting of two KH domains grouped near the N-terminus and a third KH domain located at the C-terminus (Funke et al., 1996). To further understand the physiological roles of PCBP in MOR gene regulation, we therefore performed a yeast two hybrid screening using a mouse brain cDNA library to identify the potential interacting protein(s) of PCBP.

Materials and Methods

Bacterial Transformation

The DNA plasmid was added to 50 µl of *E. coli* competent cells and was incubated on ice for 30 minutes. The tube of cells was incubated at 42°C water bath for 60 seconds and then back on ice for 60 seconds. Five hundred milliliters of Luria Broth (LB) medium was added to cells and was incubated at 37°C with shaking about 225 rpm for 60 minutes. The cells were spread onto LB/amp (50 µl/ml) plate and incubated at 37°C overnight.

Small Scale Plasmid Preparation

Colonies were picked off the plate and grown in LB media with ampicillin overnight at 37°C. The plasmid was purified from an overnight culture using QIAprep Spin Miniprep Kit purchased form Qiagen (Valencia, CA). The cells were harvested by centrifugation and were resuspended in "Resuspension Buffer P1" (50 mM glucose, 10 mM EDTA, pH 8, 25 mM Tris HCL, pH 8 with RNase A). After resuspension, cells were lysed in "Lysis Buffer P2" (0.2 M NaOH with 1% SDS) and neutralized in "Neutralizing Buffer N3" (5 M potassium acetate, glacial acetic acid). The cell mixture was then centrifuged for 10 minutes at maximum speed in a tabletop microcentrifuge. The plasmid was then purified using the QIAprep column, which was washed using Buffer PE. The plasmid was eluted by adding 50 µl of Buffer EB (10 mM Tris HCl and 1mM EDTA, pH8.5).

Large Scale Plasmid Preparation

Cultured transformed *E coli* colonies were added to 250 ml of LB media with ampicillin and were allowed to grow overnight at 37°C. The plasmid in cell culture was purified using QIAGEN Plasmid Maxiprep Kit purchased form Qiagen (Valencia, CA) according to the manufacture's instructions. Briefty, the cells were harvested by centrifugation and were resuspended in 10 ml of resuspension Buffer P1 (50 mM glucose, 10 mM EDTA, pH 8, 25 mM Tris HCL, pH 8 with RNase A). Ten milliliters of lysis Buffer P2 (0.2 M NaOH with 1% SDS) was added to the cells and 10 ml of chilled Neutralizing Buffer P3 (5 M potassium acetate, glacial acetic acid) was also added. The cell mixture was then centrifuged, and the supernatant was applied onto the QIAGEN-tip column, which was washed twice with Buffer QC. The plasmid DNA was eluted from column by adding Buffer QF. The eluted plasmid DNA was then precipitated via alcohol precipitation. The DNA pellet was redissolved using TE buffer (1 mM EDTA and 10 mM Tris HCl, pH 8).

Construction of bait plasmid pGBKT7-CBP

pGBKT7-PCBP (shown in Figure 1) was constructed by cloning the PCBP cDNA fragment into the multiple cloning site of the yeast expression vector, pGBKT7 plasmid. Restriction enzyme analysis was performed to confirm the insertion of PCBP into pGBKT7 vector.

Small-scale yeast mating of positive and negative controls

One colony of AH109[pGBKT7-53] (AH109-p53) and Y187[pTD1-1] were each picked to use in the positive control mating. One colony of AH109[pGBKT7-Lam] (AH109-pLam) and Y187[pTD1-1] were each picked to use in the negative control mating. One colony of each was placed into 2X YPDA medium to create nonmating cultures. One colony of AH109-p53 and pGBKT7-Lam were each placed into 2X YPDA medium together to create mating cultures. All cultures grew at 30°C overnight. One hundred microliter aliquots of 1:10 and 1:100 dilutions of each overnight mating and nonmating cultures were spread onto nutritional selection medium as explain in detail in Result section. Plates were incubated at 30°C until visible colonies appeared.

Lithium Acetate-mediated yeast transformation

Yeast cells were allowed to grow overnight at 30°C in YPD medium (Clontech). The overnight culture was then transferred to a new flask containing 300ml of YPD. To ensure the yeast was in the log phase of growth, the OD_{600} of the diluted culture was checked to ensure culture had reached between 0.2 - 0.3. Cells were centrifuged and then the cell pellet was resuspended in lithium acetate (LiAc) solution (0.1 M Tris-HCl, 10mM EDTA, 1M LiAc, pH 7.5). The yeast cells were then incubated with plasmid and carrier DNA in PEG/LiAc solution (50% PEG, 1M LiAc, pH 7.5) at 30°C for 30 minutes. After incubation, 20 μl dimethyl sulfoxide (DMSO) was added, and the cells underwent heat shock for 15 minutes in a 42°C water bath and were chilled on ice for 1-2 minutes. The cells were centrifuged and then resuspended in TE buffer. These cells were then plated on nutritional selection plates as explain in detail in Result section.

The library screening using yeast two-hybrid system

The AH109 yeast cells containing the pGBKT7-PCBP grew in nutritional selection medium SD/-Tryptophan at 30°C overnight. The pretransformed mouse brain library (Clontech) using pACT 2 vector in Y187 (2.2×10^7colonies/ml) was mixed with the overnight AH1109-PCBP culture and incubated at 30°C overnight. The mating mixture was then plated on 60-70 large (150-mm) nutrition selection plates (explaining in detail in "Library screening" of the Result section). Plates were incubated at 30°C until colonies appeared.

Library Titering

The pretransformed pACT 2 in Y187 mouse brain library shown in Figure 3 (Clontech) grew YPD medium and was incubated at 30°C overnight. The overnight culture was centrifuged at 1000 x g for 10 minutes and the cell pellet was resuspended in 0.5 X YPDA. One hundred microliter aliquots of 1:10, 1:100, 1:1000 and 1:10000 dilutions of the mixture was spread on SD/-Leucine plates. Plates were incubated at 30°C for 3-5 days until colonies appeared. Plates containing 30 – 300 colonies were counted. The following equation was used to calculate the titer:

$$\frac{\textit{Number of colonies}}{\textit{Plating volume x dilution factor}} = \text{colonies/ml}$$

Isolation of plasmid DNA from yeast

The yeast colony grew in growth medium at 30°C overnight. The cells were centrifuged at 14,000 rpm for 5 minutes and were resuspended in potassium phosphate (67mM KH_2PO_4; pH 7.5). The cells were lysed by adding Lyticase solution (4μg/mL) (Sigma) and SDS. The plasmid was purified using Yeastmaker Yeast Plasmid Isolation Kit purchased form Clontech (Palo Alto, CA) according to manufacture's instructions. Briefly, the plasmid was purified through the CHROMA SPIN-1000 DEPC-H_2O Column and then the purified DNA was used for bacteria transformation (details for "bacteria transformation" was described in the Method and Material section).

Results

Construction of bait plasmid and expression of bait protein in yeast cells

Several studies suggested that poly C binding protein 1 (PCBP) is involved in many cellular processes; however the mechanisms of its action are largely unexplored (Makeyev et al, 2002). Thus the identification of a protein partner(s) of PCBP would provide some insight information of its versatile physiological functions. The yeast two-hybrid system was used due to its ability to detect an unknown protein from a cDNA library via *in vivo* protein-protein interactions by the expression of a functional PCBP in cells.

First, it is necessary to construct the bait plasmid (pGBKT7-PCBP) by inserting the segment of PCBP cDNA into the multiple cloning site (MCS) of the yeast expression vector, the pGBKT7 plasmid, which also contained the kanamycin resistant gene (Fig.1A). The ligation of PCBP cDNA and pGBKT7 plasmid resulted in the PCBP fused to amino acids 1-147 of the GAL4 binding domain (GAL4 BD), which created a DNA Binding Domain (DNA-BD) PCBP fusion protein. This ligation mixture was then transformed into *E. coli* cells, and the transformants underwent kanamycin (25 μl/ml) selection.

The plasmids from the survived clones were extracted and subjected to restriction enzyme Nco I and Pst I digestion to verify the insertion of PCBP cDNA into the expression vector. The restriction digestion mixtures were then further analyzed using electrophoresis with ethidium bromide staining. As shown in Fig. 1B, the correct pGBKT7-PCBP plasmid contained the 0.6 and 0.9 Kb bands (Lanes 1-5). Lane 6

contained the DNA markers (Fig. 1B). Results of gel electrophoresis analysis of restriction digestion reaction confirmed the successful construction of PCBP into the bait vector (pGBKT7-PCBP).

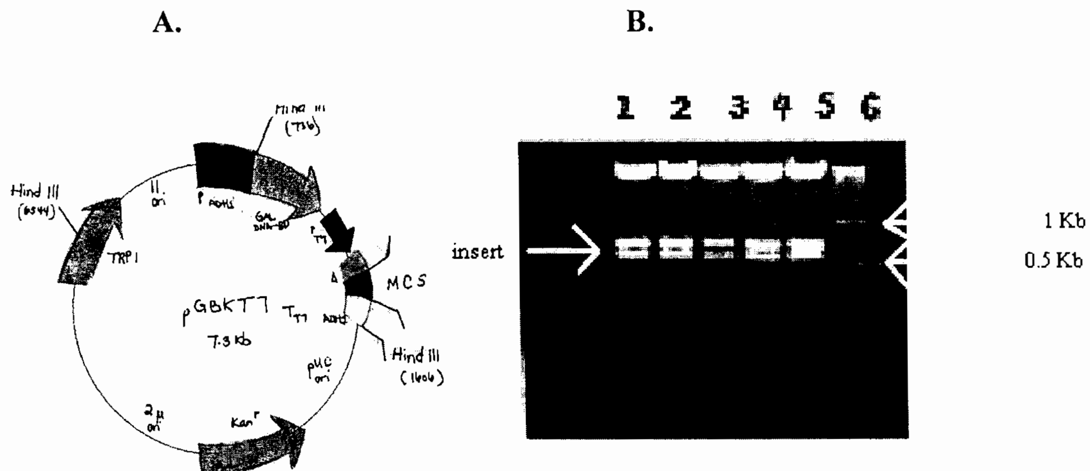


Figure 1: Construction of pGBKT7-PCBP, a bait plasmid. A. A diagram of a yeast expression vector, pGBKT7 plasmid. The pGBKT7 vector was utilized to construct a fusion protein of GAL4 DNA binding domain (GAL4 DNA-BD)-PCBP bait protein. **B. Electrophoretic analysis of the restriction digestion mixture of pGBKT7-PCBP containing clones.** Mouse PCBP cDNA fragment was ligated into the multiple cloning site (MCS) of pGBKT7 plasmid. The ligation mixture was transformed into *E. coli* cells, which were then selected using kanamycin (25 μl/ml) containing plates. The plasmids from the surviving clones were extracted and subjected to restriction enzyme Nco I and Pst I digestion, and then further analyzed using electrophoresis with ethidium bromide staining. The correct pGBKT7-PCBP plasmid contained the 0.6 and 0.9 Kb bands (Lanes

1-5) which is indicated by the arrow. Lane 6 contained DNA markers with arrows indicateing the 0.5 and 1 Kb markers .

Integration of bait plasmid into yeast genome

The pGBKT7-PCBP plasmid was then subjected to Xho I restriction enzyme digestion to generate a linearized plasmid, which was then used to transform the host yeast cell, AH109, by lithium acetate (LiAc) method as described in Method and Material section. In this process, the outer-membrane of the yeast cell AH109 is disturbed by the chemical addition of LiAc. This disruption produces holes in the membrane that are large enough for the pGBKT7-PCBP plasmid to enter. The pGBKT7-PCBP can cross the membrane of AH109 and then is integrating into yeast genome.

In addition, the pGBKT7-PCBP yeast expression plasmid also contained a selective gene encoding a protein contributing to a cell's survival under the tryptophan deffiecent nutritional selection. The pGBKT7 plasmid contains *TRP1*, the reporter gene for the production of tryptophan, which is a useful nutritional marker for selection in yeast as shown in Figure 1. Therefore, the yeasts containing the bait plasmid (pGBKT7-PCBP) will be able to produce tryptophan and to survive in media in lacking tryptophan.

Therefore, to verify integration of the bait plasmid into the yeast cells, after LiAc transformation, the transformed cells were plated onto plates containing essential nutrition except tryptophan for the selection. As shown in Fig. 2, several visible colonies were observed, indicating that pGBKT7-PCBP was integrated into the genome of these transformed AH109 yeast cells. The largest colony was then selected for the library screening use.

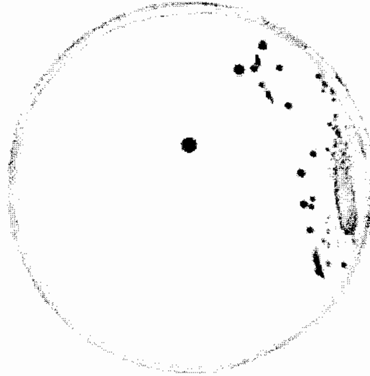


Figure 2: Integration of pGBKT7-PCBP bait plasmid into the AH109 yeast genome. The pGBKT7-PCBP plasmid, containing the tryptophan reporter gene, was linearized using Xho I restriction enzyme. The linearized DNA was then transformed into yeast cell AH109 using Lithium acetate (LiAc)-mediated transformation method. The transformed cells were then plated onto the plates containing tryptophan dropout selection medium to select for transformants containing the introduced PCBP. Visible colonies indicated that the pGBKT7-PCBP was able to integrate into the genome of AH109 yeast cells. The largest colony was then selected for the library screening use.

The setup of the positive and negative controls prior to the library screening

Before screening a cDNA library, the positive control was set up by utilizing the known strong protein-protein interaction present between p53 and large T-antigen. Two yeast strains, that contained pGBKT7-53 (p53) or pTD1-1 individually, were used as a positive control. The pTD1-1 vector, encodes full-length, wild-type GAL4 DNA-AD fused with SV40 large T-antigen. The pGBKT7-53 vector contained the fusion protein of the GAL4 DNA-BD and murine p53. Two other yeast strains, which contained the pGBKT7-Lam (DNA-BD) and pTD1-1 (AD) fusion plasmids individually, were used as

the two-hybrid negative control. The pGBKT7-Lam vector encodes a fusion of the DNA-BD with human lamin C, which served as a negative control, because Lamin C neither forms complexes nor interacts with SV40 large T-antigen.

Since p53 are known to interact with the large T-antigen, mating AH109 transformed with p53 to Y187 transformed with pTD1-1 will result in diploid cells containing both plasmids, therefore this mating was used as a positive control to verify the experimental settings prior to the cDNA library screening. Several colonies surviving under the minimal nutrition selection were obtained (data not shown). In contrast, p53 and pGBKT7-Lam do not interact with each other; mating AH109 transformed with p53 to Y187 transformed with pGBKT7-Lam will not result in diploid cells containing both plasmids, which therefore was used as a negative control. No colonies were obtained under the same nutrition selection (data not shown).

Library Screening

In the yeast two-hybrid assay, two types of fusion proteins are created: (1) the protein of interest (bait), which is fused to a DNA binding domain (DNA-BD). In this case, PCBP is the protein of interest and the pGBKT7-PCBP bait plasmid was constructed; (2) the PCBP potential interacting partner from a cDNA library, which is fused to a DNA activation domain (DNA-AD).

A pretransformed cDNA library of mouse brain was used. The cDNA library contained cDNAs of various proteins, which were fused to the GAL4 activating domain (GAL4 AD, amino acids 768-881) in the pACT2 plasmid, a yeast expression vector (Fig.3A). The pACT2 yeast vector contains the selective gene of *LEU2,* which allows

cells to grow on the media lacking leucine. The mouse cDNA library was transformed into the Y187 yeast cells.

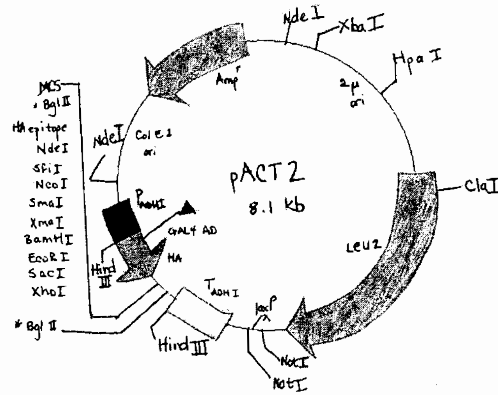
In order for the yeast strains to mate, one strain must be MATa and the other MATα. The Y187 yeast strain is MATα strain, and the AH109 yeast strain is a MATa strain. The AH109 yeast strain is therefore used as the mating partner of Y187.

The AH109 cells containing bait plasmid pGBKT7-PCBP is then used for the library screening. This transformed AH109 and Y187 cells containing cDNA library were allowed to mate in media lacking leucine, tryptophan and histidine to form diploid cells. The *TRP1* (tryptophan) and *LEU2* (leucine) genes were introduced via the bait and cDNA plasmid individually, and the AH109 cell also contains the endogenous *HIS* (histidine) reporter gene which is driven by the GAL4-responsive promoter, and could be activated by the interaction of GAL4-AD and BD. The mated yeast cells were then plated on selection media which lacked the essential nutrients tryptophan, leucine and histidine. The total numbers of 2.2×10^7 colonies from a mouse brain cDNA library were screened.

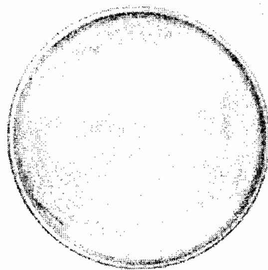
Nutritional selections occurred in three stages. The first stage is the medium stringency selection, which took place on minimal media lacking tryptophan, leucine and histidine. In total, 60-70 plates were screened under the medium stringency where several plates contained the surviving colonies. One example, as shown in Fig. 3B right panel, illustrates that several colonies survived on the minimal media plates lacking tryptophan, leucine and histidine (triple dropout selection plate). In addition, many plates containing no colonies were also observed, one of which was shown in Fig. 3B left panel. Colonies that survived from this selection underwent a high stringent selection, a second stage selection. This second stage selection occurred on minimal media plate lacking

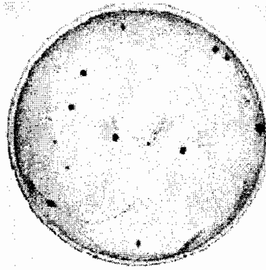
tryptophan, leucine, histidine and adenine (quadruple dropout selection plate). When the first stage survived colonies were subjected to these high stringent selection plates, as shown in Fig. 3C, some cells survived under this selection, whereas some cells did not. These colonies that survived the stringent selection underwent a higher selection called X-galactosidase color selection. This final stage of selection occurs on a quadruple selection media plate with the addition of X-α-galactosidase selection. The colonies that survived this selection produced a blue color. As shown in Fig. 3D, there were three colonies that passed the X-α-galactosidase screening, which showed the blue colored colonies.

A.



B. **C.** **D.**





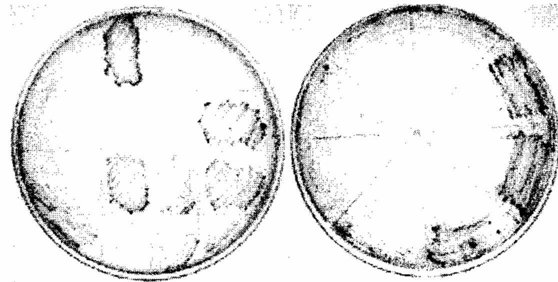


Figure 3: A. A diagram of a yeast expression vector, pACT2 plasmid, for the construction of mouse brain cDNA library. The pACT2 plasmid was used to insert various cDNA fragments, encoding various proteins expressed in the mouse brain, to generate GAL4-AD fusion proteins. This mouse cDNA library was then transformed into Y187 yeast cells. **B. Screening the PCBP-interacting protein using mouse brain cDNA library.** The AH109-PCBP cells were mated with Y187 cells containing the library cDNA, and the mating mixture, containing diploid cells with both plasmids expressed, was then plated on 60-70 plates containing histidine, leucine and tryptophan dropout medium plates for medium stringency selection. The plate on the left was an example of library screening, which contained no surviving clones, whereas the plate on

the right contained several survived colonies. The mouse brain cDNA library contained 2.2×10^7 colonies. **C. Screening candidate clones using the high stringency selection.** The colonies survived under the medium-stringency selection were further subjected to high stringency selection. This plate was an example indicating that not all colonies survived under medium stringent selection would survive under the high stringency selection. **D. Screening candidate clones using X-α-galactosidase screening.** The surviving colonies were further subjected to the X-α-galactosidase screening. Three colonies indicated survival of screening by producing a blue color.

Purification of cDNA plasmid from the candidate clones

Three colonies (named HY87, HY94 and HY18) that passed the X-α-galactosidase screening were subjected to further examination. In order to determine the sizes of cDNA fragments in the pACT2-cDNA plasmids, the plasmids were extracted from the yeast cells using a combination of lyticase and SDS. Theoretically, the plasmid DNA isolated from each positive yeast colony should contain one type of AD/library plasmid.

To our surprise, only the plasmid from HY87 yeast cells was successfully isolated In order to obtain a larger quantity of this particular plasmid, HY87 plasmid was then transformed into *E. coli* and plated onto LB medium containing ampicillin. The plasmids from the survived *E. coli* colonies were extracted, subjected to Bgl II digestion, and then further analyzed using electrophoresis with ethidium bromide staining. As shown in fig. 4, the sizes of the cDNA inserts (lanes 1-2 and 4-6; as indicated by the arrow), approximately 1.1Kb, were the same among different *E. coli* colonies. These results

suggested that one type of plasmid resided in the HY87 cells, and its size was approximately 1.1 Kb.

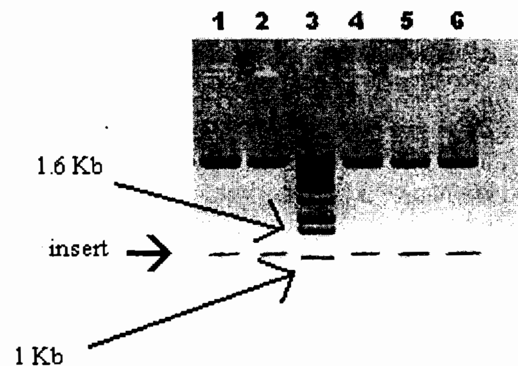


Figure 4: Electrophoretic analyses of the size of cDNA insert of pACT2 plasmid from the HY87 clone. The plasmid from the HY87 yeast clone was extracted and then transformed into *E. coli*. The transformats were then plated onto LB medium containing ampicillin. The plasmids from survivingl *E. coli* colonies were extracted and digested with Bgl II, which were then, further analyzed using electrophoresis with ethidium bromide staining. The size of this insert was approximately 1.1Kb (indicated by the arrow). The middle lane (lane 3) contained the DNA markers with arrows indicating the 1 and 1.6 Kb markers.

The cDNA sequence of the HY87 Clone

Sequences of the cDNA insert of pACT2 plasmid from HY87 clone were determined. This cDNA insert sequence, approximately 1.1 Kb in size, was then submitted to NCBI database for the sequence comparison. To our surprise, this sequence

matched the partial sequences of the mouse chromosome 5 published in NCBI database. This result therefore suggested this cDNA insert was located in the mouse chromosome 5.

However, the database search also indicated that this cDNA fragment shares no significant homology to any of known proteins. Therefore, this unknown DNA sequence was further subjected to the open reading frames (ORF) program to search for the possible ORFs found in this cDNA fragment, which will help to reveal what protein this cDNA encoded. The search outcomes were listed in Figure 5. Totally, 4 different amino acid sequences were obtained from the coding strand. In addition, these amino acid sequences were subjected to ProtParam program to compute the physical and chemical parameters of each sequence (Table 1).

As shown in Fig. 5, sequence 1 contains 60 amino acids making this the largest protein among all predicted proteins using the ORF program. It is localized from nucleotide 682 to nucleotide 864 of the sequenced DNA region. This protein has a molecular weight of approximately 6600 and a calculated theoretical pI of 5.97 as listed in Table 1. A major component of the amino acid composition is leucine and serine, which is approximately 18.3 %. The amino acid sequence was further subjected to Protein/Protein databases, and no significant similarities to other proteins were found.

The 2nd predicted protein, named sequence 2 (Fig. 5), is composed of 50 amino acids, in which a major component (12%) of the amino acid composition is glycine and alanine, respectfully. It is located from nucleotide 19 to nucleotide 171. Due to an undetermined nucleotide in the DNA sequence, there is an undetermined amino acid in sequence 2. As listed in Table 1, this sequence has a molecular weight of approximately 5479 and a calculated theoretical pI of 9.60. In addition, this sequence was in frame with

the activation domain of the pACT2 vector. The sequence 2 was then further subjected to Blast Protein/Protein databases and this sequence was found to have a low homology (46%) to both the arginine biosynthesis bifunctional protein and the cell division ftsY homolog, which is illustrated in Fig. 5B and Fig. 5C. Amino acids 4 – 32 in sequence 2 were homologous to amino acids 10 to 39 in the arginine biosynthesis bifunctional protein. Also, amino acids 15 – 50 in sequence 2 were homologous to amino acids 161 to 199 in the cell division ftsY homolog.

The 3rd predicted protein, sequence 3, is composed of 39 amino acids and leucine makes up a major component of the amino acid composition by 15.4%. Within sequence 3, there are 9 ambiguous positions where 9 amino acids have no associated codons again, due to a DNA sequencing ambiguity. It is localized from nucleotide 970 to nucleotide 1086 of the sequenced DNA region. Sequence 3 has a molecular weight of approximately 4447 and a calculated theoretical pI of 10.02, which is illustrated by Table 1. The amino acid sequence was also subjected to the Protein/Protein databases search and no significant sequence similarities to other proteins were found.

The last predicted protein is named sequence 4, which is composed of 35 amino acids making this the smallest one among all predicted proteins. One amino acid with no associated codon is found due to a DNA sequencing ambiguity. A major component of the amino acid composition for sequence 4 is alanine and serine, which is approximately 11.4 % respectfully. It is located from nucleotide 35 to nucleotide 142 of the cDNA fragment. The amino acid sequence has a molecular weight of approximately 4056 and a calculated theoretical pI of 8.34 as shown in Table 1. No significant similarities to other known proteins were found using the Blast Protein/Protein databases.

matched the partial sequences of the mouse chromosome 5 published in NCBI database. This result therefore suggested this cDNA insert was located in the mouse chromosome 5.

However, the database search also indicated that this cDNA fragment shares no significant homology to any of known proteins. Therefore, this unknown DNA sequence was further subjected to the open reading frames (ORF) program to search for the possible ORFs found in this cDNA fragment, which will help to reveal what protein this cDNA encoded. The search outcomes were listed in Figure 5. Totally, 4 different amino acid sequences were obtained from the coding strand. In addition, these amino acid sequences were subjected to ProtParam program to compute the physical and chemical parameters of each sequence (Table 1).

As shown in Fig. 5, sequence 1 contains 60 amino acids making this the largest protein among all predicted proteins using the ORF program. It is localized from nucleotide 682 to nucleotide 864 of the sequenced DNA region. This protein has a molecular weight of approximately 6600 and a calculated theoretical pI of 5.97 as listed in Table 1. A major component of the amino acid composition is leucine and serine, which is approximately 18.3 %. The amino acid sequence was further subjected to Protein/Protein databases, and no significant similarities to other proteins were found.

The 2nd predicted protein, named sequence 2 (Fig. 5), is composed of 50 amino acids, in which a major component (12%) of the amino acid composition is glycine and alanine, respectfully. It is located from nucleotide 19 to nucleotide 171. Due to an undetermined nucleotide in the DNA sequence, there is an undetermined amino acid in sequence 2. As listed in Table 1, this sequence has a molecular weight of approximately 5479 and a calculated theoretical pI of 9.60. In addition, this sequence was in frame with

the activation domain of the pACT2 vector. The sequence 2 was then further subjected to Blast Protein/Protein databases and this sequence was found to have a low homology (46%) to both the arginine biosynthesis bifunctional protein and the cell division ftsY homolog, which is illustrated in Fig. 5B and Fig. 5C. Amino acids 4 – 32 in sequence 2 were homologous to amino acids 10 to 39 in the arginine biosynthesis bifunctional protein. Also, amino acids 15 – 50 in sequence 2 were homologous to amino acids 161 to 199 in the cell division ftsY homolog.

The 3rd predicted protein, sequence 3, is composed of 39 amino acids and leucine makes up a major component of the amino acid composition by 15.4%. Within sequence 3, there are 9 ambiguous positions where 9 amino acids have no associated codons again, due to a DNA sequencing ambiguity. It is localized from nucleotide 970 to nucleotide 1086 of the sequenced DNA region. Sequence 3 has a molecular weight of approximately 4447 and a calculated theoretical pI of 10.02, which is illustrated by Table 1. The amino acid sequence was also subjected to the Protein/Protein databases search and no significant sequence similarities to other proteins were found.

The last predicted protein is named sequence 4, which is composed of 35 amino acids making this the smallest one among all predicted proteins. One amino acid with no associated codon is found due to a DNA sequencing ambiguity. A major component of the amino acid composition for sequence 4 is alanine and serine, which is approximately 11.4 % respectfully. It is located from nucleotide 35 to nucleotide 142 of the cDNA fragment. The amino acid sequence has a molecular weight of approximately 4056 and a calculated theoretical pI of 8.34 as shown in Table 1. No significant similarities to other known proteins were found using the Blast Protein/Protein databases.

Figure 5: The predicted open reading frames of cDNA of HY87

A.

```
     682
   tatatcttttcaagtaagaaagtgaagctcactctctccccc
     Y  I  F  S  S  K  K  V  K  L  T  L  S  P 727
ctctccgctccttccccattctctctctcccttcccccc    tct
 L  S  A  P  S  P  F  S  L  S  L  P  P     S 772
ctctcc   ttc   tgtttatggtgt   tttgaatgtgttgaa
 L  S     F     C  L  W  C     F  E  C  V  E 817
gagggttcatgtacatgtgtacaggcttacaatgggagtcttctt
 E  G  S  C  T  C  V  Q  A  Y  N  G  S  L  L 862
     864
```

Sequence 1

```
    19
   gcttacccatacgatgttccanattacgctagcttgggtggt
     A  Y  P  Y  D  V  P  X  Y  A  S  L  G  G 64
cat   gcc   gaggccccggggatccgaattcgcggccgcgtc
 H     A     E  A  P  G  I  R  I  R  G  R  V 10
gacgtcaaaggaacagtcaacaaaaacagattagggccagctctt
 D  V  K  G  T  V  N  K  N  R  L  G  P  A  L 15
taccttgcactctac      171
Y  L  A  L  Y
```

Sequence 2

```
     970
    gacactggtccttggtggggaaat    tanntgcnnaaatgt
     D  T  G  P  W  W  G  N     X  X  X  K  C 1015
cttcncngncttganctggacgtgttgganannaagcattacan
 L  X  X  L  X  L  G  R  V  G  X  X  A  L  X 1060
nga   ccctattttaaaaaaaaaaaa 1086
 X     P  Y  F  K  K  K  K
```

Sequence 3

```
     35
   ttccanattacgctagcttgggtggtcatatggccatggagg
     P  X  I  T  L  A  W  V  V  I  W  P  W  R 80
ccccggggatccgaattcgcggccgcgtcgacgtcaaaggaacag
 P  R  G  S  E  F  A  A  A  S  T  S  K  E  Q 125
tcaacaaaaacagat      142
 S  T  K  T  D
```

Sequence 4

B.

Color key for alignment scores

<40	40-50	50-80	80-200	>=200

Query

0 10 20 30 40 50

arginine biosynthesis bifunctional protein

cell division ftsY homolog

C.

```
Query   4    PYDVPXYASLGG-HMAMEAPGIRIRGRVDV   32
Sbjct   10   PTDVPAMPMIAGVRLATAAAGIRYKGRTDV   39
```

Arginine biosyntheis bifunctional protein argJ

```
Query   15   GHMAMEAPGIR---IRGRVDVKGTVNKNRLGPALYLALY   50
Sbjct   161  GDMGMEATEYLMKDVRGRVSLKGLKDGNELRGALKEALY   199
```

Cell division protein ftsY homolog

Figure 5: The predicted open reading frames of cDNA of HY87.

A. The cDNA sequence of HY87 was subjected to the open reading frame program from NCBI. Four predicted amino acid sequences (listed as sequences 1-4) were obtained from the search. **B. Alignment of amino acid sequence homologous region between sequence 2 and arginine biosynthesis bifunctional protein or the cell division ftsY homolog.** Using the domain search via NCBI database, the sequence 2 was found to have partial homology to the arginine biosynthesis bifunctional protein or the cell division ftsY homolog. A black line indicated the homologous region, respectively. **C. Amino acid sequence comparison of the sequence 2 and the arginine biosynthesis bifunctional protein and the cell division protein ftsY homolog protein.** The amino acids 4 – 32 of sequence 2 were homologous to amino acids 10 to 39 of the arginine biosynthesis bifunctional protein argJ. The amino acids 15 – 50 of sequence 2 were also homologous to amino acids 161 to 199 in the cell division ftsY homolog. Both sequences share approximately 46% homology to sequence 2.

Sequence 1			Sequence 2			Sequence 3			Sequence 4		
Number of amino acids: 60			**Number of amino acids:** 50			**Number of amino acids:** 39			**Number of amino acids:** 35		
Molecular weight: 6599.8			**Molecular weight:** 5478.6			**Molecular weight:** 4447.4			**Molecular weight:** 4055.8		
Theoretical pI: 5.97			**Theoretical pI:** 9.60			**Theoretical pI:** 10.02			**Theoretical pI:** 8.34		
Amino acid Composition:			**Amino acid composition:**			**Amino acid composition:**			**Amino acid composition:**		
Ala (A)	2	3.3%	Ala (A)	6	12.0%	Ala (A)	1	2.6%	Ala (A)	4	11.4%
Arg (R)	0	0.0%	Arg (R)	4	8.0%	Arg (R)	1	2.6%	Arg (R)	2	5.7%
Asn (N)	1	1.7%	Asn (N)	2	4.0%	Asn (N)	1	2.6%	Asn (N)	0	0.0%
Asp (D)	0	0.0%	Asp (D)	2	4.0%	Asp (D)	1	2.6%	Asp (D)	1	2.9%
Cys (C)	5	8.3%	Cys (C)	0	0.0%	Cys (C)	1	2.6%	Cys (C)	0	0.0%
Gln (Q)	1	1.7%	Gln (Q)	0	0.0%	Gln (Q)	0	0.0%	Gln (Q)	1	2.9%
Glu (E)	3	5.0%	Glu (E)	1	2.0%	Glu (E)	0	0.0%	Glu (E)	2	5.7%
Gly (G)	2	3.3%	Gly (G)	6	12.0%	Gly (G)	4	10.3%	Gly (G)	1	2.9%
His (H)	0	0.0%	His (H)	1	2.0%	His (H)	0	0.0%	His (H)	0	0.0%
Ile (I)	1	1.7%	Ile (I)	2	5.7%	Ile (I)	0	0.0%	Ile (I)	2	5.7%
Leu (L)	11	18.3%	Leu (L)	5	10.0%	Leu (L)	6	15.4%	Leu (L)	1	2.9%
Lys (K)	3	5.0%	Lys (K)	2	4.0%	Lys (K)	5	12.8%	Lys (K)	2	5.7%
Met (M)	3	5.0%	Met (M)	3	6.0%	Met (M)	1	2.6%	Met (M)	1	2.9%
Phe (F)	4	6.7%	Phe (F)	0	0.0%	Phe (F)	1	2.6%	Phe (F)	3	8.6%
Pro (P)	5	8.3%	Pro (P)	4	8.0%	Pro (P)	2	5.1%	Pro (P)	2	5.7%
Ser (S)	11	18.3%	Ser (S)	1	2.0%	Ser (S)	0	0.0%	Ser (S)	4	11.4%
Thr (T)	2	3.3%	Thr (T)	1	2.0%	Thr (T)	1	2.6%	Thr (T)	3	8.6%
Trp (W)	1	1.7%	Trp (W)	0	0.0%	Trp (W)	2	5.1%	Trp (W)	3	8.6%
Tyr (Y)	2	3.3%	Tyr (Y)	5	10.0%	Tyr (Y)	1	2.6%	Tyr (Y)	0	0.0%
Val (V)	3	5.0%	Val (V)	4	8.0%	Val (V)	1	2.6%	Val (V)	2	5.7%
Asx (B)	0	0.0%	Asx (B)	0	0.0%	Asx (B)	0	0.0%	Asx (B)	0	0.0%
Glx (Z)	0	0.0%	Glx (Z)	0	0.0%	Glx (Z)	0	0.0%	Glx (Z)	0	0.0%
Xaa (X)	0	0.0%	Xaa (X)	1	2.0%	Xaa (X)	10	25.6%	Xaa (X)	1	2.9%

Table 1: The physical and chemical parameters of Sequences 1-4.

Four predicted amino acid sequences from the open reading frame program of NCBI (sequences 1- 4) were further subjected to ProtParam to compute the physical and chemical parameters of each predicted protein. This calculated molecular weight, theoretical pI, and amino acid composition for each sequence is listed. “Xaa” represented the amino acids with no associated codons due to DNA sequencing ambiguity.

Discussion

The poly C binding protein 1 (PCBP) belongs to the KH domain family. The KH domain has been identified in a wide spectrum of RNA-binding proteins (Gibson et al., 1993), and proteins containing KH domains have been shown to be involved in multiple important regulatory mechanisms (Funke et al., 1996). Recent studies further showed that PCBP is also involved in the transcriptional regulation of MOR gene expression (Ko and Loh, 2005). In this study, we therefore explored the physiological role of the multi-function PCBP by investigating its potential protein using the yeast two hybrid screening method.

With the success of construction of the bait plasmid, integration of bait plasmid into the host yeasts and the screening of a mouse brain cDNA library, three candidate colonies (HY87, HY94 and HY18) that were selected using the high stringent condition were obtained. From the three colonies that survived the X-α-galactosidase screening, only the cDNA plasmid from the HY87 yeast colony was successfully isolated. No cDNA plasmids were obtained from the other two yeast colonies, indicating that they may be false positives. Using a NCBI database search, the DNA sequence from the pACT2 plasmid of HY87 clone was found to match partial sequences of the mouse chromosome 5, suggesting this cDNA located in mouse chromosome 5.

The database search further showed that this cDNA fragment shares no significant homology to any known proteins, signifying that this cDNA may possible encodes an unknown protein. Therefore, to investigate if this cDNA encoded any proteins, the DNA sequence was subjected to the open reading frame program. Four predicted amino acid sequences were obtained from the sense strand. Among the predicted sequences,

Sequence 2 is in frame with the AD domain in pACT2, indicating that this protein is most likely to be expressed in the yeast cells. In addition, Sequence 2 shared a partial homology (46% similarity at the amino acid level) to both arginine biosynthesis bifunctional protein Arg J (Arg J protein) and cell division protein ftsY homolog (CDPF homolog). The amino acids 4 – 32 of Sequence 2 were homologous to amino acids 10 to 39 of the Arg J protein. The amino acids 15 – 50 of Sequence 2 were also homologous to amino acids 161 to 199 in the CDPF homolog.

The arginine biosynthesis bifunctional protein *argJ* (ArgJ protein) belongs to the *argJ* family of proteins. It has the length of 413 amino acids and the molecular weight of 43320 Da. The arginine biosynthesis bifunctional protein *argJ* possesses *N*-acetyltransferase (EC 2.3.1.35, encoded by the *argJ* gene) commonly referred to as ornithine acetyltransferase (OATase). OATase catalyses the transfer of the acetyl-group from *N*-acetylornithine to L- glutamate in arginine biosynthesis in microorganisms (Marc et al., 2000). Also involved in the arginine biosynthesis is acetylglutamate synthase (AGSase, EC 2.3.1.1) which initiates this pathway by acetylating glutamate (Marc et al., 2000). The enzyme catalyzed the first and fifth steps involved in the cyclic version of arginine biosynthesis (Sato et al., 1996); therefore the ArgJ protein is classified as a bifunctional enzyme (Sakanyan et al., 1993).

No acetylated derivatives pathway specialized in ornithine biosynthesis have been reported in mammalian. Some ornithine is made from glutamyl-phosphate semialdehyde (an intermediate in the pathway to proline), but most of the arginine produced is hydrolyzed in the urea cycle, which ensures detoxification of excess ammonium. Hence, arginine is an essential amino acid obtained through food. Although

no acetylated derivatives cycle has been identified, AGSase is present in mammals. It is viewed as the first committed step in the urea cycle because carbamoyl-phosphate synthase I, forming the carbamoyl phosphate needed together with ornithine to produce citrulline, must be activated specifically by acetylglutamate to function (Grisolia and Cohen, 1953). Thus, factors regulating the activity of the AGSase are relevant to the control of the urea cycle. It has also been shown that arginine stimulates the activity of rat AGSase (Shigesada and Tatibana, 1978). AGSase transcripts could be detected in liver, small intestine and kidney tissue (Häberle et al., 2003) and its activity in mammals is modulated by protein intake and is directly dependent on the L-arginine concentration (Shigesada and Tatibana, 1978). The homology of Sequence 2 to the ArgJ protein suggested that Sequence 2 may be related to the mammal AGSase. However, this needs to be further investigated.

The Sequence 2 also shares partial homology to the cell division protein ftsY homolog (CDPF homolog), which belongs to the GTP-binding signal recognition particle (SRP) family. It is 416 amino acids in length and has the molecular weight of 44360 Da. The SRP is found in the cytoplasm of eukaryotes and prokaryotes, where it plays a central role in the co-translational targeting of ribosomes; translating and exporting membrane proteins to the membrane of the endoplasmic reticulum and the plasma membrane, respectively (Doudna and Bately, 2004). The prokaryotic SRP pathway consists of the following two proteins: Ffh and FtsY, and a 4.5S RNA. FtsY is the membrane associated SRP receptor or docking protein and it has three domains. The highly conserved N-domain is involved in the protein-protein interaction, which directs protein trafficking (Freymann et al., 1999). The G-domain of FtsY mediates the GTPase

activities, which are required for their function in protein targeting (Powers and Walter, 1997). The negatively charged amino-terminal A- domain of FtsY is important for membrane localization (Millman and Andrews, 1999, Powers and Walters, 1997, Zelazny et al., 1997) and many also be involved in modulating the GTPase activity in response to the interaction with the cytoplasmic membrane (deLeeuw et al., 2000). Collectively, the homology of Sequence 2 to the N-domain suggests that Sequence 2 may be involved in the protein-protein interaction and protein trafficking.

Future experiments will need to be carried out to examine the cloned protein and its possible functional roles. First, the expression of Sequence 2 needs to be confirmed using Western blot analysis. Secondly, the protein interaction between PCBP and Sequence 2 protein needs to be confirmed via co-immunoprecipitation. The protein encoded by Sequence 2 should be co-precipitated by anti-PCBP antibody. Thirdly, Sequence 2 may be a partial sequence of an entire clone, due to a short predicted peptide sequence. One possible solution is to redo the library screening. In re-performing the library screening with different batch of cDNA library, an interaction between PCBP and a full length protein may be detected. Hence, it may isolate this specific interaction which may contain the complete sequence of an entire clone. Once isolated, the putative full length clone could be sequenced and entered into a protein database to identify whether any homology exists.

Conclusion

In summary, we performed a yeast two hybrid screening using a mouse brain cDNA library to identify the potential interacting protein(s) of poly C binding protein 1 (PCBP). An isolated cDNA sequence from the HY87 clone was acquired from this screening and was found to match partial sequences of the mouse chromosome 5, suggesting this clone is located at chromosome 5 and is possibly expressed in the mouse brain. However, this cDNA sequence matched no known cDNA sequence of NCBI database. This cDNA was further subjected to the open reading frame program and 4 different amino acid sequences were obtained. Among the predicted sequences, Sequence 2 was in frame with the AD domain in pACT2, and shared a partial homology (46% similarity at the amino acid level) to both arginine biosynthesis bifunctional protein arg J (Arg J protein) and cell division protein ftsY homolog (CDPF homolog). Interestingly, Sequence 2 has a homology to the N-domain localized in the CDPF homolog, implicating that Sequence 2 may be involved in protein-protein interactions and protein trafficking. Future experiments are necessary to confirm the protein expression from the cDNA clone, the interaction with PCBP, and the possible presence of a full length cDNA.

Reference

Adinolfi, S., C. Bagni, M.A. Castiglione Morelli, F. Fraternali, G. Musco, A. Pastore. 1999. 1999. Novel RNA-binding motif: The KH module. *Biopolymers* 51: 153 -164.

Baber J.L., D. Levens, D. Libutti, N. Tjandra. 2000. Chemical shift mapped DNA-binding sites and 15N relaxation analysis of the C-terminal KH domain of heterogeneous nuclear ribonucleoprotein K. *Biochemistry,* 39:6022-6032.

Buckanovich, R.J., J. B. Posner and R.B. Darnell. 1993. Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system. *Neuron*, 11: 657-672.

Bustelo, X.R., K.L. Suen, W.M. Michael, G. Dreyfuss and M. Barbacid. 1995. Association of the vav proto-oncogene product with poly (rC)- specific RNA-binding proteins. *J.Biol. Chem.* 15: 1324-1332.

deLeeuw, E., K. teKaat, C. Moser, G. Menestrina, R. Demel, B. Dekruijff, B. Oudega, J. Luirink, and I. Sinning. 2000. Anionic phospholipids are involved in membrane association of FtsY and stimulate its GTPase activity. *EMBO J* 19: 531 – 541.

Denisenko O.N. and K. Bornsztyk. 1997. The product of the murine homolog of the Drosophila extra sex combs gene displays transcritptional repressor activity. *Mol Cell Bio.,* 17:4707-4717.

Denisenko, O.N., B. O' Neill, J. Ostrowski, I. Van Seuningen, K. Bornsztyk. Ziki 1, a transcriptinal repressor that interacts withe the heterogenous nuclear ribonucleoprotein particle K protein. *J. Biol Chem.,* 1996. 271 (44): 27701 – 27706.

Doudna, J. A., and R. T. Batey. (2004). Structural insights into the signal recognition particle. *Annu. Rev. Biochem.* 73: 539-557.

Freymann, D.M., R.J. Keenan, R.M. Stoud, and P. Walter. 1999. Functional changes in the structure of the SRP GTPase on binding GDP and Mg^{2+} GDP. *Nat Struct. Biol.* 6: 793-801.

Funke B, B. Zuleger, R. Benaventer, T. Shuster, M. Goller, J. Stevenin , E. Horak. (1996). The mouse poly (C)-binding protein exists in multiple isoforms and intereacts with several RNA-binding proteins. *Nucleic Acid Research,* 24: 3821-3828.

Ghosh AK., M. Majumder, R. Steele, R.A. White, R.B. Ray. 2001. A Novel 16-Kilodalton Cellular Protein Physically Interacts with and Antagonizes the Functional Activity of c-*myc* Promoter-Binding Protein 1. *Molecular and Cellular Biology*, 21:655-662.

Gibson, T.J., J.T. Thompson, and J. Heringa. 1993. The KH domain occurs in a diverse set of RNA-binding proteins that include the antiterminator NusA and is propably involved in Binding to nucleic acid. *FEBS Lett.* 324: 361 – 366.

Grishin NV. 2001. KH domain: One motif, two folds. *Nucleic Acids Res,* 29:638-643.

Habelhah H, K. Shah, L. Huang, A. Ostareck-Lederer, A.L. Burlingame, K.M. Shokar, M.W. Hentze, Z. Ronai. 2001. ERK phosporylation drives cytoplasmic accumulation of hnRNP-K and inhibition of mRNA translation. *Nat Cell Biol,* 3:325-330.

Häberle, J., E. Schmidit, S. Pauli, J.G. Kreuder, B. Plecko,, A. Galler., B. Wermuth, E. Harms, H.G. Koch. 2003. Mutation analysis in patients with *N*-acetylglutamate synthase deficiency. *Hum. Mutat.* 21: 593-597.

Jannatipour, M., R. Dion, S. Khan, H. Jindal, J. Laganiere Fanx, A. H. Chishti, G.A. Rouleau. (2001). Schwannomin Isoform-1 Interacts with Syntenin via PDZ Domains. *J. Biol. Chem.*, 276 33093-33100.

Kim J.H, B. Hahm, Y.K. Kim, M. Choi, S.K. Jang. 2000. Protein-protein interaction amoung hnRNPs shuttling between nucleus and ccytoplasm. *J Mol Biol*, 298:395-405.

Ko, J.L., H.C. Liu, S.R. Minnerath, H.H. Loh. 1998. Transcriptional Regulation of Mouse μ-Opioid Receptor Gene. *J. Biol. Chem.*, 273: 27678-27685.

Ko, J.L. and H.H. Loh. 2001. Single-stranded DNA-binding Complex Involved in Transcriptional Regulation of Mouse μ-Opioid Receptor Gene. *J. Biol. Chem.*, 276: 788-795.

Ko, J.L., and H.H. Loh. 2005. Poly C binding protein, a single-stranded DNA-binding protein, regulates mouse μ-Opioid Receptor Gene. *J. Neurochem.*, 93 (3) 749 - 761.

Lewis H.A., H. Chen, C. Edo, R.J. Buckanovich, Y.Y. Yang, K. Musunuru, R. Zhong,, R.B. Darnell, S.K. Burley. 1999. Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding domains. *Structure Fold Des.*, 7:191 – 203.

Makeyev A.V., D.L. Eastmond and S.A. Liebhaber. (2002). Targeting a KH-domain protein with RNA decoys. *RNA*, 8: 1160-1173.

Makeyev A.V. and S.A Liebhaber. (2002). The poly(C)-binding protein: A multiplicity of functions and a search for mechanisms. *RNA*, 8: 265-278.

Marc F., P. Weigel, C. Legrain, Y. Almeras, M. Santrot, N. Glansdorff, and V. Sakanyan. 2000. Characterization and kinetic mechanism of mono- and bifuctional ornithine acetyltransferases from thermophilic microorganisms. *Eur. J. Biochem.* 267: 5217-5226.

Miau L.H., C.J. Chang, B.J. Shen, W.H. Tsai and S.C. Lee. 1998. Identification of heterogeneous nuclear riboncleoprotein K (hnRNP K) as a repressor of C/EBPbeta-mediated gene activation. *J Biol Chem*, 273: 10784-10791.

Michelotti E.F., G.A. Michelotti, A.L. Aronsohn and D. Levens. 1996. Heterogeneous nuclear ribonucleoprotien K is a transcription factor. *Mol Cell Biol,* 16:2350-2360.

Millman, J.S., and D.W.Andrews. 1999. A site-specific, membrane-dependent cleavage event defines the membrane binding domain of FtsY. *J. Biol. Chem* . 274: 33227- 33234.

Morin F., B. Vannier, F. Houdart, M. Regnacq, T. Berges. (2003). A Proline-rich domain in the Gamma Subunit of Phosphodiesterase 6 Mediates Interactions with SH3-containing Proteins. *Molecular Vision*, 9: 449-459.

Musco G, A. Kharrat, G. Stier, F. Fraternali, T.J. Gibson, M. Nilges, A. Pastore. 1997. The solution structure of the first KH domain of FMR1, the protein responsible for the fragile X syndrome. *Nat. Struct Biol,* 4:712-716.

Musco G, G. Stier, C. Joseph , M.A. Castiglione Morelli, M. Nilges, T.J. Gibson, A. Pastore. 1996. Three-dimensional structure and stability of the KH domain: Molecular insights into the fragile X syndrome. *Cell,* 85:237-245.

Powers, T., and P. Walter. 1997. Co-translational protein targeting catalyzed by the *Escherichia coli* signal recognition particle and its receptor. *EMBO J.* 16: 4880-4886.

Sakanyan, V. D., C. Charlier, A. Legrain, I. Kochikyan, A. Mett, A. Pierard, and N. Glansdorr. 1993. Primary structure, partial purification and regulation of key enzymes of the acetyl cycle of arginine biosynthesis in *Bacillus* stearothermophilus: dual function of orithine acetyltransferase. *J Gen Microbiol.* 139, 393-402.

Sato, S., H. Kotani, A. Tanaka, E. Asamizu, Y. Nakamura, N. Miyajima, M. Hirosawa, M. Sugiura, S. Sasamoto, T. Kimura, T. Hosouchi, A. Mastsuno, A. Muraki, A. Nakazaki, K. Naruo, S. Okumura , S. Shimpo, C. Takeuchi, S. Tabata. 1996. Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. Strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.*, 3:109 – 136.

Schullery D.S., J. Ostrowski, O.N. Denisenko, L. Stempka, M. Shnyreva, H. Suzuki, M. Gschwendt, K. Bomsztyk. 1999. Regulated interaction of protein kinase Cδ with the heterogeneous nuclear ribonucleoprotein K protein. *J Biol Chem* 274:15101-15109.

Shigesda, K., and M. Tatibana. 1978. Role of acetylglutamate in ureotelism. *Eur. J. Biochem.* 84: 285-291.

Shnyreva M., D.S. Schullery, H. Suzuki, Y. Higaki, K. Bomsztyk. 2000. Interaction of two multifunctional proteins. Heterogeneous nuclear ribonucleoprotein K and Y-box binding protein. *J Biol Chem,* 275: 15498-155503.

Sidiqi, M., J.A. Wilce, J.P. Vivian, C.J. Porter, A. Barker, P.J. Leedman and M.C.J. Wilce. Structure and RNA-binding of the third KH domain of poly (C) – binding protein 1. 2005. *Nucleic Acids Research.* 33 (4): 1213 – 1221.

Siomi H., M. Choi , M.C. Siomi, R.L. Nussbaum, G. Dreyfuss, 1994. Essential role for KH domains in RNA binding: Impaired RNA binding by a mutation in the KH domain of RMR1 that causes fragile X syndrome. *Cell,* 77:33-39.

Siomi H., M.J. Maturnis, W.M. Michael, G. Drefuss. 1993. The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. *Nucleic Acids Res*, 21: 1193-1198.

Tomonaga T, and D. Levens. 1996. Activating transcriptions from singlestranded DNA. *Proc Natl Acad Sci USA* 93:5830-5835.

Van Seuningen I., J. Ostrowski, X.R. Bustelo, P.R. Sleath, K. Bomsztyk. 1995. The K protein domain that recruits the interleukin 1-responsive K protein kinase lies adjacent to a cluster of c-Src and Vav SH3-binding sites. Implications that K protein acts as a docking platform. *J Biol Chem* 270:26976-26985.

Wimberly B.T., D.E. Brodersen and W.M. Clemons Jr, R.J. Morgan-Warren, A.P. Carter, C. Vonrhein, T. Hartsch, V. Ramakrishnan. 2000. Structure of the 30S ribosomal subunit. *Nature, 407:*327-339.

Zelazny A., A. Seluanov, A. Cooper, and E. Bibi,. 1997. The NG domain of the prokaryotic signal recognition particle receptor, FtsY, is fully functional when fused to an unrelated integral membrane polypeptide. *Proc. Natl. Acad. Sci. USA.* 94: 6025 - 6029.