STRUCTURAL STUDIES OF PROTEIN PARTNERSHIPS OF THE PNT DOMAIN OF GABPα

by

Scott Earl MacIntosh

B.Sc. (Honours), Dalhouse University, 1998

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Department of Biochemistry and Molecular Biology)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

September 2001

© Scott Earl MacIntosh, 2001

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Biochemistry and Molecular Biology

The University of British Columbia
Vancouver, Canada

Date September 27, 2001

DE-6 (2/88)

Abstract

Background:

The Ets family of transcription factors is an established model system for studying the regulation of gene expression. In addition to the DNA-binding ETS domain, several Ets proteins contain a second region of conserved sequence, the Pointed (PNT) domain. Although the exact functions of this domain remain to be established, there is a mounting body of evidence that the PNT domain mediates interactions with partner transcription factors. However, a detailed structural study of PNT domain interactions has yet to be carried out. This information is essential for a complete understanding of the role of the PNT domain in the control of gene expression by the Ets transcription factors.

Goals:

Previous work in the McIntosh laboratory has focused on the PNT domain of GABPα. To continue studies on this domain, three main goals of this thesis work were formulated:

i) Use of a yeast two-hybrid system to screen for protein partners of the domain.

ii) A detailed study of the domain to investigate the possibility of self-association.

iii) A detailed study of the domain to investigate the possibility of a small molecule binding to the domain.

Results and conclusions:

A yeast two-hybrid screen was used in an attempt to identify proteins that interacted with the minimal PNT domain of GABPα. One potential interacting protein, PBP, was identified. However, attempts to demonstrate a direct interaction between the proteins *in vitro* were unsuccessful.

The behaviour of a minimal GABPα PNT domain construct during purification suggested that this protein formed oligomers. However, no evidence of self-association of the protein was obtained using several different techniques. As a result, this is likely an artifact with no biological relevance.

Resonances not assignable to any amino acid residues were present in the NMR spectra of the GABPα PNT domain, raising the possibility that a small molecule was binding to the domain. Using NMR and mass spectrometry, this molecule was tentatively identified as an N-acetyl hexosamine oligomer. However, NMR relaxation experiments and further purification of the protein failed to provide any evidence that this molecule was binding the PNT domain of GABPα with appreciable affinity.

Table of Contents

List of Tables

List of Figures

Abbreviations

amu	atomic mass units
CD	circular dichroism
cDNA	complementary deoxyribonucleic acid
D_2O	deuterium oxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
ERK	extracellular signal-regulated kinase
ESI-MS	electrospray ionization mass spectrometry
FPLC	fast protein liquid chromatography
GABP	GA binding protein
GST	glutathione-S-transferase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HSQC	heteronuclear shift quantum correlation
IPTG	isopropyl β-D-thioglucopyranoside
JNK	c-Jun N-terminal kinase
LB	Luria-Bertani
MAPK	mitogen-activated protein kinase
mRNA	messenger ribonucleic acid
NMR	nuclear magnetic resonance
NOE	nuclear overhauser enhancement

NTA	nitrilotriacetic acid
OAc	acetate
OD_{600}	optical density at 600 nm
ONPG	*o*-nitrophenyl-β-D-galactopyranoside
PBP	Pointed Binding Protein
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
PMSF	phenylmethylsulfonyl fluoride
PNT	Pointed
RP-HPLC	reversed-phase high performance liquid chromatography
SAM	sterile alpha motif
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SS-DNA	single-stranded deoxyribonucleic acid
TBS	Tris buffered saline
UV	ultraviolet

Acknowledgments

I would like to acknowledge the assistance of a number of people who helped make this thesis work possible. I would like to thank my research supervisor, Dr. Lawrence McIntosh, for his guidance and thesis committee members Dr. Ross MacGillivray and Dr. Natalie Strynadka for their helpful advice. I would also like to thank previous McIntosh laboratory workers Isabelle Pot and Lisa Gentile. This thesis is based on their preliminary work. I am also indebted to Shouming He, who performed the mass spectrometry analysis. Finally, I thank all members of the McIntosh laboratory past and present.

CHAPTER 1

General Introduction

1.1 Ets Proteins

The Ets family of transcription factors and oncoproteins is an established model system for studying the regulation of gene expression. The founding member of this family, Ets-1, was discovered in the early 1980's as part of the tripartite oncogene of avian retrovirus E26 (Leprince *et al.*, 1983; Nunn *et al.*, 1983). Since then, many proteins with sequence homology to Ets-1 have been discovered. They have been used as model systems to study how transcription factors bind DNA, modulate transcriptional activity at promoters and respond to signaling pathways. These proteins are present in all metazoan phyla and are involved in a variety of cellular processes including developmental pathways, oncogenesis and viral gene expression (Graves & Petersen, 1998).

The family is characterized by the highly conserved DNA-binding ETS domain, an ~85-residue domain that recognizes a core 5'-GGA-3' site. The structures of the ETS domains from several Ets proteins have been solved to date as summarized in Table 1.1. The NMR-derived structure of the ETS domain from murine Ets-1 (Donaldson *et al.*, 1996) is shown in Figure 1.1.

These structural studies reveal that all ETS domains have a similar topology consisting of four anti-parallel β-strands that pack against three α-helices to form a winged helix-turn-helix motif. DNA binding activity to sequences with a core 5'-GGA-3' motif is very similar within the family and ETS domain residues that contact the DNA are highly conserved. Given that many Ets proteins are known to display distinct

Source	Ets Protein	Method	Reference	PDB ID
Human	Elk-1	X-ray Diffraction	(Mo *et al.*, 2000)	1DUX
Human	Ets-1	NMR	(Werner *et al.*, 1997)	2STT
Mouse	Ets-1	NMR	(Donaldson *et al.*, 1996)	1ETD
Human	Fli-1	NMR	(Liang *et al.*, 1994)	1FLI
Mouse	GABPα	X-ray Diffraction	(Batchelor *et al.*, 1998)	1AWC
Mouse	PU.1	X-ray Diffraction	(Kodandapani *et al.*, 1996)	1PUE
Human	SAP-1	X-ray Diffraction	(Mo *et al.*, 1998)	1BC8

Table 1.1. Reported ETS domain structures from various Ets proteins.

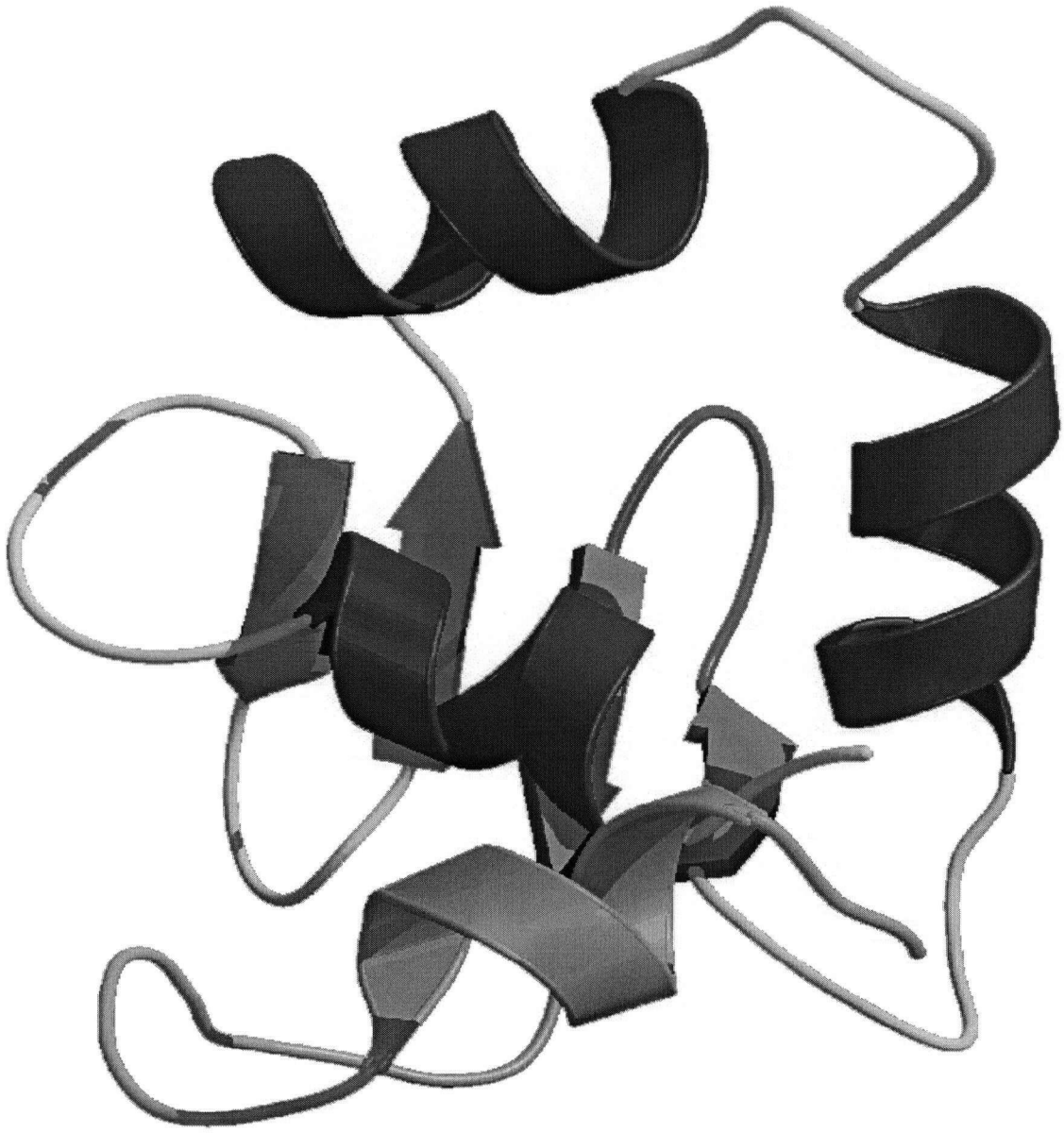


Figure 1.1. NMR solution structure of the ETS domain of murine Ets-1. β-strands appear as arrows and α-helices as coils. From Donaldson *et al.* (1996).

activities, that multiple Ets proteins are often present in each cell type, and that the proteins function by binding a conserved site on the DNA with a conserved domain, the question arises of how biological specificity is achieved within the family. A variety of mechanisms could be involved including subtle differences in sequence-specific DNA binding, differential responsiveness to signal transduction cascades, and combinatorial control of protein partnerships.

1.2 GA Binding Protein (GABP)

GABP provides an insightful example as to the role that protein partnerships play in providing specificity within the Ets family. GABP is involved in the expression of nuclear genes encoding mitochondrial proteins (Virbasius *et al.*, 1993), herpes simplex virus immediate-early genes (Triezenberg *et al.*, 1988), and adenovirus early genes (Watanabe *et al.*, 1993). A diagram of the GABP $\alpha_2\beta_2$ heterotetramer bound to its DNA site is shown in Figure 1.2. The α subunit contains the DNA-binding ETS domain, but only weakly activates transcription. The β subunit is a member of the Notch-Ankyrin repeat family of proteins and contains a transactivation domain but does not directly bind DNA itself. The amino-terminus of GABPβ consists of four and a half ankyrin repeats that mediate heterodimerization with GABPα (Thompson *et al.*, 1991). This heterodimer is able to bind DNA with an affinity greater than the GABPα subunit alone (Batchelor *et al.*, 1998; Thompson *et al.*, 1991). In addition, the carboxy-terminus of GABPβ also contains a leucine zipper motif which allows self-association of GABPα/β heterodimers to form a heterotetramer that can recognize promoters containing tandem 5'-GGA-3' sequences (Thompson *et al.*, 1991).

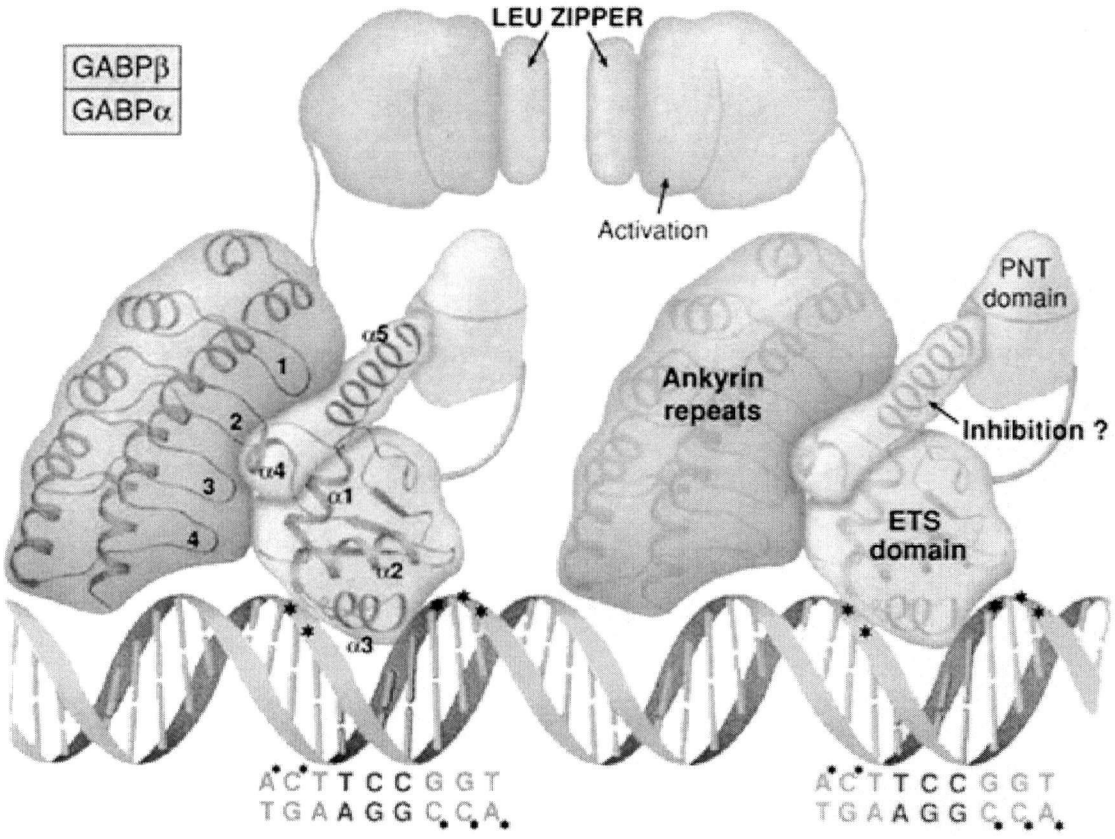


Figure 1.2. GABP $\alpha_2\beta_2$ heterotetramer bound to DNA. The ETS domain and carboxy-terminal helix of the α subunit and the ankyrin repeat region of the β subunit are shown as ribbon diagrams with the remainder of the proteins shown schematically. The asterisks indicate DNA backbone contacts observed in the crystal structure. From Graves (1998).

1.3 The PNT domain

In addition to the signature DNA-binding ETS domain, several members of the Ets family contain a second region of conserved sequence, known as the Pointed (PNT) domain (Klambt, 1993). The PNT domain is found in over 250 regulatory proteins including the Ets family of transcription factors, *Drosophila* Polycomb group transcriptional repressors, pheromone response pathway proteins from *S. pombe* and *S. cerevisiae*, and the ephrin family of receptor tyrosine kinases (Schultz *et al.*, 1997). However, when present in non-Ets proteins, the PNT domain is usually referred to as the SAM, SPM or HLH domain (Ponting, 1995). A sequence alignment of the PNT domains from several Ets proteins appears in Figure 1.3. Although the exact functions of this domain remain elusive, there is a mounting body of evidence that the PNT domain mediates interactions with partner transcription factors and responsiveness to phosphorylation-dependent signaling pathways by acting as a MAPK docking site. Thus, the PNT domain may play an important role in regulating the biological specificity of Ets proteins.

1.3.1 Structures of the PNT domain

The structures of several PNT domains have been solved to date as summarized in Table 1.2. These structures have a similar global topology consisting of a four or five-helix bundle. The PNT domain of GABPα, solved in the McIntosh laboratory using NMR spectroscopy, is shown in Figure 1.4. Helices H2-H5 form a compact four-helix bundle against which packs the most amino-terminal helix, H1. The PNT domain of Ets-1, also solved in the McIntosh laboratory, has a similar structure (Slupsky *et al.*, 1998).

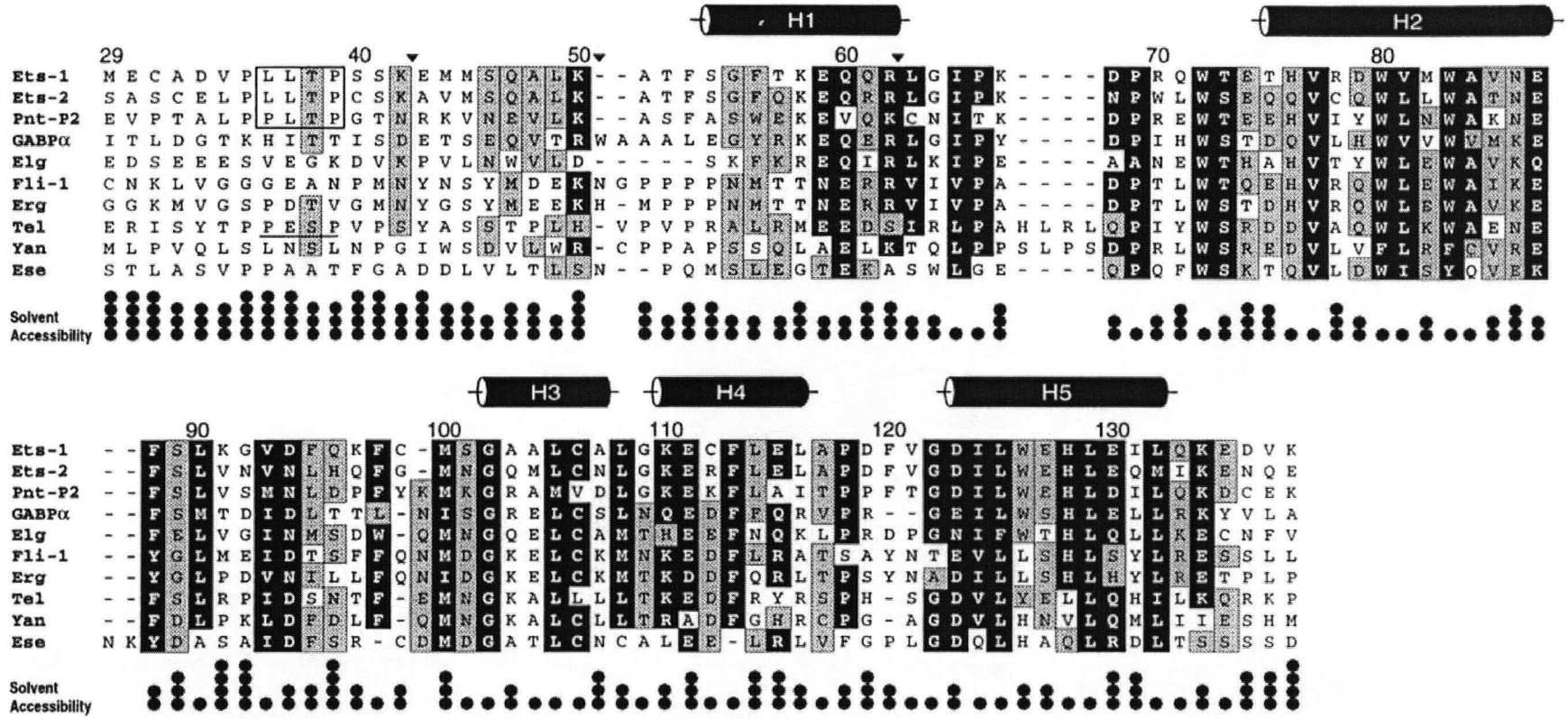


Figure 1.3. Sequence alignment of PNT domains from several Ets proteins. Highly conserved residues are boxed in black and moderately conserved residues in gray. MAPK phosphorylation sites in the preceding amino-terminal region of Ets-1, Ets-2 and Pnt-P2 are boxed and a potential MAPK site in Tel is underlined. Numbering and positions of the 5 α-helices for the PNT domain of Ets-1 are shown. From Slupsky *et al.* (1998).

Protein	Association State	Method	Reference	PDB ID
GABPα	Monomeric	NMR	(McIntosh lab)	N/S
Ets-1	Monomeric	NMR	(Slupsky *et al.*, 1998)	1BQV
Tel	Oligomeric	X-ray Diffraction	(Kim *et al.*, 2001)	1JI7
Ephrin Receptor A4	Dimeric	X-ray Diffraction	(Stapleton *et al.*, 1999)	1B0X
Ephrin Receptor B2	Oligomeric	X-ray Diffraction	(Thanos *et al.*, 1999b)	1B4F
Ephrin Receptor B2	Oligomeric	NMR	(Smalla *et al.*, 1999)	1SGG
Ephrin Receptor B2	Monomeric	X-ray Diffraction	(Thanos *et al.*, 1999a)	1F0M
p73	Monomeric	NMR	(Chi *et al.*, 1999)	1COK

Table 1.2. Reported PNT and SAM domain structures. N/S = Not Submitted

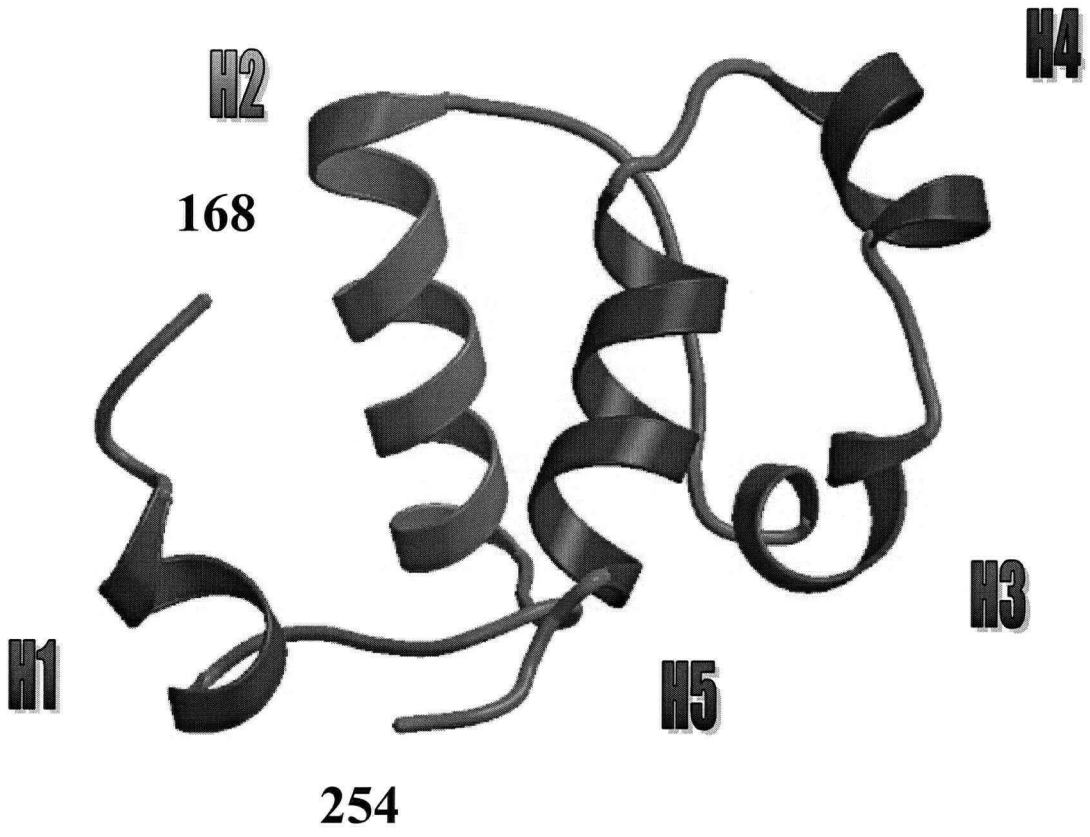


Figure 1.4. Backbone structure of the PNT domain of GABPα (residues 168-254).

Although these domains are monomeric, as shown in Figure 1.5, their structures reveal possible surfaces for interaction with as yet unknown proteins. The Ets-1 PNT domain surface contains two hydrophobic patches surrounded by polar and charged residues which are common protein-protein interfaces (Larsen *et al.*, 1998).

Although the structures of the Ets-1 and GABPα PNT domains are similar, there are some variations in other reported structures of this module. For example, the recently solved structure of the Tel PNT domain lacks the most amino-terminal helix (Kim *et al.*, 2001), resulting in a four helix bundle. Recent studies conducted in the McIntosh laboratory on the PNT domain from the Ets protein Erg also suggest that, similar to the Tel PNT domain, the amino-terminal helix is absent in this protein. In addition, the p73 and ephrin SAM domain structures show some differences in the locations and length of the helices compared to the PNT domain of GABPα. Like Tel and Erg, they lack helix H1 but also have an additional short helix present in the loop region connecting helices H3 and H4. These structural studies suggest that, although the PNT domain does show a high degree of structural homology, minor variations may have evolved in different protein families to provide specialized functions.

Despite these differences, studies conducted on PNT domains from several protein families strongly suggest that an underlying role of the domain is the mediation of protein-protein interactions in multi-protein complexes. In addition, recent reports suggest that one specific function of the PNT domain in some Ets proteins may be to serve as a MAPK docking site. The experimental evidence supporting these hypotheses is presented in the following sections.

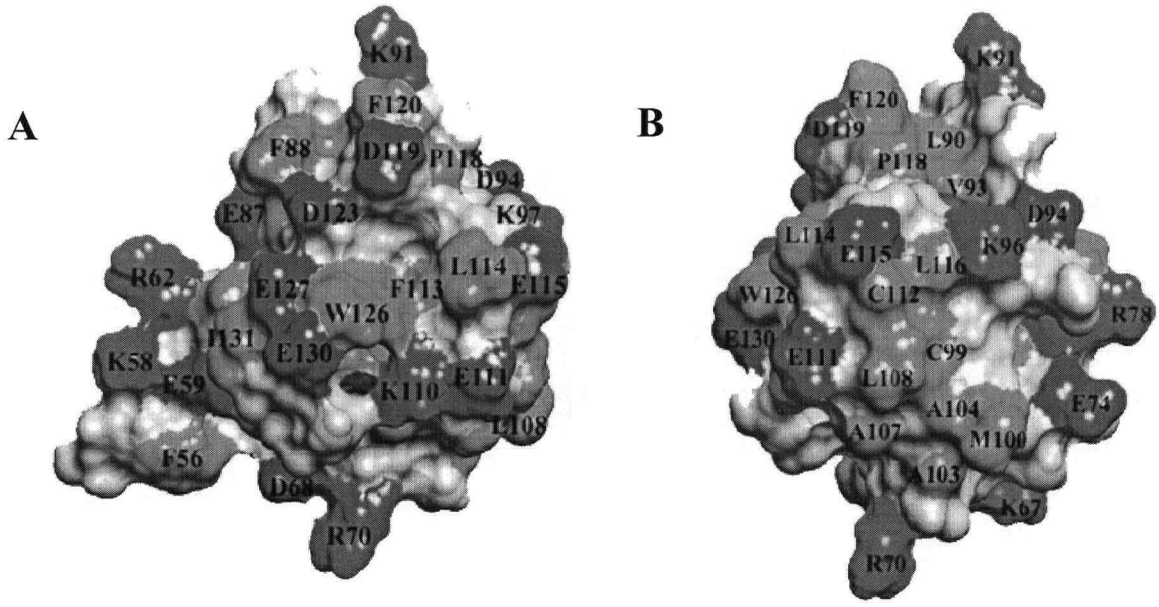


Figure 1.5. Two views of the van der Waals surface of the murine Ets-1 PNT domain, residues 54-135, showing putative protein interaction interfaces. **A)** One hydrophobic patch formed by W126, F113 and L114 that is surrounded by seven charged residues (K110, E111, E115, D119, D123, E127, E130). **B)** On another part of the surface, a large hydrophobic patch is formed by eight residues (C99, M100, A103, A104, A107, L108, C112, L116) that are surrounded by seven charged residues (K67, R70, E74, R78, K96, E111, E115). From Slupsky *et al.* (1998).

1.4 Evidence implicating the PNT domain in protein-protein interactions

1.4.1 Self-association of the Tel PNT domain

There are many reports of chromosomal translocations involving the Ets protein Tel that are associated with leukemia. In many of these, a chimeric gene is formed when the PNT domain of Tel is fused to a fragment of another protein. A summary of these fusions is given in Table 1.3. As can be seen from this table, the majority of these translocations fuse the Tel PNT domain to a protein tyrosine kinase domain. In each case, the PNT domain of Tel mediates self-association of these fusion proteins in vitro and in vivo, leading to constitutive tyrosine kinase activity (Carroll *et al.*, 1996; Golub *et al.*, 1996; Ho *et al.*, 1999; Jousset *et al.*, 1997; Lacronique *et al.*, 1997). Studies using deletion mutants and chimeric proteins indicate that the PNT domain of Tel is sufficient for oligomerization regardless of the context in which it is found (Jousset *et al.*, 1997). The recently reported crystal structure of an oligomeric form of the Tel PNT domain (Kim *et al.*, 2001) supports this view.

1.4.2 Association of the Tel and Fli-1 PNT domains

In addition to the well-studied Tel fusion proteins, several groups have reported formation of hetero-oligomers between some Ets proteins. The Ets proteins Tel and Fli-1 have been reported to interact both *in vitro* and *in vivo* and the interaction appears to be mediated in part by the PNT domains of each protein (Kwiatkowski *et al.*, 1998). In this study, the first 371 amino acids of Tel, including the PNT domain, was used as bait in a yeast two-hybrid screen against a library constructed from a murine pluripotent hematopoietic cell line. Two Ets proteins were identified in this screen, a fragment of Tel

Partner Protein	Function of Partner	Reference
ABL	Non-receptor tyrosine kinase	(Golub *et al.*, 1996; Papadopoulos *et al.*, 1995)
AML1	Transcription factor	(Golub *et al.*, 1995; Romana *et al.*, 1995)
ARG	Non-receptor tyrosine kinase	(Cazzaniga *et al.*, 1999; Iijima *et al.*, 2000)
ARNT	Transcription factor	(Salomon-Nguyen *et al.*, 2000)
JAK2	Receptor-associated kinase	(Lacronique *et al.*, 1997; Peeters *et al.*, 1997)
PDGFRβ	Receptor tyrosine kinase	(Golub *et al.*, 1994)
SYK	Non-receptor tyrosine kinase	(Kuno *et al.*, 2001)
TRKC	Receptor tyrosine kinase	(Eguchi *et al.*, 1999; Knezevich *et al.*, 1998)

Table 1.3. Reported fusions of the Tel PNT domain to various proteins.

itself and two fragments of another Ets protein, Fli-1 (Figure 1.6). It is important to note that all of these fragments contain the PNT domains of their respective Ets proteins. To confirm these interactions *in vitro*, co-immunoprecipitation experiments using proteins expressed in rabbit reticulocyte lysates were conducted. Anti-Fli-1 antibody specifically precipitated full-length Fli-1 and Tel 1-371. In addition, a GST pull-down assay was performed to confirm these interactions. A GST-fusion of full-length Fli-1 specifically bound *in vitro* translated Tel. Also, a GST-fusion of Tel 1-371 bound *in vitro* translated Tel, confirming results obtained by other groups working with Tel protein constructs.

In vivo studies confirmed Tel Fli-1 interactions in transfected cells. Full length wild-type Tel and Tel with a Myc domain at the amino-terminus were transiently transfected into a 293 cell line. Anti-Myc antibody precipitated wild-type Tel as well as Myc tagged Tel. In a similar experiment using full-length Fli-1 and Tel proteins in K562 cells, Tel and Fli-1 were specifically immunoprecipitated using anti-Fli-1 antibody.

A possible biological role of the Fli-1-Tel interaction was demonstrated using megakaryocytic promoter constructs in 293 cells. Transfected Fli-1 resulted in an increase in promoter activity, however, this increase was completely inhibited by the presence of full-length Tel. Inclusion of a fragment of Tel containing the PNT domain resulted in a 50% inhibition of Fli-1 transactivation. These biological effects were confirmed and investigated in more detail in a later study (Kwiatkowski *et al.*, 2000) in which Tel and Fli-1 were expressed separately and together in the K562 leukemia cell line, which does not express endogenous Tel or Fli-1. Fli-1 expressed alone in these cells induced a megakaryocytic phenotype that was reversed upon expression of Tel.

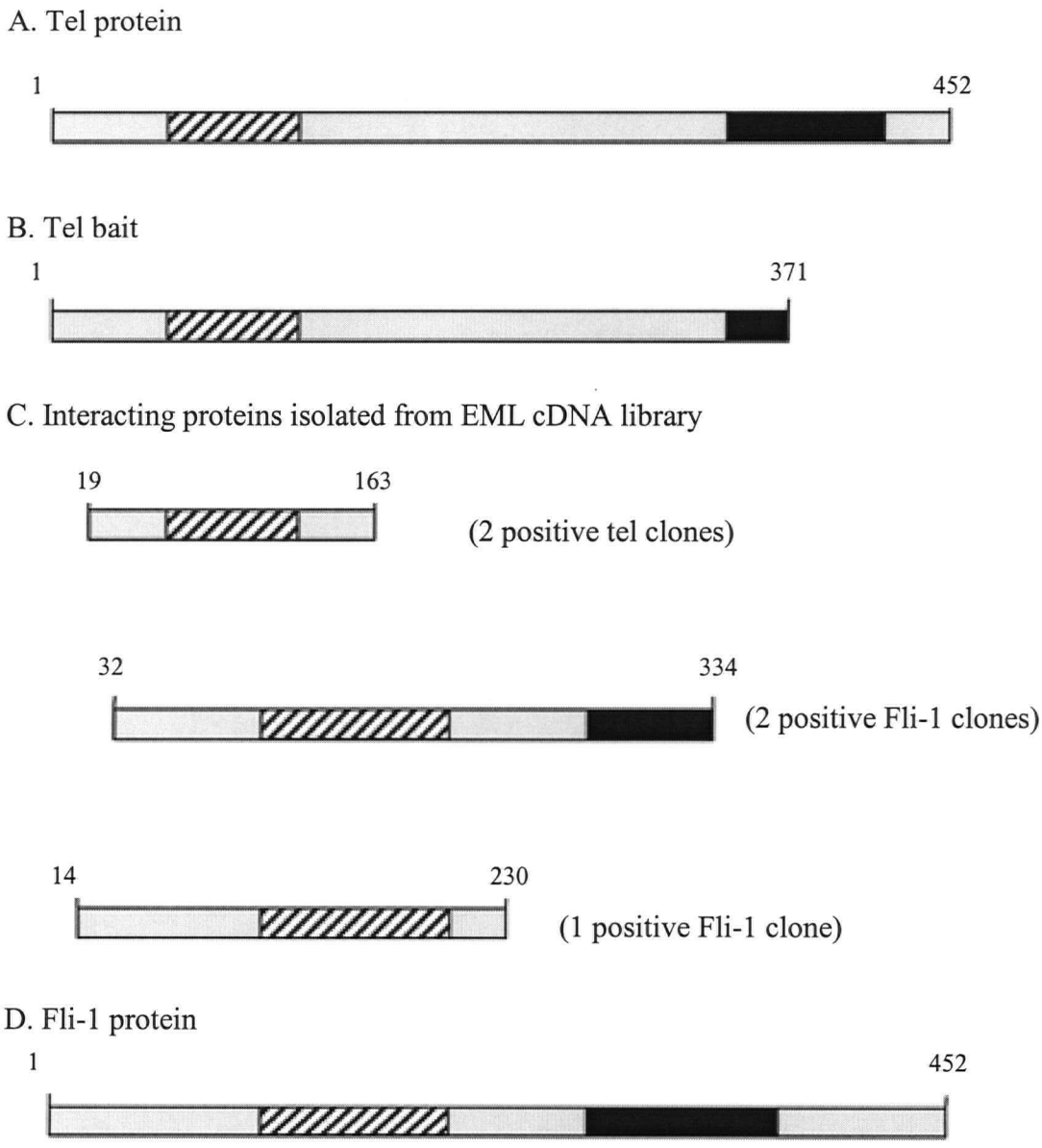


Figure 1.6. Tel bait and partners isolated in a yeast two-hybrid screen. Striped bars represent the PNT domains of the proteins while the solid black bars represent the ETS domains. **A)** Representation of the full-length Tel protein. **B)** Tel bait used in the screen. **C)** Tel and Fli-1 fragments isolated in the screen. **D)** Representation of the full-length Fli-1 protein. From Kwiatkowski *et al.* (1998).

Although these experiments were not conducted using the minimal PNT domains of Tel and Fli-1, they do suggest that the PNT domain is playing an important role in the interaction of these proteins and that this interaction could alter the transactivation activities of these transcription factors. However, attempts to reproduce this interaction in the McIntosh laboratory using the minimal PNT domains from Tel and Fli-1 have been unsuccessful. Possible explanations for the discrepancy are discussed in chapter 5.

1.4.3 Association of the PNT domains of Erg and other Ets proteins

Two studies have reported direct interactions between Ergp55 and other Ets proteins. An initial study demonstrated that Ets-2 was able to activate the stromelysin-1 gene but that co-expression with Erg inhibited this activation (Buttice *et al.*, 1996). This situation is analogous to that described above for Tel and Fli-1. Basuyaux *et al.* (1997) investigated this potential interaction between Erg and Ets-2. A GST-fusion of full-length Ets-2 interacted strongly with *in vitro* translated Erg and weakly with *in vitro* translated Ets-1. To identify domains on Ets-2 responsible for the interactions, a series of deletions was constructed. Erg was still able to bind the first 129 amino acids of Ets-2, a region that includes the PNT domain, in a DNA-independent manner.

Carrere *et al.* (1998) investigated Erg protein interactions in more detail. Co-immunoprecipitation assays demonstrated self-association of Erg proteins and interactions between Erg and Fli-1 and Erg and Ets-2. Interactions were also found between Erg and Er81 and Erg and PU-1, Ets proteins that lack a PNT domain, indicating that the domain is not essential for Ets protein interactions. However, deletion constructs

of Erg demonstrated that the PNT domain does play a role in these interactions, as an Erg construct containing only the first 200 amino acids was still able to interact with a full-length Erg protein. Thus, Ergp55 seems to self-associate via both the PNT and ETS domains. However, like the reported Tel Fli-1 interaction, work in the McIntosh laboratory using the minimal Erg PNT domain has failed to reproduce the above results. This is discussed in greater detail in chapter 5.

1.4.4 Interaction of UBC9 with the PNT domains of Ets-1 and Tel

The PNT domain has also been found to be involved in interactions with non-Ets proteins. Two reports suggest that a ubiquitin conjugating enzyme, UBC9, may interact with the PNT domains of several Ets proteins. Hahn *et al.* (1997) conducted a yeast two-hybrid screen using various Ets-1 constructs against a library constructed from EBV-transformed peripheral blood B-cells. The majority of clones encoded the human homologue of the *S. cerevisiae* E2 ubiquitin conjugating enzyme, UBC9. Further two-hybrid experiments demonstrated an interaction between huUBC9 and the first 130 amino acids of Ets-1, which includes the PNT domain.

The interaction of these proteins *in vitro* was confirmed using a GST-huUBC9 fusion and *in vitro* translated full-length Ets-1. Other Ets proteins were tested for interaction with huUBC9 as well. Fli-1, which contains a PNT domain, and Elf-1, PU.1 and Net, which do not, interacted with huUBC9. So, similar to the results reported for Erg as described in section 1.4.3, the PNT domain is not essential for these interactions but does play a role when it is present. Transfection experiments indicated that huUBC9

was able to increase the transactivation function of Ets-1 and that this was not dependent on the ubiquitin-conjugating activity of huUBC9.

A second group also identified huUBC9 as a PNT domain interacting protein using a yeast two-hybrid screen with Tel as bait (Chakrabarti *et al.*, 1999). However, in this case, the PNT domain of Tel was found to be essential for the interaction with huUBC9. These interactions were confirmed *in vitro* and *in vivo.*

1.4.5 Interaction of Daxx with the Ets-1 PNT domain

Use of the human Ets-1 PNT domain and amino terminal MAPK site in a yeast two-hybrid screen using a human fetal brain cDNA library as prey identified Daxx as an interacting protein (Li *et al.*, 2000). Daxx was previously reported to bind to the Fas cell death receptor and plays a role in apoptosis, possibly through activation of the JNK pathway (Yang *et al.*, 1997). *In vitro* GST-fusion pull-down experiments confirmed the interaction between the Ets-1 PNT domain and Daxx and also demonstrated that the carboxy-terminal region of Daxx mediated Ets-1 PNT domain binding. In addition, Daxx was shown to repress Ets-1 mediated transcription in transfected cells.

1.4.6 Self-association of Polycomb group proteins via the SAM domain

The *Drosophila* Polycomb group transcriptional repressors form multimeric structures that regulate chromatin structure and gene expression. Several of these proteins, including polyhomeotic (ph) and Sex comb on midleg (Scm), contain SAM domains at their carboxy-termini. Yeast two-hybrid, mutagenesis and *in vitro* binding experiments have demonstrated that the minimal SAM domains of ph and Scm are

necessary and sufficient to mediate homo- and heterotypic interactions between these proteins (Kyba & Brock, 1998; Peterson *et al.*, 1997).

1.4.7 Association of yeast pheromone response pathway components

Pheromone response pathways in *S. cerevisiae* and *S. pombe* consist of a MAPK module. Using a yeast two-hybrid screen, Barr *et al.* (1996) demonstrated an interaction between Byr2, a MAP/ERK kinase kinase, and a previously cloned protein, Ste4. Simple deletion studies established that the amino-terminal regions of each protein, which contain SAM domains, are responsible for mediating this interaction. A later study using more extensive deletions and point mutations demonstrated that the minimal SAM domain of Byr2 is sufficient for interaction with Ste4 and that this interaction is essential for Byr2 activity (Tu *et al.*, 1997). Similar results have been reported for the *S. cerevisiae* homologues of Byr2 and Ste4, Ste11 and Ste50 (Wu *et al.*, 1999).

1.4.8 Possible ephrin receptor association via the SAM domain

The ephrin family of tyrosine kinase receptors has been implicated in the regulation of a variety of processes including axon guidance, angiogenesis and cell migration. All ephrin receptors have a carboxy terminal SAM domain but the function of the domain in this protein family is not well understood. As can be seen from Table 1.2, a number of different multimeric forms of isolated ephrin SAM domains have been reported, even for the same protein in the case of ephrin B2. The interaction interfaces reported in these structures differ and the isolated domains associate very weakly in solution with dissociation constants in the mM range (Smalla *et al.*, 1999; Stapleton *et*

al., 1999; Thanos *et al.*, 1999a; Thanos *et al.*, 1999b). While these observations do not preclude an important role for the SAM domain in ephrin receptor function, the biological significance of these results remains unknown.

1.5 Evidence for the involvement of the PNT domain in MAPK signaling

1.5.1 The PNT domain as a MAPK docking site

Recent work has suggested that one specific function of the PNT domain may be as a MAP kinase docking site. Affinity chromatography using the minimal PNT domain from Ets-1 without its adjacent phosphorylation site demonstrated that the domain specifically bound the MAP kinase ERK2 from a calf thymus extract. Enzymatic studies demonstrated that the dissociation constant for ERK2 binding to the Ets-1 PNT domain and MAPK site was approximately ten times stronger than binding to the peptide MAPK site alone (B. Graves, personal communication). The substrate specificities of many protein kinases are increased by docking sites, short sequences in the substrate that bind part of the kinase but which are secondary to the phosphoacceptor site (Holland & Cooper, 1999).

1.5.2 Interaction of the PNT domain of Yan with the SAM domain of Mae

A recent study has reported the discovery of another protein that interacts with the PNT domain which may have an important role in MAPK signaling. Baker *et al.* (2001) conducted a yeast two-hybrid screen using the full-length Ets protein Yan as bait against a *Drosophila* embryo library. They recovered a novel interacting protein, called Mae, containing a SAM domain at its carboxy-terminus. The interaction was confirmed *in vitro*

and deletion studies demonstrated that the PNT domain of Yan and the SAM domain of Mae mediated the interaction. Interestingly, Mae appears to be a signaling intermediate that links the MAPK signaling pathway to transcription factor targets, as it is required for phosphorylation of Yan. Binding of Mae to the PNT domain affects the ability of Yan to bind DNA through its ETS domain, suggesting that Mae causes a local conformational change in Yan. This could expose the phosphorylation site and allow Yan to bind MAPK. Alternatively, Mae itself could be acting as the MAPK docking site.

1.6 Questions, goals and significance

The evidence stated above suggests that the general function of the PNT domain is the mediation of protein-protein interactions. Based on this evidence, the following hypothesis was formed: **The PNT domain mediates interactions of Ets proteins with either PNT domain containing-proteins or with other proteins involved in signal transduction and transcriptional regulation.**

However, despite the above evidence that the PNT domain mediates protein-protein interactions, the precise function of this domain in the context of Ets transcription factors remains unclear. Although several PNT and SAM domain structures are known, a detailed biophysical study of PNT domain interactions has yet to be carried out. The studies cited above have all used qualitative techniques such as yeast two-hybrid screens, binding assays and immunoprecipitations that demonstrate the existence of the interactions, but do not characterize them in detail. In addition, these studies often use large fragments of Ets proteins that contain the PNT domain. As a result, with the exception of the Tel PNT domain, studies on minimal PNT domains are lacking. In

addition, although several oligomeric forms of the domain have been reported, the interaction interfaces differ considerably, even for the same protein, bringing into question the biological relevance of the interactions. Further structural and biophysical studies of PNT domains would help to clarify the mechanism of interaction.

Despite the fact that the Ets protein GABPα has a very wide tissue distribution, no protein interactions involving its PNT domain have been reported to date. However, recent work in the McIntosh laboratory has suggested that the GABPα PNT domain may self-associate and that a small molecule may bind the domain.

As a result, three main goals of this thesis work were formulated:

i) To use a yeast two-hybrid system to screen for potential protein partners of the PNT domain of GABPα. This would be followed by structural and thermodynamic characterization of the *in vitro* complexes of the PNT domain and any protein partners identified in the screen.

ii) A detailed study of the PNT domain of GABPα to investigate the possibility of self-association. This would include structural and thermodynamic characterization of the complex.

iii) A detailed study of the PNT domain of GABPα to investigate the possibility of a small molecule binding to the GABPα PNT domain. Identification of this molecule and a biophysical description of its binding could reveal a novel function of the domain.

Work toward each of these goals is described in chapters 2, 3 and 4, respectively.

Gene expression relies on the formation of multi-protein complexes and combinatorial protein interactions. Thus, an understanding of the formation of these

complexes may aid in a more complete understanding of gene expression and identification of possible targets for therapeutic intervention. In addition, this information is essential for a complete understanding of the role the PNT domain plays in the control of gene expression by the Ets transcription factors.

Chapter 2

Yeast Two-Hybrid Screen Using the GABPα PNT Domain

As stated in section 1.6, identification of a novel PNT domain binding protein followed by structural elucidation of the complex of a PNT domain and this protein may help clarify the molecular mechanism of PNT domain interactions. This chapter describes the use of the yeast two-hybrid system to identify proteins that interact with the PNT domain of GABPα. This Ets family member was chosen due to its expression in a broad range of tissue types, thus maximizing the probability of identifying an interacting protein from available two-hybrid libraries.

2.1 Introduction

The yeast two-hybrid system, originally described by Fields and colleagues (Fields & Song, 1989), has become a powerful tool for identifying protein-protein interactions. This system is based on earlier research, which found that the DNA-binding and transactivation domains of various transcription factors could function individually (Brent & Ptashne, 1985; Keegan *et al.*, 1986; Ma & Ptashne, 1987). A diagram of the yeast two-hybrid system, which is based on two fusion proteins, appears in Figure 2.1. The bait consists of the DNA binding domain of a transcription factor fused a protein of interest. This species is able to bind DNA by virtue of the DNA binding domain, but is unable to activate transcription. The second fusion protein, the prey, consists of the protein products of a cDNA library fused to the activation domain of a transcription

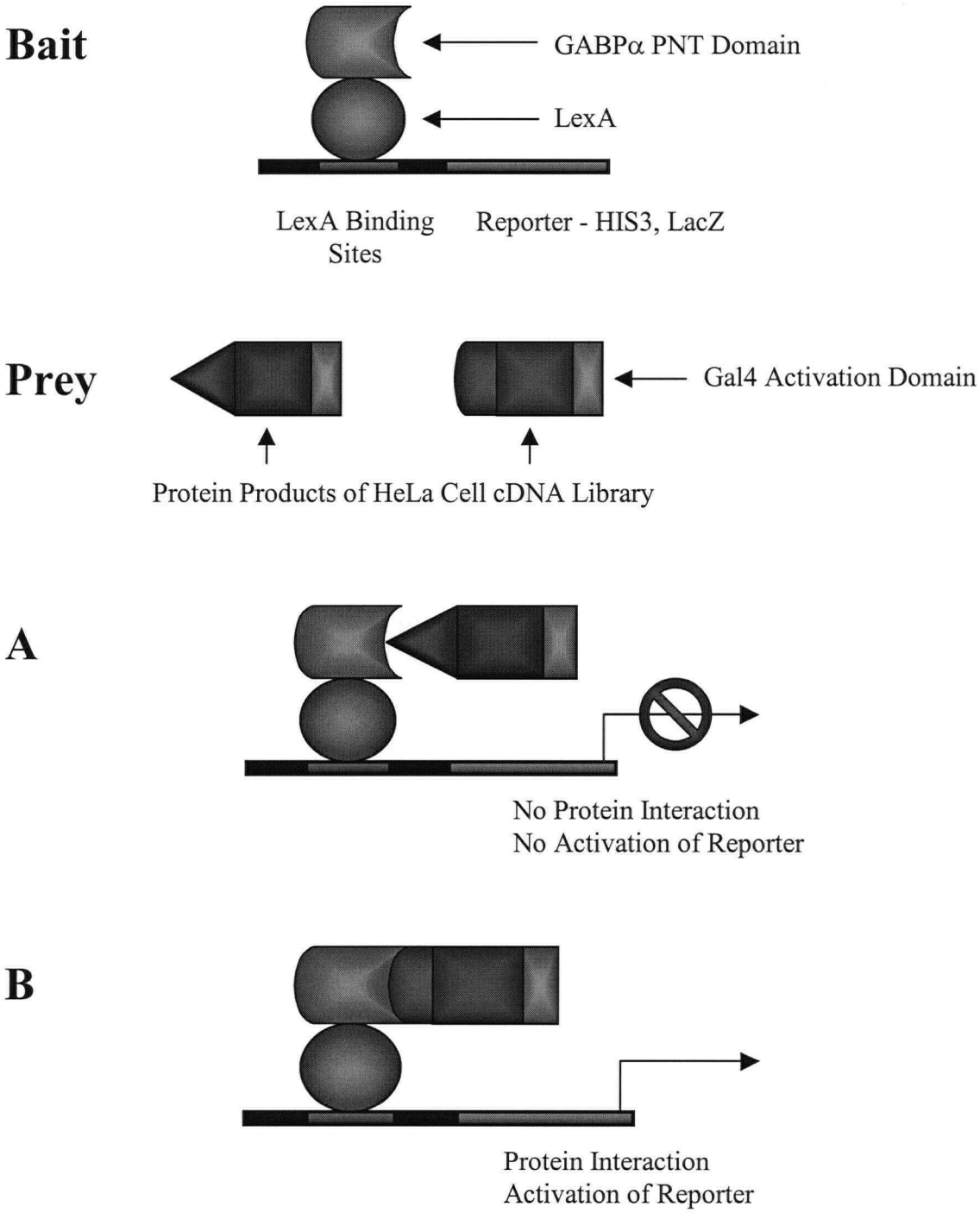


Figure 2.1. Schematic diagram of the yeast two-hybrid system. **A)** Protein products that do not interact with the GABPα PNT domain do not activate the reporters. **B)** A protein product that does interact with the PNT domain results in activation of the reporter.

factor. Members of a pool of fusion proteins are able to activate transcription, but do not result in reporter activation due to the lack of a DNA-binding domain. Interaction of a prey protein with the bait results in the prey fusion protein being localized to the DNA, allowing activation of the reporter.

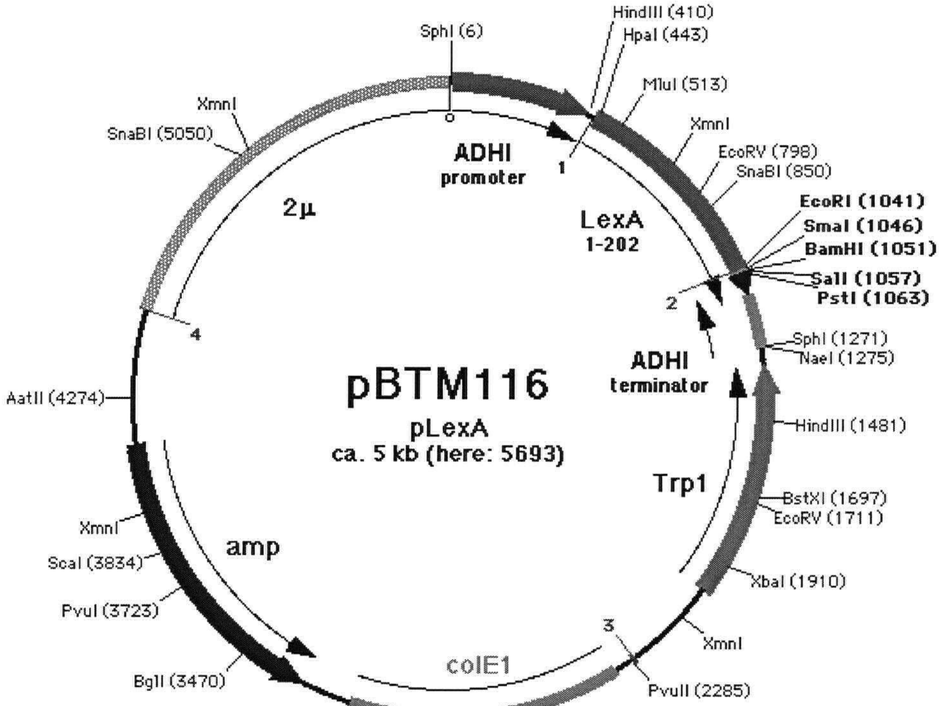
2.2 Materials and Methods

2.2.1 Host Strain

The *Saccharomyces cerevisiae* reporter strain L40 (Hollenberg *et al.*, 1995; Vojtek *et al.*, 1993) contains two integrated reporter constructs, the yeast *HIS3* gene and the bacterial *lacZ* gene, under control of LexA operators. Association of the two fusion proteins activates transcription of the His3 and LacZ genes. Thus, interacting proteins can be identified using His3 selection, which is rapid and allows the use of large, high complexity libraries. Proteins identified in this initial screen can then be assayed independently using an assay for β-galactosidase.

2.2.2 Bait Plasmid

The pBTM116 bait plasmid, shown in Figure 2.2, contains the entire coding region of the *E. coli* LexA protein which was fused to the GABPα PNT domain (amino acids 168-254). Amino acids 66-230 of human lamin C were fused to LexA (Vojtek *et al.*, 1993) for use as a negative control.




```
                                                                                  LexA
                      HindIII (410)             1             HpaI (443)
 391 CAAGCATACAATCAACTCCAAGCTTGAAGCAAGCCTCCTGAAAG ATG AAA GCG TTA ACG GCC AGG CAA CAA GAG GTG TTT GAT
                                                1▶Met Lys Ala Leu Thr Ala Arg Gln Gln Glu Val Phe Asp
                                                             MluI (513)
 474 CTC ATC CGT GAT CAC ATC AGC CAG ACA GGT ATG CCG CCG ACG CGT GCG GAA ATC GCG CAG CGT TTG GGG TTC
  14▶Leu Ile Arg Asp His Ile Ser Gln Thr Gly Met Pro Pro Thr Arg Ala Glu Ile Ala Gln Arg Leu Gly Phe
 546 CGT TCC CCA AAC GCG GCT GAA GAA CAT CTG AAG GCG CTG GCA CGC AAA GGC GTT ATT GAA ATT GTT TCC GGC
  38▶Arg Ser Pro Asn Ala Ala Glu Glu His Leu Lys Ala Leu Ala Arg Lys Gly Val Ile Glu Ile Val Ser Gly
 618 GCA TCA CGC GGG ATT CGT CTG TTG CAG GAA GAG GAA GAA GGG TTG CCG CTG GTA GGT CGT GTG GCT GCC GGT
  62▶Ala Ser Arg Gly Ile Arg Leu Leu Gln Glu Glu Glu Glu Gly Leu Pro Leu Val Gly Arg Val Ala Ala Gly
     XmnI
 690 GAA CCA CTT CTG GCG CAA CAG CAT ATT GAA GGT CAT TAT CAG GTC GAT CCT TCC TTA TTC AAG CCG AAT GCT
  86▶Glu Pro Leu Leu Ala Gln Gln His Ile Glu Gly His Tyr Gln Val Asp Pro Ser Leu Phe Lys Pro Asn Ala
                                                    EcoRV (798)
 762 GAT TTC CTG CTG CGC GTC AGC GGG ATG TCG ATG AAA GAT ATC GGC ATT ATG GAT GGT GAC TTG CTG GCA GTG
 110▶Asp Phe Leu Leu Arg Val Ser Gly Met Ser Met Lys Asp Ile Gly Ile Met Asp Gly Asp Leu Leu Ala Val
                      SnaBI (850)
 834 CAT AAA ACT CAG GAT GTA CGT AAC GGT CAG GTC GTT GTC GCA CGT ATT GAT GAC GAA GTT ACC GTT AAG CGC
 134▶His Lys Thr Gln Asp Val Arg Asn Gly Gln Val Val Val Ala Arg Ile Asp Asp Glu Val Thr Val Lys Arg
 906 CTG AAA AAA CAG GGC AAT AAA GTC GAA CTG TTG CCA GAA AAT AGC GAG TTT AAA CCA ATT GTC GTT GAC CTT
 158▶Leu Lys Lys Gln Gly Asn Lys Val Glu Leu Leu Pro Glu Asn Ser Glu Phe Lys Pro Ile Val Val Asp Leu
                                                                                     EcoRI
                                                                                     2     SmaI
 978 CGT CAG CAG AGC TTC ACC ATT GAA GGG CTG GCG GTT GGG GTT ATT CGC AAC GGC GAC TGG CTG GAA TTC CCG
 182▶Arg Gln Gln Ser Phe Thr Ile Glu Gly Leu Ala Val Gly Val Ile Arg Asn Gly Asp Trp Leu Glu Phe Pro
      BamHI  SalI    PstI
1050 GGG ATC CGT CGA CCT GCA GCC AAG CTA ATT CCG GGC GAA TTT CTT ATG ATT TAT GAT TTT TAT TAT TAA ATAA
 206▶Gly Ile Arg Arg Pro Ala Ala Lys Leu Ile Pro Gly Glu Phe Leu Met Ile Tyr Asp Phe Tyr Tyr •••
1123 GTTATAAAAAAAATAAGTGTATACAAATTT
```

Figure 2.2. Map and multiple cloning site of the pBTM116 shuttle vector.

2.2.3 Prey Plasmid

The prey plasmid library was constructed by Hannon *et al.* (1993) and consists of the pGADGH plasmid containing the Gal4 activation domain (amino acids 768-881) fused to a HeLa cell library. Poly(A)$^+$ RNA was isolated from HeLa S3 suspension cells and double-stranded cDNA constructed using the Unizap XR cloning kit (Strategene). This cDNA consisted of full-length transcripts with an average size of ~1.5 kb. These were cloned into the pGADGH vector, a map of which appears in Figure 2.3. This library was electroporated into *E. coli* DH10B cells (Bio-Rad) which resulted in ~7×10^6 primary recombinants. Library DNA was then obtained from these transformants.

2.2.4 Preparation and transformation of competent yeast cells

Yeast cells were prepared and transformed using a modification of the LiOAc/SS-DNA/PEG procedure (Schiestl & Gietz, 1989). To prepare competent *S. cerevisiae* cells, 5 ml L40 starter cultures were grown overnight in appropriate dropout media in test tubes at 30°C with shaking at 225 rpm. In the morning, 1 ml of this culture was diluted into 25 ml of appropriate media in a 500 ml Erlenmeyer flask and incubated at 30°C for 4.5 h. Cells were pelleted by centrifugation in 50 ml screw-capped tubes at 2100 × g for 5 min. The supernatant was discarded and the cells washed with 20 ml of sterile water and re-pelleted. Cells were then resuspended in 2 ml of 50 mM LiOAc, 5 mM Tris-HCl pH 7.5, 0.5 mM EDTA, 500 mM sorbitol, 20% (v/v) glycerol, aliquoted, and stored at -70°C.

For transformation, competent yeast cells were thawed on ice. Three µl of plasmid DNA was added to 100 µl of thawed cells along with 4 µl of boiled salmon sperm DNA (10 mg/ml). One hundred µl of 70% PEG 3350 was then added. Tubes

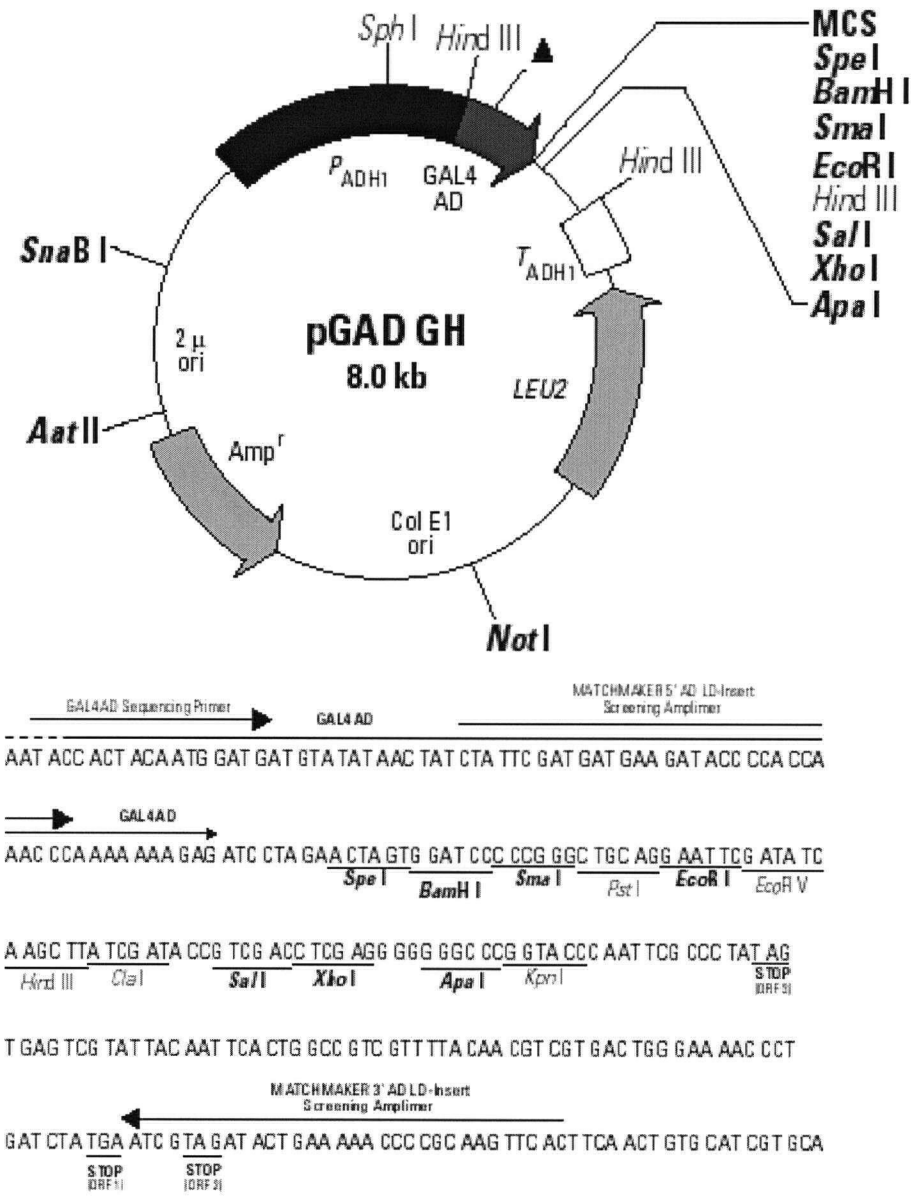


Figure 2.3. Map and multiple cloning site of the pGADGH shuttle vector. This vector is a derivative of pGAD10 and contains a complete ADH promoter and a portion of the pBluescript (Strategene) polylinker.

were incubated at 30°C for 30 min followed by heat shock at 42°C for 15 min. Four hundred µl of 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 1 M sorbitol was then added. The entire suspension was then plated on appropriate dropout media to select for transformants and incubated at 30°C until colonies appeared.

2.2.5 Recovery of prey plasmids from yeast

Prey plasmids from positive transformants were isolated by repeated plating on dropout media lacking only leucine, which eventually resulted in loss of the bait plasmid. Prey plasmid was isolated from these colonies using a modification of the "smash and grab" method (Hoffman & Winston, 1987). L40 cells from a single colony were added to 400 µl of extraction solution (100 mM NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 2% Triton X-100, 1% SDS). One hundred fifty µl of 425-600 micron acid washed glass beads (Sigma) and 200 µl of a phenol:chloroform:isoamyl alcohol mix (25:24:1) were then added. This suspension was vortexed for 1.5 min and then spun at 13 000 rpm in a microfuge for 5 min. The upper aqueous layer was transferred to a fresh tube and 40 µl of 3 M NaOAc and 800 µl of absolute ethanol were added. The tube was then incubated on ice for 5 min and the precipitated DNA pelleted by spinning at 13 000 rpm for 10 min in a microfuge. This pellet was then washed with 200 µl of 70% ethanol and repelleted. The supernatant was discarded and the pellet dried for 15 min in a vacuum dessicator. The dried pellet was resuspended in 50 µl of sterile water. Plasmid from this procedure was then electroporated into competent *E. coli* DH5α cells using a Bio-Rad Gene Pulser II electroporation apparatus operating at 1800 V. Colonies from these transformations were

then used to inoculate overnight cultures in LB media from which plasmid DNA was recovered using the QIAprep Spin Miniprep Kit (Qiagen).

2.2.6 β-galactosidase test of bait-prey interaction

Prey plasmids from the above procedures were transformed L40 cells containing bait plasmid as described in section 2.2.4. To confirm protein interactions, these transformants were then tested for bait-prey interaction using a qualitative ONPG test. L40 transformants were placed in 100 µl of Z buffer (100 mM phosphate pH 7.0, 10 mM KCl, 1 mM $MgSO_4$) with β-mercaptoethanol (2.70 ml to 1 litre Z buffer) and 1 mg/ml ONPG. Five µl of 5% sarcosyl and 5 µl of toluene were then added to gently disrupt the cells. This lysate was vortexed for 30 s and incubated at 37°C. Tubes were then monitored for the appearance of yellow *o*-nitrophenol product at regular intervals. A positive control consisting of Ras bait and Raf prey was included for comparative purposes (Vojtek *et al.*, 1993).

2.2.7 Cloning of PBP

Primers to clone the coding region of PBP out of the pGADGH vector are shown in Figure 2.4. These were used, along with template purified as described in section 2.2.5, to amplify the coding region of PBP using PCR. Reaction products were resolved using agarose gel electrophoresis and the gel stained using ethidium bromide. Bands were visualized under UV illumination and a 700 bp band corresponding to the expected size of the PBP coding region was excised and purified using the QIAEX II Gel Extraction Kit (Qiagen). This purified fragment and a pET28a expression vector (Figure 2.5) were

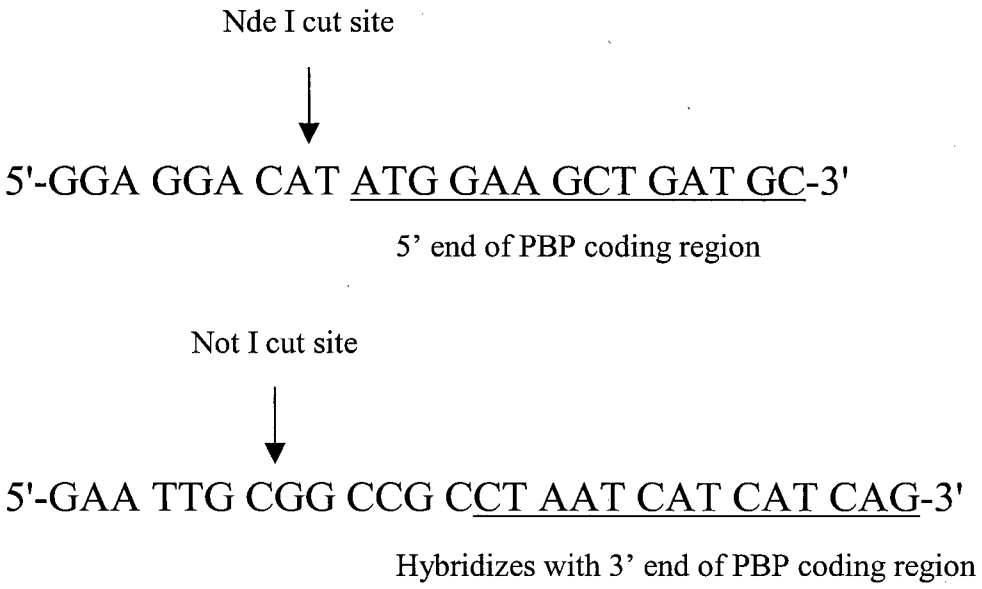


Figure 2.4. Primers to clone the coding region of PBP out of the pGADGH vector. These oligonucleotides were synthesized by GibcoBRL.

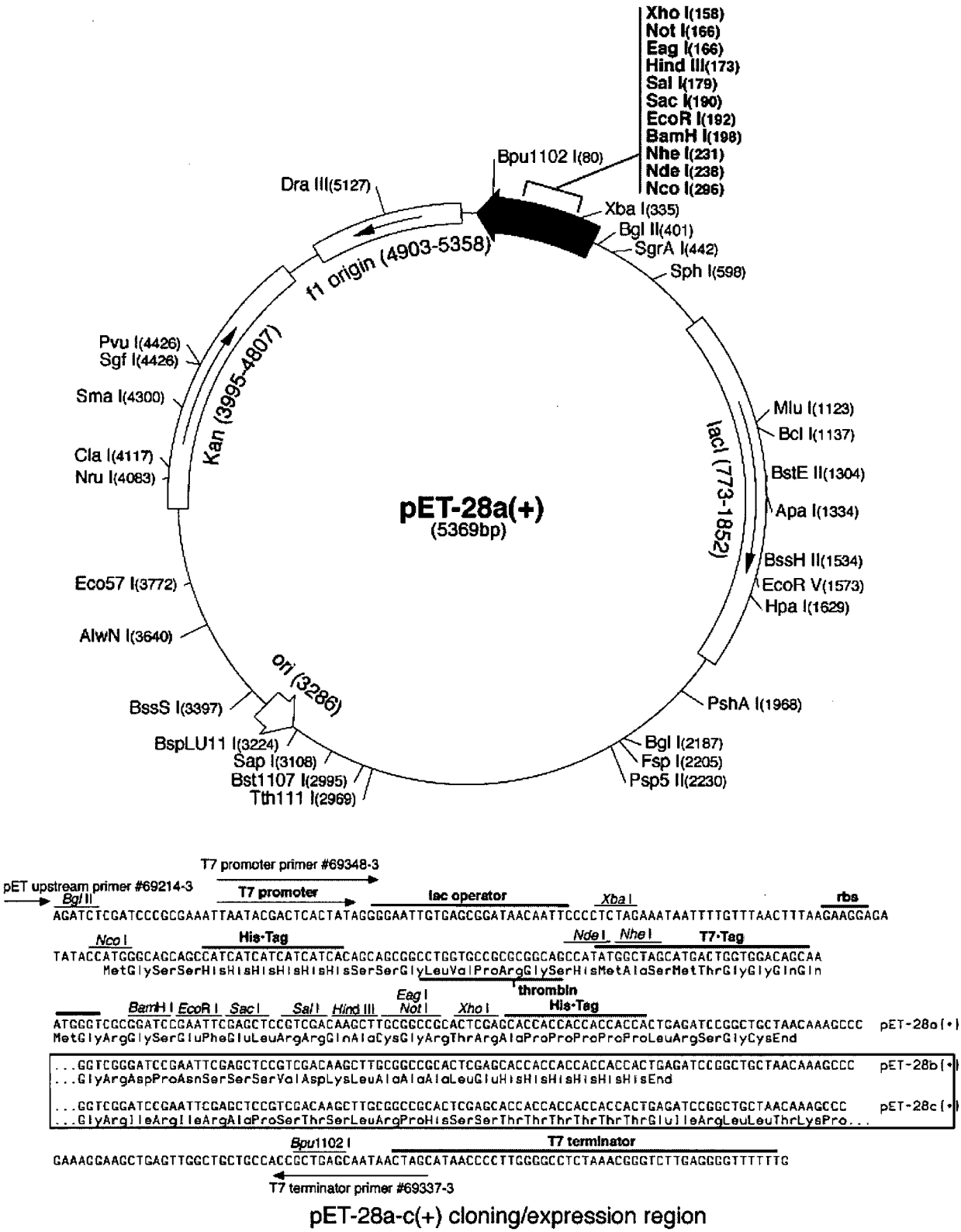


Figure 2.5. pET28a vector map and multiple cloning site (Novagen).

digested with NdeI and NotI restriction enzymes and the fragments resolved using agarose gel electrophoresis. The gel was again stained with ethidium bromide and bands corresponding to the PBP coding region and the linearized pET28a vector were excised and purified using the gel extraction kit. These fragments were then mixed and incubated overnight at 16°C in the presence of T4 DNA ligase. This placed a thrombin cleavable 6×His tag at the amino-terminus of the PBP coding region.

Ligated vectors were electroporated into competent DH5α cells and plasmid DNA was recovered as described in section 2.2.5. DNA sequencing was performed on the plasmids to confirm the sequence of PBP and correct placement of the 6×His tag. This plasmid was then electroporated into *E. coli* BL21(λDE3) cells.

2.2.8 Expression of PBP

PBP was expressed from *E. coli* BL21(λDE3) cells by inoculating 25 ml of LB medium in a 125 ml Erlenmeyer flask with a single colony of *E. coli* BL21(λDE3) cells containing the pET28a-PBP plasmid. This starter culture was incubated overnight at 30°C with constant shaking. In the morning, the culture was transferred to a 50 ml screw-capped tube. Cells were pelleted by low-speed centrifugation and resuspended in 1 ml of fresh LB medium. These resuspended cells were then diluted into 1 litre of LB medium and grown at 30°C with shaking until the OD_{600} of the culture reached 0.6-0.8. Expression of PBP was then induced by adding IPTG to a concentration of 1 mM and continuing incubation at 30°C for 3h. The culture was transferred into 250 ml centrifuge bottles and cells harvested by centrifugation at 1 700 × g for 20 min at 4°C in a Sorvall RC-5B Plus centrifuge using a chilled Sorvall GSA fixed angle rotor. Cells were

resuspended in 20 ml of binding buffer (5 mM imidazole, 50 mM HEPES pH 7.5, 500 mM NaCl, 5% (v/v) glycerol) with 0.1 mM PMSF and were stored at -70°C.

2.2.9 Purification of PBP

Cells expressing PBP were lysed using a French pressure cell (Ribi *et al.*, 1959). Cells stored at -70°C were slowly thawed in ice water and the suspension poured into a chilled Aminco French pressure cell (40 ml capacity, 1 in diameter piston). Passing the suspension through the French pressure cell three times using a hydraulic Aminco French pressure cell press operating at 3500 lb/in^2 lysed the cells. To reduce the viscosity of the lysate, it was sonicated at 4°C for 5 min using a Branson 250 Sonifier equipped with a microtip. Pulse duration was 0.5 per second and the power output of the tip was 60 watts. Unbroken cells, outer membrane and peptidoglycan components were pelleted by centrifugation at 7 800 × g for 45 min at 4°C in a Sorvall RC-5B Plus centrifuge using a chilled Sorvall SS-34 fixed angle rotor. The supernatant was passed through a 0.8 μm syringe filter to remove unpelleted cell wall material and applied to a column of Ni-NTA agarose (Qiagen) equilibrated with binding buffer (5 mM imidazole, 50 mM HEPES pH 7.5, 500 mM NaCl, 5% (v/v) glycerol) at a rate of 1.5 ml/min using a Pharmacia LKB Pump P-1 peristaltic pump. The column was then washed with 110 ml of 30 mM imidazole, 50 mM HEPES pH 7.5, 500 mM NaCl, 5% (v/v) glycerol and the His-tagged PBP eluted using 250 mM imidazole, 50 mM HEPES pH 7.5, 500 mM NaCl, 5% (v/v) glycerol. Immediately upon collection of fractions, EDTA was added to a concentration of 1 mM to bind any Ni^{2+} ions that leached off the column. Five minutes later DTT was added to a concentration of 0.15 mM. SDS-PAGE analysis of the eluted proteins using

the discontinuous buffer system of Laemmli (1970) revealed a prominent 29 kDa band (data not shown), matching the predicted size of 6×His-tagged PBP. This band was estimated to comprise over 90% of the eluted protein. The identity of 6×His-tagged PBP was confirmed using electrospray ionization mass spectrometry (ESI-MS, experimental mass 28,574, predicted mass 28,566).

2.2.10 Metal affinity pull-down experiment using PBP and GABPα PNT$^{(168-254)}$

GABPα PNT$^{(168-254)}$ (Figure 2.6, expressed and purified as described in sections 3.2.1 and 3.2.2, respectively) and 6×His-tagged PBP were dialyzed into buffer A (20 mM Na_2HPO_4, 20 mM NaCl, 1 mM DTT, pH 7.2). The proteins were incubated with Talon metal affinity resin (Clontech), the resin pelleted and the supernatant removed. The resin was washed with buffer A, and finally the supernatant, wash and resin fractions were analyzed by SDS-PAGE. As a control, GABPα PNT$^{(168-254)}$ alone was incubated with the Talon resin.

2.2.11 Cleavage of PBP 6×His-tag

One mg of lyophilized thrombin was added to 3 ml of PBP in 250 mM imidazole, 50 mM HEPES pH 7.5, 500 mM NaCl, 5% (v/v) glycerol. This solution was dialyzed overnight at 4°C against 4 litres of Tris buffered saline with 2.5 mM $CaCl_2$. Following dialysis, the solution was incubated with 200 µl of *p*-aminobenzamidine agarose bead slurry (Sigma) for 45 min at 4°C to remove the thrombin. Following this incubation, the beads were pelleted by centrifugation and the supernatant incubated with 200 µl of Talon metal affinity resin (Clontech) for 45 min at 4°C to remove the cleaved 6×His-tag. The

resin was pelleted by centrifugation leaving cleaved PBP, with only three extra amino acid residues at the amino-terminus, in the supernatant (see Figure 2.5 for a map of the pET 28a multiple cloning site and thrombin cleavage site). Removal of the 6×His-tag was confirmed using ESI-MS (experimental mass 26 821, predicted mass 26 815).

2.2.12 Native Gel Electrophoresis

Native gel electrophoresis was conducted using a continuous buffer system consisting of the separating buffer system of Laemmli (1970) with SDS omitted.

2.2.13 Circular Dichroism Spectroscopy

CD spectra were recorded on a Jasco J-720 spectropolarimeter using a cylindrical, water-jacket quartz cell. Parameters for wavelength scan were 190-300 nm range, 20 mdeg sensitivity, 0.1 nm resolution, accumulation 3, 1.0 nm bandwidth, 4 s response and 50 nm/min scan rate.

2.3 Results and discussion

2.3.1 Yeast two hybrid screens and discovery of "PNT Binding Protein" (PBP)

Isabelle Pot, a former technician in the McIntosh laboratory, performed a yeast two-hybrid screen using a GABPα PNT domain construct consisting of amino acids 138-254 from the full length protein (GABPα $PNT^{(138-254)}$, see Figure 2.6) with a mouse embryo library. Three different, but overlapping fragments of a single protein of unknown function were found to interact with this PNT domain construct. To demonstrate a direct interaction between the proteins *in vitro*, the region common to all

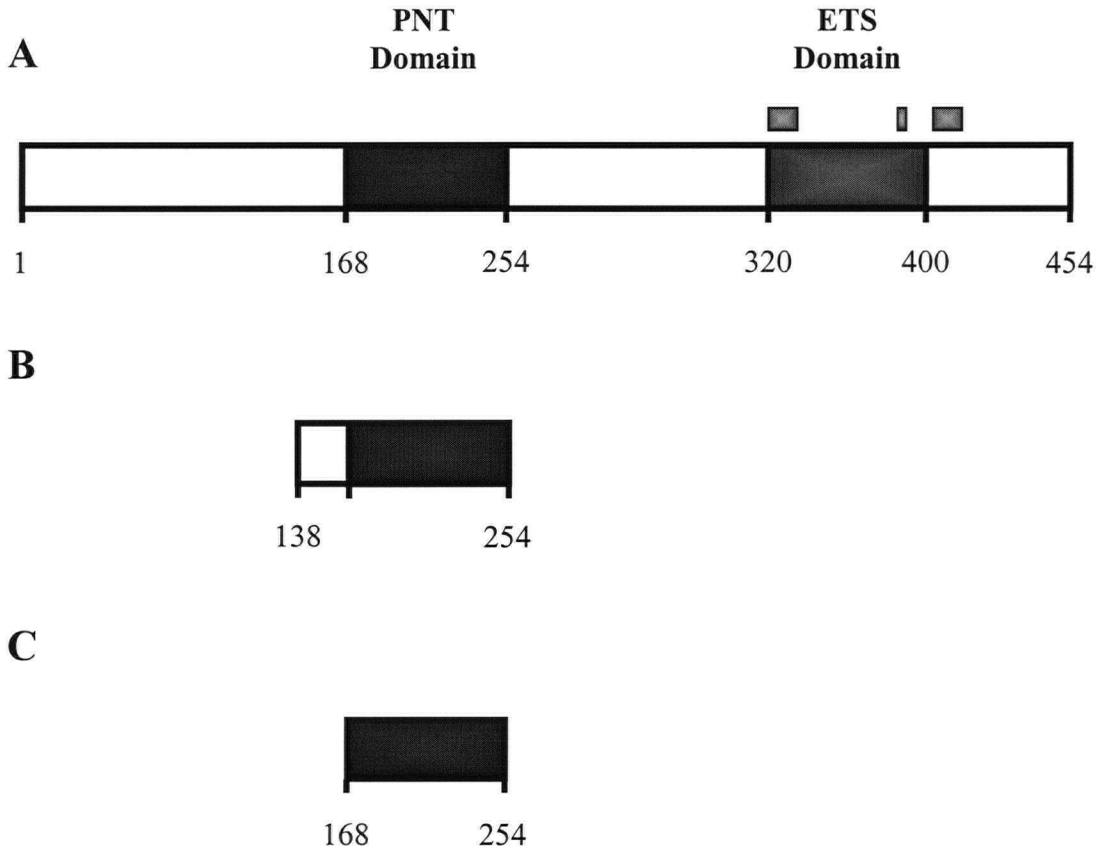


Figure 2.6. GABPα PNT domain constructs. **A)** Schematic diagram of the full-length GABPα protein showing the locations of the PNT and ETS domains defined through sequence homology and structural studies. Residues at the GABPβ interface have boxes above them. **B)** A construct consisting of the minimal PNT domain plus 30 unstructured amino terminal residues. **C)** The minimal PNT domain.

three fragments was cloned and expressed in *E. coli* BL21(λDE3) cells. However, the resonances in the ^{1}H NMR spectrum of the protein were clustered in the region expected for a random coil conformation, suggesting the absence of a well-defined tertiary structure. In addition, attempts to demonstrate an interaction between the protein and GABPα PNT$^{(168\text{-}254)}$, the minimal GABPα PNT domain (see Figure 2.6), were unsuccessful. As a result, work on this protein was not pursued any further.

A second two-hybrid screen was then initiated using the minimal GABPα PNT domain construct, GABPα PNT$^{(168\text{-}254)}$, as bait against a HeLa cell library. One clone that interacted with this domain was identified, but due to time restrictions the screen was not completed. Upon commencement of this thesis work the screen was continued and a second independent clone identified which encoded the same 226 amino acid "Pointed Binding Protein" (PBP). The results of the screen with controls appear in Figure 2.7.

2.3.2 Sequence and Homology Analyses of PBP

The sequence of the PBP cDNA appears in Figure 2.8. Eukaryotic mRNAs have consensus sequences, termed Kozak sequences, surrounding the initiating methionine codon that are required for efficient initiation of translation (Kozak, 1986). A Kozak sequence was present at the initiating methionine of PBP with A in the –3 position and G in the +4 position. Two forms of PBP were identified (approximately 900 and 1400 bp in length) which appear to be alternatively spliced variants with different 3'-untranslated regions. A putative non-canonical polyadenylation signal (Zhao *et al.*, 1999) was also present. These sequences are often found in alternatively spliced mRNAs. Examination of the 3'-untranslated region of the 1400 bp transcript reveals a very high A/T content

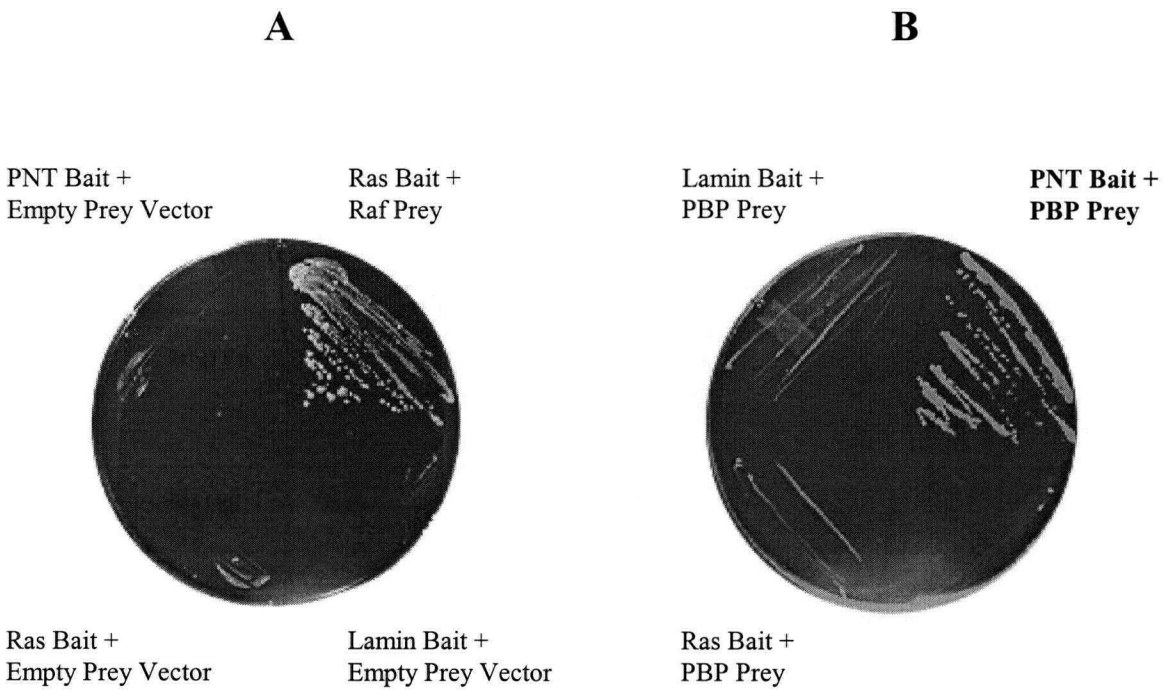


Figure 2.7. Use of a yeast two-hybrid system to identify proteins that interact with the PNT domain of GABPα. **A)** A control plate showing a Ras/Raf positive interaction and negative controls with empty prey vectors. **B)** The positive interaction between the GABPα PNT domain bait and PBP prey.

A

```
AATTCGGCACGAGCTACTGCAAGGCAAAAGCCGGAGTGGACGTGTCTTTTGAAACTGCTGCTCTTTCACTTCTCA
                          |
GGCGTCACCGAGAGCTCAGCACCCAGGCTGAACTCTGTACCATTTGGAAGAATGGAAGCTGATGCATCTGTTGAC
                                                  M  E  A  D  A  S  V  D

ATGTTTTCCAAAGTCCTGGAGCATCAGCTGCTTCAGACTACCAAACTGGTGGAAGAACATTTGGATTCTGAAATT
 M  F  S  K  V  L  E  H  Q  L  L  Q  T  T  K  L  V  E  E  H  L  D  S  E  I

CAAAAACTGGATCAGATGGATGAGGATGAATTGGAACGCCTTAAAGAAAAGAGACTCCAGGCACTAAGGAAAGCT
 Q  K  L  D  Q  M  D  E  D  E  L  E  R  L  K  E  K  R  L  Q  A  L  R  K  A
                |
CAACAGCAGAAACAAGAATGGCTTTCTAAAGGACATGGGGAATACAGAGAAATCCCTAGTGAAAGAGACTTTTTT
 Q  Q  Q  K  Q  E  W  L  S  K  G  H  G  E  Y  R  E  I  P  S  E  R  D  F  F
                                                                  |
CAAGAAGTCAAGGAGAGTGAAAATGTGGTTTGCCATTTCTACAGAGACTCCACATTCAGGTGTAAAATACTAGAC
 Q  E  V  K  E  S  E  N  V  V  C  H  F  Y  R  D  S  T  F  R  C  K  I  L  D

AGACATCTGGCAATATTGTCCAAGAAACACCTCGAGACCAAATTTTTGAAGCTGAATGTGGAAAAAGCACCTTTC
 R  H  L  A  I  L  S  K  K  H  L  E  T  K  F  L  K  L  N  V  E  K  A  P  F

CTTTGTGAGAGACTGCATATCAAAGTCATTCCCACACTAGCACTGCTAAAAGATGGGAAAACACAAGATTATGTT
 L  C  E  R  L  H  I  K  V  I  P  T  L  A  L  L  K  D  G  K  T  Q  D  Y  V

GTTGGGTTTACTGACCTAGGAAATACAGATGACTTCACCACAGAAACTTTAGAATGGAGGCTCGGTTCTTCTGAC
 V  G  F  T  D  L  G  N  T  D  D  F  T  T  E  T  L  E  W  R  L  G  S  S  D
                |
ATTCTTAATTACAGTGGAAATTTAATGGAGCCACCATTTCAGAACCAAAAGAAATTTGGAACAAACTTCACAAAG
 I  L  N  Y  S  G  N  L  M  E  P  P  F  Q  N  Q  K  K  F  G  T  N  F  T  K

CTGGAAAAGAAAACTATCCGAGGAAAGAAATATGATTCAGACTCTGATGATGATTAG
 L  E  K  K  T  I  R  G  K  K  Y  D  S  D  S  D  D  D  *
```

B

```
AGCTCAATAATTCTTTGTAAATTGTCTTTTTTTTTCTGCTTCAGATTTAAATGTGTTTTTAAAATTCTATTAATG
TCTA||ATACATTGGTCACCTAAATACTCATATTCTCGAGTTTTATACAGTTGTATCACATCGAAAAGTGTCTTT
ACTGTTTTCTGTGTGGCCATCATGTTTAAGTTGAGGAAAACTCAGTTCTTAAATTATCTGGGAAGGGTCTGGATT
CTCTATTTTTGAGATTGACTTTATCACAATATGATTCTTACATCTTTATACCATTTACAATTGTGTTTTAGATCT
ACAGAGTTAGAAATTCGAAAACTATTCCAGGACTAATTCTTAATCGGCATTATTTATACAAGAGGTCAAGTAACA
TTTACTAGCGCAATACTGCACTTGTAAATGAATTATAAACGCTCTTCTGGAATATATTTAAATAACCATTAAAGA
ACTGCTTATTCATTCTGGACACTGCATGTTGATGTTGAATCAACTGATGCCAGCAGAAAGCTATTTTGATTTGTG
AACATACTGCCTTATTTAAAGGGTCCTGATTGCTTGTATTTTAAGACATTCATTAAAAAGAAACCAGGAAACACT
TTTGAAATAACAGCATAAGGAACTTCACTGTCTCTGCTCAATAAAATACCTGTAACTAAAAAAAAAAAAAAAAAA
```

Figure 2.8. Nucleotide and predicted amino acid sequence of PBP cDNA clones. **A)** cDNA sequence of the coding region of PBP and predicted amino acid sequence given in standard one-letter code. The Kozak sequence is underlined. Exon boundaries are marked with | above the sequence. **B)** cDNA sequence of the 3'-untranslated region of PBP. The point at which the poly(A)$^+$ tail of the 900 bp transcript begins is marked with ||. Putative ATTTA RNA destabilizing elements are highlighted in grey.

(67%) and the presence of five ATTTA pentanucleotide sequences. These sequences are found in the 3'-untranslated regions of many highly unstable mRNAs (Chen & Shyu, 1994). Thus, alternative splicing of PBP transcripts could be one mechanism to control expression of the PBP protein. This is supported by the fact that human phosducin-like protein, a homologue of PBP (see below), contains similar features. Phosducin-like protein transcripts have alternate 3'-untranslated regions with the long form containing instability elements (Lazarov *et al.*, 1999).

The cDNA sequence of PBP matched that of an entry in the Genbank sequence database (Genbank protein ID 13543639). This sequence was a direct submission and encoded a putative ATP binding protein of unknown function associated with cell differentiation. This last point was intriguing because it suggested that PBP may play a role in gene expression, which would be expected for a protein that interacted with the PNT domain of an Ets transcription factor. A putative PBP isoform was also found that differed in only two amino acids from PBP. Several putative homologues of PBP from a variety of eukaryotes are listed in Table 2.1. In addition, the sequence of the PBP cDNA matched those of expressed sequence tags from embryonic tissues and from several human and mouse cancer cell lines. A BLAST search of the human genomic database revealed that PBP was located on chromosome 2 (2p24.3-p24.1) and consisted of five exons, the boundaries of which are shown in Figure 2.8.

The majority of these GenBank entries were direct submissions and corresponded to putative proteins of unknown function. However, all these proteins have sequence similarity to phosducin, a protein best known for its role in regulating the

Organism	Amino Acid Identity to PBP†	Genbank Protein ID
Mus musculus	89%	12833213
Drosophila melanogaster	53%	7299373
Arabidopsis thaliana	46%	5902677
Arabidopsis thaliana	43%	4835247
Leishmania major	42%	5852127
Caenorhabditis elegans	41%	485090
Arabidopsis thaliana	37%	10177522
Mus musculus	33%	13937367
Saccharomyces cerevisiae	31%	6320389
Danio rerio	28%	14280517
Homo sapiens - phosducin	24%	130134

Table 2.1. Some putative PBP homologues. This is not a comprehensive list but is intended to illustrate the presence of PBP homologues in a variety of organisms. †Amino acid identities were calculated using the BLAST2 Program (Tatusova & Madden, 1999).

phototransduction cascade through binding to the $G_t\beta\chi$ subunit. In recent years, several phosducin homologues have been discovered and several studies suggest that the phosducin family may have functions other than G-protein regulation. Specifically, phosducin and phosducin-like orphan protein 1 have been found to bind the cone-rod homeobox transcription factor and alter its transactivation activity (Zhu & Craft, 2000). This raises the possibility that members of the phosducin family may be involved in the regulation of gene expression. In addition, a recent study on the *S. cerevisiae* homologue of PBP demonstrated that this protein was indeed a phosducin homologue (Flanary *et al.*, 2000).

To obtain additional information on PBP a variety of sequence analysis tools were used. Information determining the subcellular localization of a protein is frequently encoded in its amino acid sequence as a short protein sorting signal. Prediction of subcellular localization sites can suggest a function for an uncharacterized protein (Nakai, 2000). The PSORT II algorithm (Nakai & Kanehisa, 1992) is the only publicly available program that makes use of both known sorting signals and amino acid composition to predict the subcellular location of a protein. Analysis of the PBP amino acid sequence using PSORT II did not reveal the presence of a signal sequence, transmembrane regions or organelle targeting sequences, indicating that PBP was likely a soluble, cytosolic protein.

Analysis of PBP using Lupas' NEWCOILS algorithm (Lupas *et al.*, 1991) resulted in a high probability (greater than 90%) that an amino-terminal region in PBP had the potential to form coiled-coils. However, Berger's more recent PAIRCOILS algorithm (Berger *et al.*, 1995) predicted a much lower probability (less than 40%).

Proteins that score well with NEWCOILS but low with PAIRCOILS tend to be false positives, although most do contain an amphipathic α-helix (Berger *et al.*, 1995). Thus, PBP is unlikely to contain a region that can form coiled-coils but may possess an amphipathic α-helix, which could play a role in mediating protein interactions.

Motif searches revealed several putative phosphorylation sites for protein kinase C and casein kinase II, and a P-loop, which is part of an ATP/GTP binding motif. However, the following observations make it unlikely that PBP binds ATP or GTP:

1 – The P-loop consensus sequence is known to pick up a large number of false positives in database screening (Saraste *et al.*, 1990).

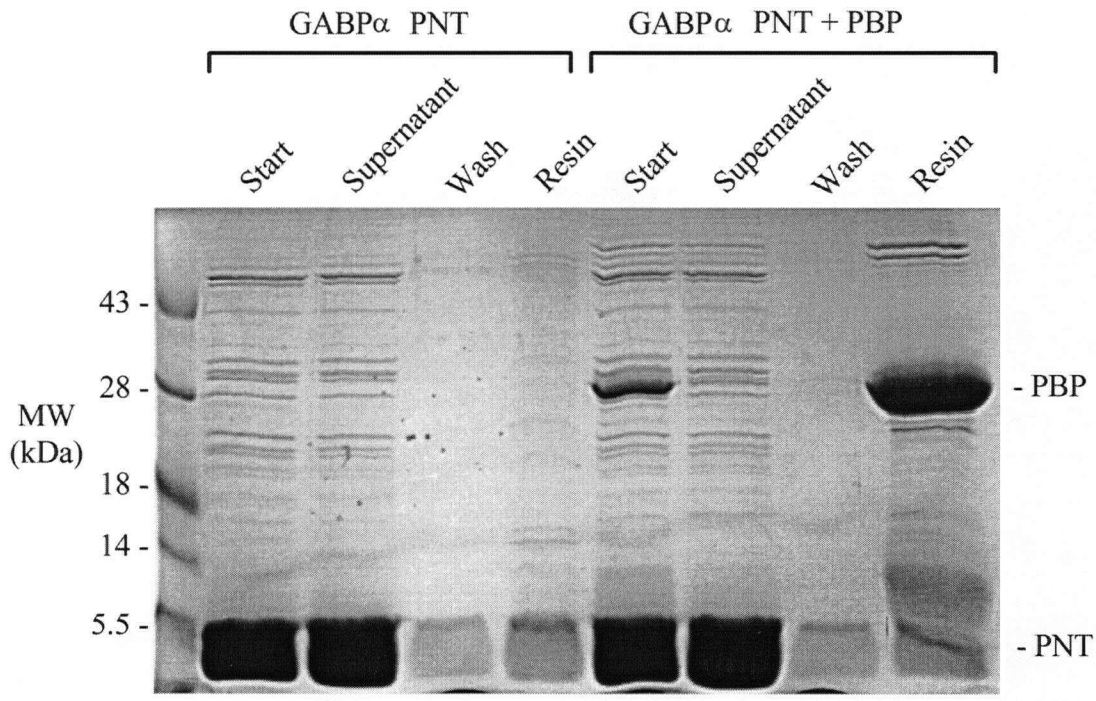
2 – The location of the P-loop in PBP is inconsistent with that of most ATP/GTP binding proteins. In addition, other motifs important for ATP/GTP binding are not present (Traut, 1994).

3 – The P-loop is not conserved in the putative PBP homologues.

Thus, although these sequence analyses suggest that PBP is a phosducin homologue and provide some insights, the exact function of this protein remains enigmatic.

2.3.3 Metal affinity pull-down experiments using GABPα PNT$^{(168-254)}$ and PBP

An attempt was then made to reproduce the yeast two-hybrid interaction *in vitro* using a metal affinity "pull-down" experiment. The results of one such experiment with 6×His-tagged PBP and untagged GABPα PNT$^{(168-254)}$ appear in Figure 2.9. As the control experiment using only GABPα PNT$^{(168-254)}$ demonstrated, there was little non-specific binding of GABPα PNT$^{(168-254)}$ to the resin as most of the protein was found in the



.

Figure 2.9. Metal affinity pull-down experiments fail to demonstrate a direct interaction between 6×His-tagged PBP and GABPα $PNT^{(168-254)}$. SDS-PAGE analysis shows no significant GABPα $PNT^{(168-254)}$ binding to the resin which contains PBP.

supernatant. When both proteins were incubated with the resin, again, very little GABPα $PNT^{(168-254)}$ was present in the resin fraction. The small amount that was present was likely due to non-specific binding as a similar level was found in the control resin fraction. Most GABPα $PNT^{(168-254)}$ remained in the supernatant suggesting that there was no direct interaction between PBP and GABP α $PNT^{(168-254)}$ under the conditions of this experiment. Similar results were obtained when the experiment was repeated using various buffer conditions, concentrations of PBP and GABPα $PNT^{(168-254)}$, and incubation times.

2.3.4 Native Gel Electrophoresis using GABPα $PNT^{(168-254)}$ and PBP

One concern with the "pull-down" experiment was the possibility that binding of the amino-terminal 6×His-tag to the resin prevented interaction of GABPα $PNT^{(168-254)}$ with PBP due to steric constraints. To investigate this possibility, native gel electrophoresis of GABPα $PNT^{(168-254)}$ and PBP was conducted. Both GABPα $PNT^{(168-254)}$ and PBP have low isoelectric points and will migrate in the same direction on a native gel. If these proteins form a stable complex, the complex would be expected to migrate with a different mobility than the individual proteins. As can be seen from Figure 2.10, the electrophoretic mobilities of GABPα $PNT^{(168-254)}$ and PBP are the same, whether separate or mixed together. This suggests that the proteins are not forming a stable complex under these conditions.

To eliminate the possibility that the 6×His-tag was interfering with the interaction between PBP and GABPα $PNT^{(168-254)}$, the tag was removed using thrombin as described in section 2.2.11. Native gel electrophoresis was repeated using this untagged form of

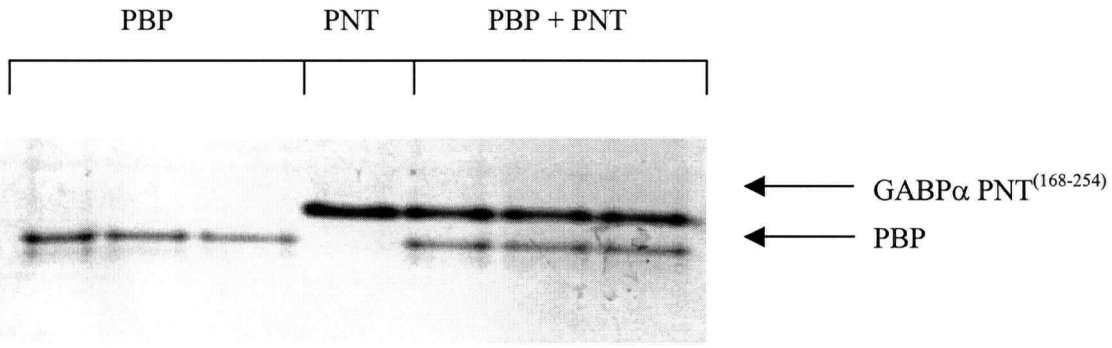


Figure 2.10. Native gel electrophoresis failed to demonstrate a direct interaction between PBP and GABPα $PNT^{(168-254)}$. The electrophoretic mobilities of both proteins were unaltered whether separate or mixed.

PBP and GABPα $PNT^{(168-254)}$. Again, the electroporetic mobilities of the proteins were the same whether separate or mixed suggesting that the proteins are not interacting under these conditions.

2.3.5 Circular Dichroism (CD) Studies of PBP Stability

One possible explanation for these negative results is that PBP, expressed in *E. coli*, may not have been properly folded in its native conformation during these experiments. Evidence of PBP instability was obtained when concentrated protein was being prepared for the "pull-down" experiments described in section 2.4.1. PBP appeared to have a solubility limit of ~40 μM in Buffer A, which is about 10-fold lower than the concentration required for NMR studies. It was suspected that PBP may have been unstable but still soluble at concentrations below 40 μM, but then aggregated at higher concentrations. To investigate the possibility that PBP was unstable, CD spectroscopy was used. PBP was dialyzed into Buffer A (20 mM Na_2HPO_4, 20 mM NaCl, pH 7.2), and stored at 4°C. Aliquots were then periodically removed and analyzed by far-UV CD spectroscopy. As can be seen from Figure 2.11, PBP appears to be unstable in Buffer A. From the initial spectrum (1 day), it was apparent that PBP had a significant amount of α-helical secondary structure as evidenced by the two peaks of negative ellipticity at 207 and 220 nm. The protein also cooperatively denatured at a midpoint temperature of 52°C (data not shown). These data indicate that initially, there was at least some secondary structure present in PBP. However, it is important to point out that the presence of secondary structure does not necessarily indicate that a protein is in its native tertiary structure. Over time the intensity of the spectra decreased and a "smoothing" of the

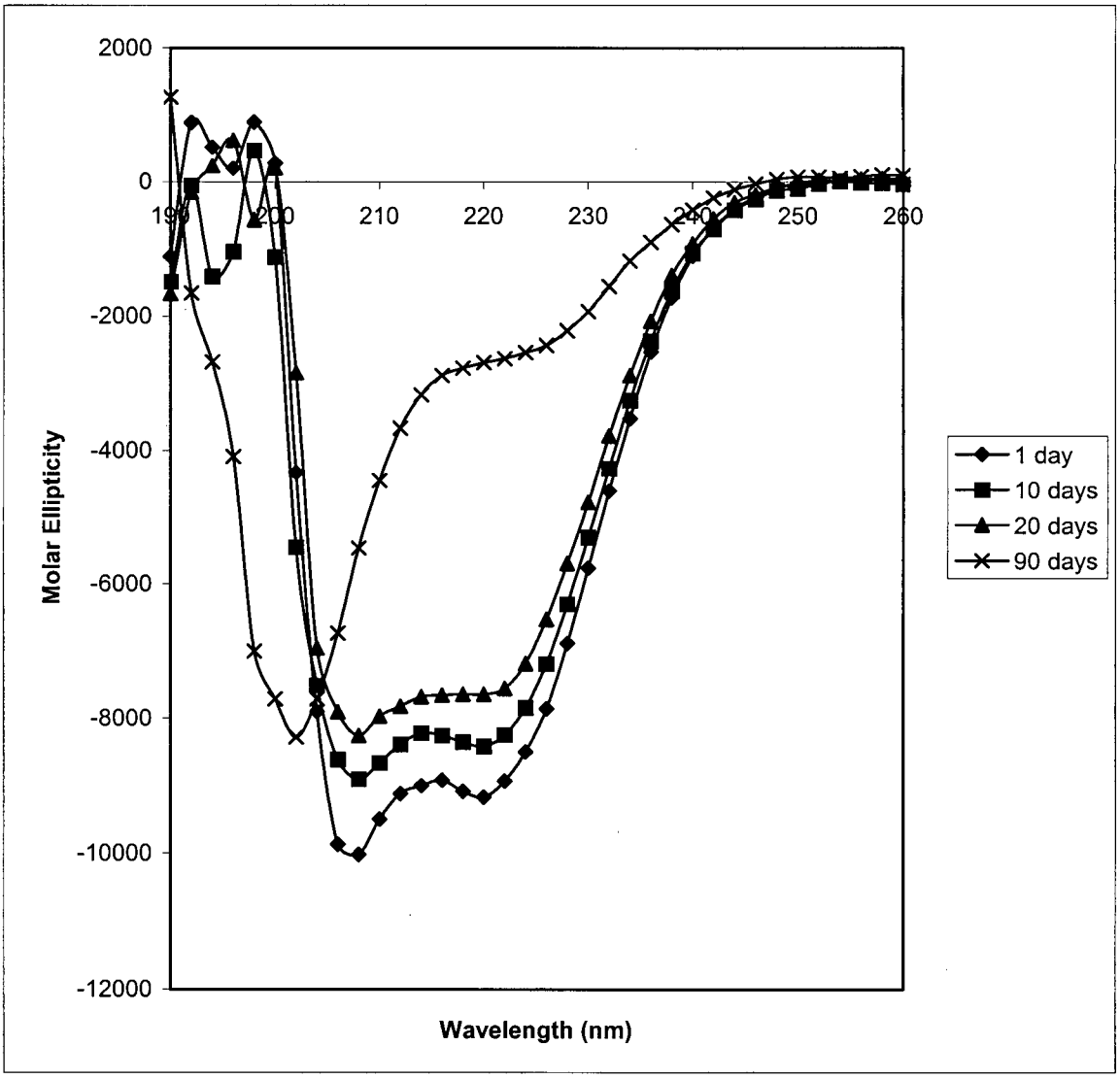


Figure 2.11. The secondary structure of PBP changed over time. CD spectra of PBP in 20 mM Na_2HPO_4, 20 mM NaCl, pH 7.2 recorded over time.

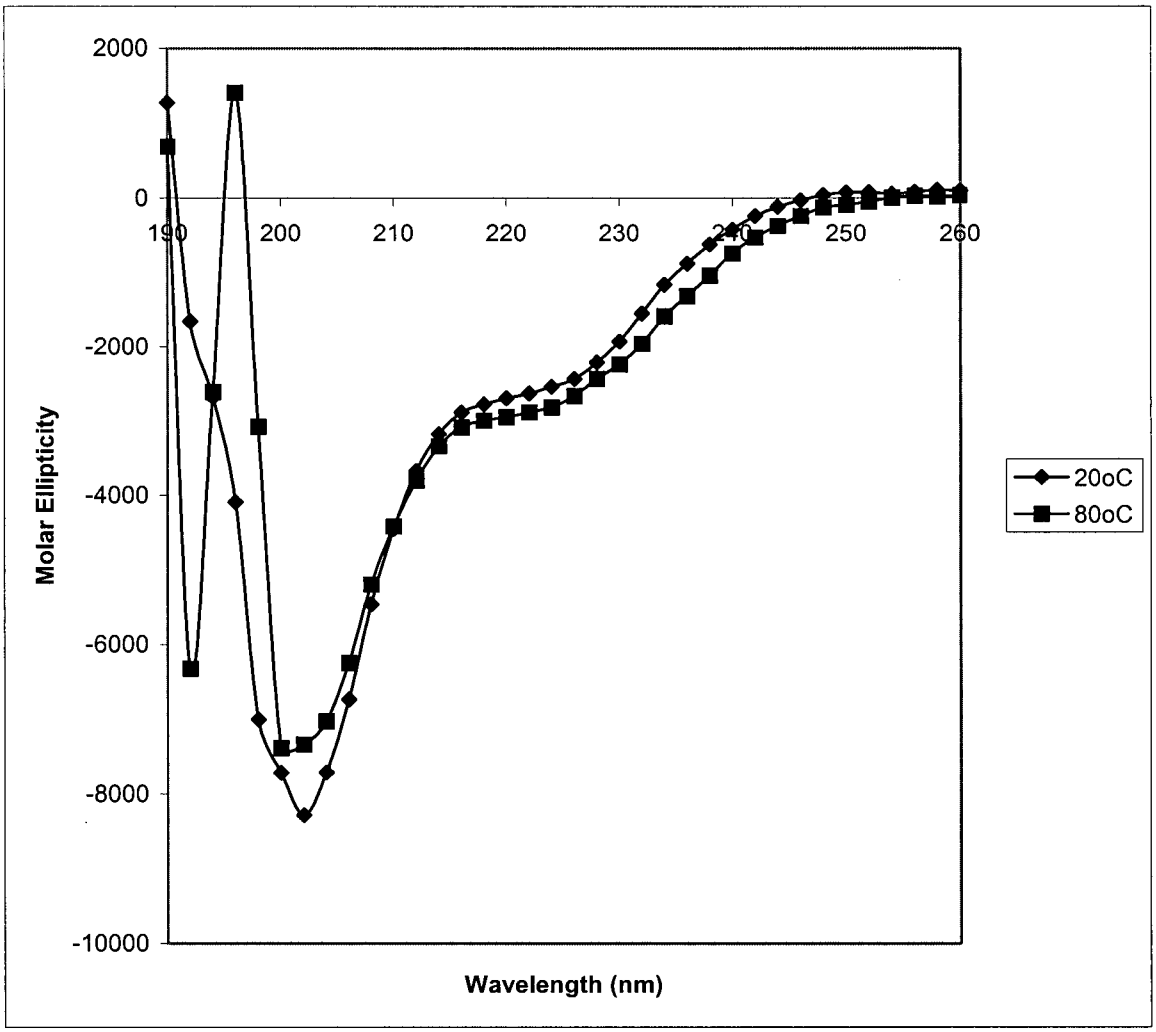


Figure 2.12. After 3 months of storage, PBP was completely denatured. The CD spectrum of PBP after 3 months, recorded at 20°C was similar to that of a heat-denatured sample of fresh PBP recorded at 80°C.

second peak at 220 nm occurred. These changes suggest that a gradual loss of the α-helical secondary structure occurred. Although the decrease in intensity is partly explained by the change in the concentration of the protein as it precipitated (see below), the shape of the spectra changed as well. This suggests that there was indeed a change in the secondary structure of PBP. Eventually, after 90 days at 4°C, the spectrum of PBP had changed dramatically and resembled that of a random coil. Indeed, at this stage the protein did not thermally melt at all (data not shown) and the spectrum of the protein at 20°C was virtually identical to that of a fresh sample of PBP denatured at 80°C, as shown in Figure 2.12. In addition, after a storage period of about 30 days, a white precipitate gradually appeared in the PBP sample. Collectively, these data strongly suggest that PBP was unstable under these conditions. PBP stability was also investigated using other buffer conditions including Phosphate Buffered Saline (PBS) and Tris Buffered Saline (TBS). In both cases, similar changes in the CD spectra over time were observed.

Although the spectral changes occurred over a period of weeks while all binding experiments were conducted with fresh protein, this data does suggest that PBP was unstable and raises the possibility that PBP, while possessing α-helical secondary structure, may not have folded into its native, tertiary structure under these conditions. These stability problems with PBP, along with the solubility limit mentioned above, would make a structure determination very difficult. Since there is no reference spectrum of native PBP for comparison with the initial spectrum, it is difficult to determine if PBP was initially in a native conformation and gradually unfolded and aggregated or if PBP was never properly folded at all. Thus, the negative results obtained in the above *in vitro*

interaction experiments may be due to the fact that PBP was not folded properly under the experimental conditions.

2.4 Conclusions and summary

A yeast two-hybrid screen was performed in an attempt to identify proteins that interacted with the PNT domain of GABPα. One potential protein partner was identified and was called PBP for PNT-Binding Protein. This protein was expressed in *E. coli*, but attempts to reproduce the two-hybrid interaction *in vitro* using metal affinity pull-down experiments and native gel electrophoresis were unsuccessful. CD experiments suggested that one possible explanation for these negative results is that PBP was not properly folded in *E. coli* and not in its native conformation.

There are also other possible explanations for the negative result. A post-translational modification such as phosphorylation or glycosylation may be necessary for the interaction. Since these proteins were expressed in *E. coli*, these modifications would not be present. Alternatively, since there is a PBP homologue in *S. cerevisiae*, the two-hybrid interaction may be mediated by an endogenous yeast protein. Another possibility is that the interaction between GABPα $PNT^{(168-254)}$ and PBP may be transient and thus not amenable to study using the above methods. It is known that the yeast two-hybrid system is a more sensitive assay of protein interactions than "pull-downs" and other similar methods.

Chapter 3

Investigation of Possible GABPα PNT Multimerization

As discussed in section 1.2, the leucine zipper motif of GABPβ has been found to mediate association of GABP α/β heterodimers. Previous work in the McIntosh laboratory, described below in section 3.1, has raised the possibility that the PNT domain of GABPα may self-associate as well. This interaction could make an important contribution to GABP α/β heterodimer association, perhaps by acting as a control point to help regulate subunit binding. The work described in this chapter was undertaken to investigate the possibility of GABPα PNT domain association in greater detail.

3.1 Background - Previous GABPα PNT Domain Work

Lisa Gentile, a previous post-doctoral fellow in the McIntosh laboratory, performed structural studies on the GABPα PNT domain prior to the beginning of this thesis work. Initial studies focused on a construct consisting of amino acids 138-254 of the full-length GABPα protein (GABPα PNT$^{(138\text{-}254)}$, see Figure 2.6), a region containing the PNT domain. Following expression of the protein in *E. coli* and spectral assignment, NMR ^{15}N relaxation analysis revealed that the first thirty amino acids were unstructured. In addition, mass spectrometry revealed that these residues were proteolyzed during storage, confirming the random coil structure of this region.

To determine the structure of the minimal PNT domain, these unstructured residues were removed to produce a different construct, GABPα PNT$^{(168\text{-}254)}$ (see Figure

2.6). However, in contrast to the longer GABPα $PNT^{(138-254)}$ construct which eluted at a single position during anion exchange FPLC at pH 8.5, GABPα $PNT^{(168-254)}$ eluted at two distinct positions, as shown in Figure 3.1. The form eluting at the lower salt concentration was called form A while the form eluting at the higher salt concentration was called form B. A ^{15}N-^{1}H HSQC spectrum of each form was collected and these are overlayed in Figure 3.2. Form A had a much simpler spectrum and contained the expected number of resonances for a protein of this size. Form B, however, had a more complex spectrum with many extra resonances present reflecting heterogeneity. To simplify determination of the structure of the domain, form A was chosen for study and its structure solved using NMR. The backbone structure of the protein appears in Figure 1.4. NMR relaxation analysis allowed calculation of an overall rotational correlation time of 6.7 ns, indicating that GABPα $PNT^{(168-254)}$ form A was a monomer in solution. Ultra-centrifugation studies also indicated the protein was monomeric (apparent molecular weight 11 690 Da, predicted molecular weight 10 320 Da).

Due to time restrictions, the nature of GABPα $PNT^{(168-254)}$ form B, and the relationship between the forms, remained unstudied. As described above, the PNT domain is believed to mediate protein-protein interactions and the existence of GABPα $PNT^{(168-254)}$ form B raised the possibility that the PNT domain of GABPα may self-associate. The following sections describe experiments conducted to investigate this possibility.

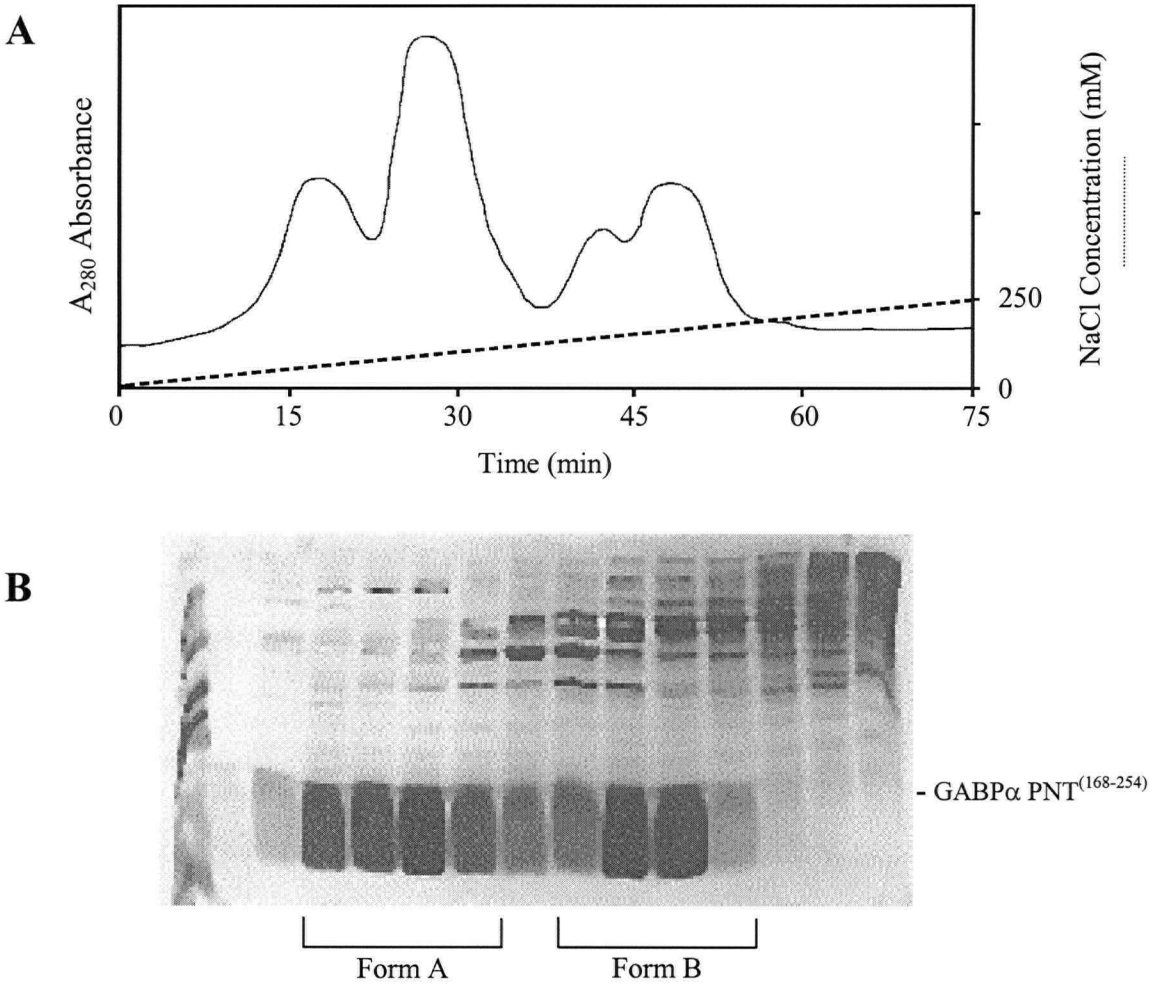


Figure 3.1. GABPα $PNT^{(168-254)}$ elutes at two distinct positions during anion exchange FPLC at pH 8.5. **A)** A_{280} trace of eluate from a Q Sepharose anion exchange column during purification of GABPα $PNT^{(168-254)}$ from *E. coli*. **B)** SDS-PAGE analysis of the corresponding eluate fractions.

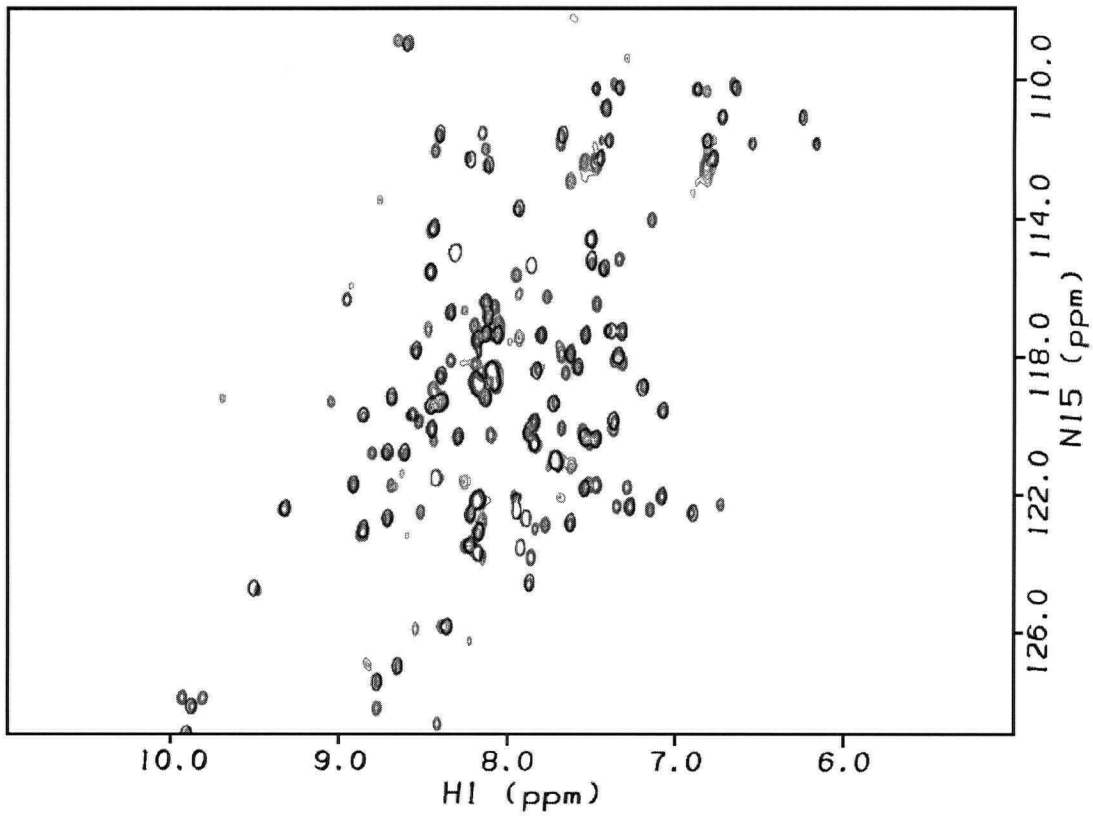


Figure 3.2. GABPα PNT$^{(168\text{-}254)}$ form A has a simpler ^{15}N-^{1}H HSQC spectra than form B. Overlay of ^{15}N-^{1}H HSQC spectra with form A in black and form B in grey.

3.2 Materials and methods

3.2.1 Expression of GABPα PNT$^{(168-254)}$

GABPα PNT$^{(168-254)}$ was expressed from a pET22b vector (Novagen) in *E. coli* BL21(λDE3) cells. For unlabelled protein, 25 ml of LB medium was inoculated with a single colony of *E. coli* BL21(λDE3) cells containing the pET22b-GABPα PNT$^{(168-254)}$ plasmid. For ^{15}N- and ^{15}N, ^{13}C-labelled protein, 25 ml of M9 minimal media was inoculated. This starter culture was incubated overnight at 30°C with constant shaking. In the morning, the culture was transferred to a 50 ml screw-capped tube and the cells pelleted by low-speed centrifugation. These cells were then resuspended in 1 ml of appropriate medium. For unlabelled protein, cells were diluted into 1 litre of LB medium. For ^{15}N-labelled protein, cells were diluted into 1 litre of M9 minimal medium containing 1 g of $^{15}NH_4Cl$ (Isotec). For ^{15}N, ^{13}C-labelled protein, cells were diluted into 1 litre of M9 minimal medium containing 1 g of $^{15}NH_4Cl$ and 3 g of $^{13}C_6$ D-glucose (Isotec). This culture was grown at 30°C with shaking until the OD_{600} reached 0.6-0.8. Expression of GABPα PNT$^{(168-254)}$ was then induced by adding IPTG to a concentration of 0.5 mM and incubation was continued at 30°C for 4h. The culture was then transferred to 250 ml centrifuge bottles and cells harvested by centrifugation at 1 700 × g for 20 min at 4°C in a Sorvall RC-5B Plus centrifuge using a chilled Sorvall GSA fixed angle rotor. Cells were resuspended in 20 ml of 50 mM Tris-HCl pH 8.5 with 10 mM DTT and 0.1 mM PMSF and were stored at -70°C.

3.2.2 Purification of GABPα PNT$^{(168-254)}$

GABPα PNT$^{(168-254)}$ was purified using a protocol developed by Lisa Gentile during previous work with the protein (unpublished results). Cell lysate was prepared using the procedure described for PBP in section 2.2.9. This lysate was applied at 4 ml/min to a Fast Flow Q Sepharose anion exchange column equilibrated with 50 mM Tris-HCl, pH 8.5. The column was then washed with 50 mM Tris-HCl, pH 8.5 at 4 ml/min until the UV absorbance of the eluate returned to baseline. GABPα PNT$^{(168-254)}$ was eluted using a linear NaCl gradient from 0-250 mM over a period of 75 min at a flow rate of 4 ml/min.

3.2.3 Site-Directed Mutagenesis of GABPα PNT$^{(168-254)}$

Mutagenesis of Cys 223 to Ala was carried out using the Quik Change Site-Directed Mutagenesis kit (Stratagene). The mutation was confirmed using DNA sequencing and the mutant protein (GABPα PNT$^{(168-254)}$ C223A) was expressed in *E. coli* BL21(λDE3) cells as described in section 3.2.1 and purified as described in section 3.2.2.

3.2.4 Glutataldedhyde Crosslinking of GABPα PNT$^{(168-254)}$

A 25 μM solution of GABPα PNT$^{(168-254)}$ in 100 mM phosphate buffer, pH 6.5 was incubated in the presence of 1 mM DTT at 4°C overnight to reduce any disulfide linkages present. In the morning, the solution was incubated with 0.25% glutaraldehyde for 45 min at room temperature. Reaction mixtures were then analyzed by SDS-PAGE.

3.2.5 Construction of a 6×His-tagged version of GABPα PNT$^{(168-254)}$

The coding region of GABPα PNT$^{(168-254)}$ was cut out of the pET22b vector using Nde I and Hind III restriction enzymes. A pET28a vector (Novagen) was also cleaved using these enzymes. The resulting fragments were resolved by agarose gel electrophoresis and the gel stained using ethidium bromide. Bands corresponding to the GABPα PNT$^{(168-254)}$ coding region and the linearized pET28a vector were excised and purified using the QIAEX II Gel Extraction Kit (Qiagen). Purified fragments were incubated together with T4 DNA ligase overnight at 16°C. Ligation reactions were transformed into $CaCl_2$-competent *E. coli* DH5α cells using the heat shock method. Colonies resulting from these transformations were used to grow overnight cultures from which plasmid DNA was extracted using the alkaline hydrolysis method.

Plasmids were subjected to PCR analysis using T7 promoter and T7 terminator primers to check the size of the DNA fragment inserted into the pET28a vector. Reactions were analyzed by agarose gel electrophoresis and a band of the size expected for a 6×His-tagged version of GABPα PNT$^{(168-254)}$ was obtained. Plasmid DNA was then electroporated into *E. coli* BL21(λDE3) cells.

3.2.6 Expression and Purification of 6×His-tagged GABPα PNT$^{(168-254)}$

6×His-tagged GABPα PNT$^{(168-254)}$ was expressed and purified as described above for GABPα PNT$^{(168-254)}$ (sections 3.2.1 and 3.2.2).

3.2.7 *Cleavage of His-tagged GABPα PNT$^{(168-254)}$*

The 6×His-tag was removed from GABPα PNT$^{(168-254)}$ using the procedure described in section 2.2.11.

3.2.8 *Preparation of NMR samples*

Following purification, protein samples were placed in 2 000 nominal molecular weight dialysis tubing and were dialyzed into 20 mM Na_2HPO_4, 20 mM NaCl, pH 7.2 at 4°C. Following dialysis, samples were concentrated to a volume of ~600 µl using a Centricon concentration device with a 3 000 nominal molecular weight cut-off filter. The concentration of the samples were then determined using absorbance at 280 nm. DTT was added to a concentration of 10 mM and D_2O to a concentration of 7.5% (v/v).

3.2.9 *NMR spectroscopy*

NMR spectra were acquired at 30°C with a Varian Inova 600 MHz spectrometer equipped with a triple resonance probe and a pulsed field gradient accessory. Data were processed using Felix 2000 (Molecular Simulations Inc). Resonances from nuclei in the backbone of GABPα PNT were assigned using a combination of HNCACB (Grzesiek & Bax, 1992b) and CBCA(CO)NH (Grzesiek & Bax, 1992a) triple resonance experiments. ^{15}N T_1, T_2 and heteronuclear NOE relaxation data were recorded and analyzed as previously described (Farrow *et al.*, 1994). T_1 time points were: 11.1, 33.3, 88.8, 155.4, 222.0, 388.5, 499.5, 666.0 and 943.5 ms. T_2 time points were: 16.8, 33.7, 50.5, 67.3, 84.2, 101.0, 117.8, 134.7 and 151.5 ms. For NOE values, spectra were acquired with and

without 3 s of ^{1}H saturation with a total recycle delay of 5 s. Intensities of 25 resonances in the spectra were measured using Felix 2000.

3.3 Results and discussion

3.3.1 Confirmation of the existence of GABPα PNT$^{(168-254)}$ forms A and B

A first step in the investigation of the nature of GABPα PNT$^{(168-254)}$ form B was reproduction of the results obtained by other workers as described in section 3.1. GABPα PNT$^{(168-254)}$ was expressed and purified as described in sections 3.2.1 and 3.2.2 using procedures identical to those used by previous workers. As expected, the GABPα PNT$^{(168-254)}$ elution profile was identical to that obtained previously as shown in Figure 3.1. The form A elution maximum occurred at approximately 120 mM NaCl while form B eluted later with a maximum at approximately 190 mM NaCl.

3.3.2 Investigation of possible GABPα PNT$^{(168-254)}$ covalent modification

To eliminate the possibility that a covalent modification of GABPα PNT$^{(168-254)}$ was responsible for the altered elution position of form B, samples of forms A and B were analyzed by ESI-MS. Both forms were found to have identical masses that matched the predicted mass of GABPα PNT$^{(168-254)}$ (experimental mass 10 323, predicted mass 10 320). The amino-terminal methionine had been removed from GABPα PNT$^{(168-254)}$ by *E. coli* proteases (Hirel *et al.*, 1989). This eliminated the possibility that GABPα PNT$^{(168-254)}$ form B contained a covalent modification that altered the mass of the protein. However, asparagine and glutamine residues are known to undergo nonenzymatic deamidation to form aspartate, isoaspartate or glutamate residues (Wright, 1991). Deamidation would

increase the negative charge on a protein and it would be expected to elute at a higher salt concentration during anion exchange FPLC. This matches the observed behaviour of GABPα $PNT^{(168-254)}$ form B. In addition, deamidation increases the mass of a protein by only 1 Da making these changes very difficult to detect by mass spectrometry. To determine if forms A and B of GABPα $PNT^{(168-254)}$ had different charges, native gel electrophoresis was used. Deamidation would be expected to alter the electrophoretic mobility of the protein and forms A and B should migrate differently on a native gel. When GABPα $PNT^{(168-254)}$ forms A and B were subjected to native gel electrophoresis, they had identical electrophoretic mobilities (data not shown). As a result, it is unlikely that GABPα $PNT^{(168-254)}$ form B arises due to deamidation.

3.3.3 Site-directed mutagenesis of GABPα $PNT^{(168-254)}$

Since GABPα $PNT^{(168-254)}$ form B did not appear to be a covalently modified form of GABPα $PNT^{(168-254)}$, self-association of the protein was investigated. GABPα $PNT^{(168-254)}$ contains a single cysteine residue, therefore, oxidation during purification may result in the formation of GABPα $PNT^{(168-254)}$ dimers held together by a disulfide linkage. Although DTT is included in the purification protocol, the disulfide linkage could be resistant to reduction. To eliminate this possibility, a point mutation was introduced into GABPα $PNT^{(168-254)}$ that replaced the single cysteine residue with an alanine residue (see section 3.2.3). This mutant was called GABPα $PNT^{(168-254)}$ C223A.

GABPα $PNT^{(168-254)}$ C223A was expressed and purified using a procedure identical to that used for purification of GABPα $PNT^{(168-254)}$. No change in the elution profile was observed. The masses of forms A and B were identical as determined by ESI-

MS and matched the predicted mass of the mutant protein (experimental mass 10 290, predicted mass 10 288). As a result, GABPα PNT$^{(168-254)}$ form B is not a disulfide linked dimer.

3.3.4 Glutaraldehyde crosslinking of GABPα PNT$^{(168-254)}$

In an attempt to demonstrate GABPα PNT$^{(168-254)}$ association, glutaraldehyde crosslinking was utilized. Chemical crosslinking has some advantages over other methods such as analytical ultra-centrifugation or gel filtration in that it can "trap" transient protein complexes. Although these complexes are weak, they may nevertheless be significant in a physiological setting. GABPα PNT$^{(168-254)}$ form A was incubated in the presence of glutaraldehyde as described in section 3.2.4 and the reaction mixtures analyzed by SDS-PAGE. The results of this analysis can be seen in Figure 3.3. With glutatraldehyde treatment, there is a band present at the mass expected for a dimer of GABPα PNT$^{(168-254)}$ that is not present in the absence of glutaraldehyde. The PNT domains of several other Ets proteins are included for comparative purposes and it can be seen that crosslinking is only seen with GABPα PNT$^{(168-254)}$ and the PNT domain from Pnt P2. It was estimated that less than 10% of the GABPα PNT$^{(168-254)}$ protein crosslinked to form dimers. However, Ste4 and Byr2 are two SAM domains known to interact (see section 1.4.7), and show 100% crosslinking under these conditions such that no monomer remains in the reaction mixture (data not shown). This suggests that the GABPα PNT$^{(168-254)}$ monomers may associate weakly compared to the Ste4-Byr2 interaction. Even a weakly associating protein complex may be resolvable during purification by anion exchange FPLC. To further investigate, crosslinking was conducted

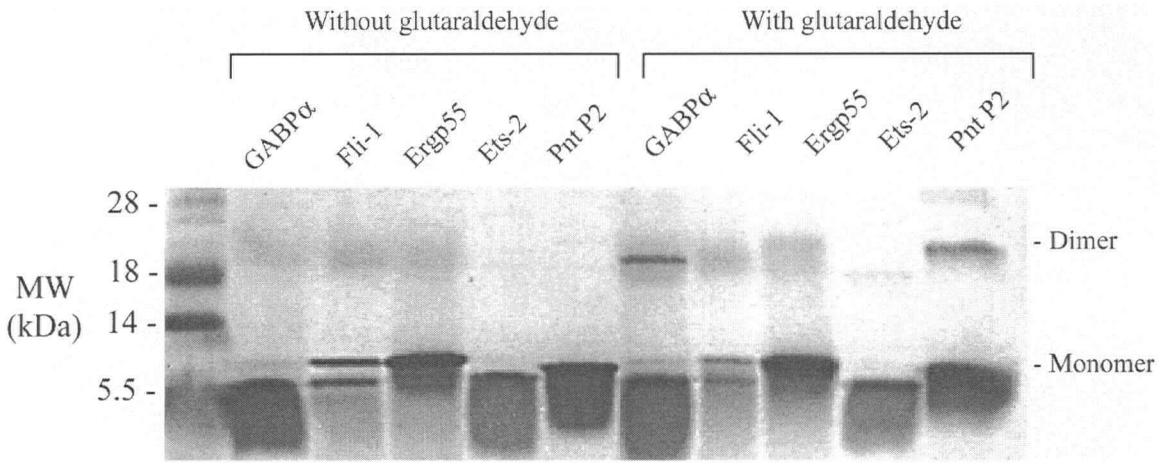


Figure 3.3. SDS-PAGE analysis of glutaraldehyde cross-linking reactions of PNT domains from various Ets proteins.

using GABPα PNT$^{(168-254)}$ form B. If form B is a oligomeric form of GABPα PNT$^{(168-254)}$, it would be expected to crosslink more than a sample of GABPα PNT$^{(168-254)}$ form A. However, similar crosslinking results were obtained using GABPα PNT$^{(168-254)}$ forms A and B suggesting that these two forms of GABPα PNT$^{(168-254)}$ have the same oligomeric state (data not shown). Crosslinking experiments were also conducted using higher concentrations of GABPα PNT$^{(168-254)}$. In these experiments, bands corresponding to the mass expected for trimers and tetramers were observed (data not shown). As with previous experiments, it was estimated that only 25% of the protein crosslinked. This behaviour was also observed with the PNT domains from Pnt P2 and Ets-2, a protein that at lower concentrations did not demonstrate significant crosslinking.

Taken together, these results do not support the hypothesis that GABPα PNT$^{(168-254)}$ form B is an oligomeric form of the domain. Similar levels of crosslinking are observed with forms A and B suggesting that both forms have the same oligomeric state. Trimer and tetramer formation is seen at higher concentrations of GABPα PNT$^{(168-254)}$ as well. Although this alone could be consistent with self-association, it could also be due intermolecular crosslinks formed during random, non-specific collision of the protein molecules. The fact that this is observed with other PNT domains, including those that show no significant crosslinking at lower concentrations, supports the latter view.

One additional interesting observation regarding Figure 3.3 is the lack of significant crosslinking observed with the minimal PNT domain from the Ets protein Ergp55. As discussed in section 1.4.3, other workers have reported self-association of Ergp55. The discrepancy could be due to the fact that experiments in the McIntosh laboratory have used minimal PNT domains as defined by sequence homology and

structural studies whereas other workers generally use larger protein fragments containing the PNT domain. Regions of these proteins outside of the minimal PNT domain may be required for interaction.

3.3.5 NMR dilution series

In a final attempt to demonstrate self-association of GABPα PNT$^{(168-254)}$, NMR spectroscopy was used. NMR is a very useful tool for studying protein interactions. When proteins associate, changes occur in their NMR spectra as residues are exposed to new environments. One can begin with a concentrated sample of protein and look for any changes in the NMR spectrum as the protein is diluted. If the protein is self-associating over this concentration range, some of these complexes should dissociate upon dilution and producing changes in the NMR spectra.

A ^{15}N-labeled sample of GABPα PNT$^{(168-254)}$ form A was prepared at a concentration of 3.7 mM and a ^{15}N-^{1}H HSQC spectrum recorded. This sample was gradually diluted with the same buffer and HSQC spectra recorded at protein concentrations of 1.7, 1.3, 1.0, 0.5 and 0.1 mM. When these spectra were overlaid, no significant peak shifts were seen suggesting that the association state of the protein is not changing under these conditions. Thus, if GABPα PNT$^{(168-254)}$ form A does associate to produce form B, the dissociation constant is not between 0.1 mM and 3.7 mM. The dissociation constant is unlikely to lie above 3.7 mM since it is extremely doubtful that an interaction this weak would even be detectable during purification of the protein. In addition, the biological relevance of the interaction would be highly questionable. The dissociation constant is also unlikely to lie below 0.1 mM because an interaction of this

strength would have been immediately apparent during previous NMR relaxation and ultra-centrifugation analyses. Consistent with the results described above, this experiment does not support the view that GABPα $PNT^{(168-254)}$ is self-associating.

3.3.6 Relaxation analysis of GABPα $PNT^{(168-254)}$ form A and B

To unequivocally determine the oligomeric state of GABPα $PNT^{(168-254)}$ form B, ^{15}N NMR relaxation analysis was performed as described in section 3.2.9. Both forms A and B were found to have similar overall rotational correlation times of 6.3 and 5.8 ns, respectively. These values are similar to that of 6.7 ns obtained by previous workers (see section 3.1) and are within the range expected for monomeric proteins. These results indicate that form A and form B are both monomeric forms of GABPα $PNT^{(168-254)}$.

3.3.7 ^{15}N-^{1}H HSQC analysis of form A and B as a function of time

To more clearly establish the relationship between form A and form B, additional NMR experiments were conducted on the proteins. Inspection of the overlay of the ^{15}N-^{1}H HSQC spectra of form A and B in Figure 3.2 reveals that the vast majority of the resonances present in form A overlap very well with resonances in form B. This suggests that both forms have a similar conformation or contain populations of proteins with similar conformations. However, form B has additional resonances, suggestive of structural heterogeneity. If this is the case, the two forms may be in slow equilibrium and separation of the forms may allow equilibration to be re-established.

To investigate this possibility, a sample of ^{15}N-labeled GABPα $PNT^{(168-254)}$ form A was stored at 4°C and subjected to ^{15}N-^{1}H HSQC analysis over a period of two weeks.

During this time, many new resonances gradually appeared in the ^{15}N-^{1}H HSQC spectra suggesting that the conformation of the protein was changing with time. These resonances generally corresponded with the resonances in the ^{15}N-^{1}H HSQC spectrum of GABPα $PNT^{(168-254)}$ form B. To confirm this, an ^{15}N, ^{13}C-labeled sample of GABPα $PNT^{(168-254)}$ was prepared and the polypeptide backbone resonances of both form A and B assigned using three-dimensional NMR experiments as described in section 3.2.9. Analysis of this data indicated that these new resonances could be paired with an original resonance using $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{13}C'$, H^{N} and N shifts. This is consistent with a slow conversion from form A, which appears to initially contain a relatively homogenous population of GABPα $PNT^{(168-254)}$, to form B, which appears to be a heterogenous mixture. The dominant species in each form is monomeric and is similar in structure. Since almost every residue in the protein has multiple resonances, the conversion of form A to form B involves a global change in the conformation of the protein rather than a localized conformational change. Following two weeks in storage at 4°C, GABPα $PNT^{(168-254)}$ form A was subjected to anion exchange FPLC using identical conditions to its original purification from the *E. coli* cell lysate. The sample now produced an elution profile with form A and form B present in similar proportions. This confirms that during storage, a sample consisting exclusively of form A is gradually converting to a mixture of form A and form B, consistent with the above hypothesis.

Similar experiments were performed with GABPα $PNT^{(168-254)}$ form B. However, over time, form B demonstrated a tendency to precipitate. Concomitant with the appearance of precipitate in the form B sample was a decrease in intensity of many resonances in the ^{15}N-^{1}H HSQC spectra. This suggests that the conformation of GABPα

PNT$^{(168-254)}$ in form B responsible for these secondary resonances is unstable and aggregates. These aggregates would not be observable by NMR. When the supernatant from the form B sample was subjected to anion exchange chromatography using conditions from the original purification, an elution profile similar to that seen in the original purification was again obtained with forms A and B present in the expected proportions. This confirms that conversion between forms A and B involves a physical change in the conformation of the protein rather than a non-reversible chemical modification. These results are consistent with the hypothesis that forms A and B slowly equilibrate. Form A appears to be a homogenous population of GABPα PNT$^{(168-254)}$ while form B appears to be a mixture of two conformations of the protein. One of these conformations is present in form A while the second shows an increased tendency to aggregate.

3.3.8 Chromatographic behaviour of other GABPα PNT constructs

As discussed in section 3.1, previous work with GABPα PNT$^{(138-254)}$ established that this longer construct eluted at a single position during anion exchange FPLC. Thus, forms A and B of GABPα PNT$^{(168-254)}$ may be unique to the shorter construct. To investigate this several other GABPα PNT domain proteins were constructed.

GABPα PNT$^{(138-254)}$ (see Figure 2.6) was stored at 4°C for several weeks allowing proteolysis of the flexible amino-terminal amino acids. Mass spectrometry of the sample indicated that a mixture of fragments of various lengths was produced, the shortest of which consisted of amino acids 161-254. This mixture was subjected to anion exchange FPLC under conditions used in previous GABPα PNT$^{(168-254)}$ purifications. These

proteins all eluted at the same single position. Analysis of the ^{15}N-^{1}H HSQC spectra of GABPα PNT$^{(138-254)}$ over time indicated that this construct did not undergo conformational changes as its spectrum was unchanged. This work confirms that forms A and B are unique behaviour of the shortest GABPα PNT constructs.

A 6×His-tagged version of GABPα PNT$^{(168-254)}$ was constructed as described in section 3.2.5. When this construct was purified using anion exchange FPLC, the protein eluted at a single position, similar to the results obtained with GABPα PNT$^{(138-254)}$. This experiment demonstrates that the extra amino acids at the amino-terminus required for production of a single conformation of the PNT domain need not be residues from GABPα, but can be any residues.

Another sample of 6×His-tagged GABPα PNT$^{(168-254)}$ was purified using nickel chelate chromatography using a procedure similar to that described for PBP in section 2.2.9. The 6×His-tag was then cleaved off with thrombin as described in section 2.2.11 resulting in a PNT domain construct with only three amino acid residues on the amino-terminus of GABPα PNT$^{(168-254)}$. Again, this construct was subjected to anion exchange FPLC and eluted at a single position. Thus, only three amino acid residues on the amino-terminus are required to produce a stable conformation of the PNT domain.

The chromatographic behaviour of the various PNT domain constructs is summarized in Figure 3.4. This work confirms that the production of forms A and B is unique to the shortest PNT domain construct, GABPα PNT$^{(168-254)}$.

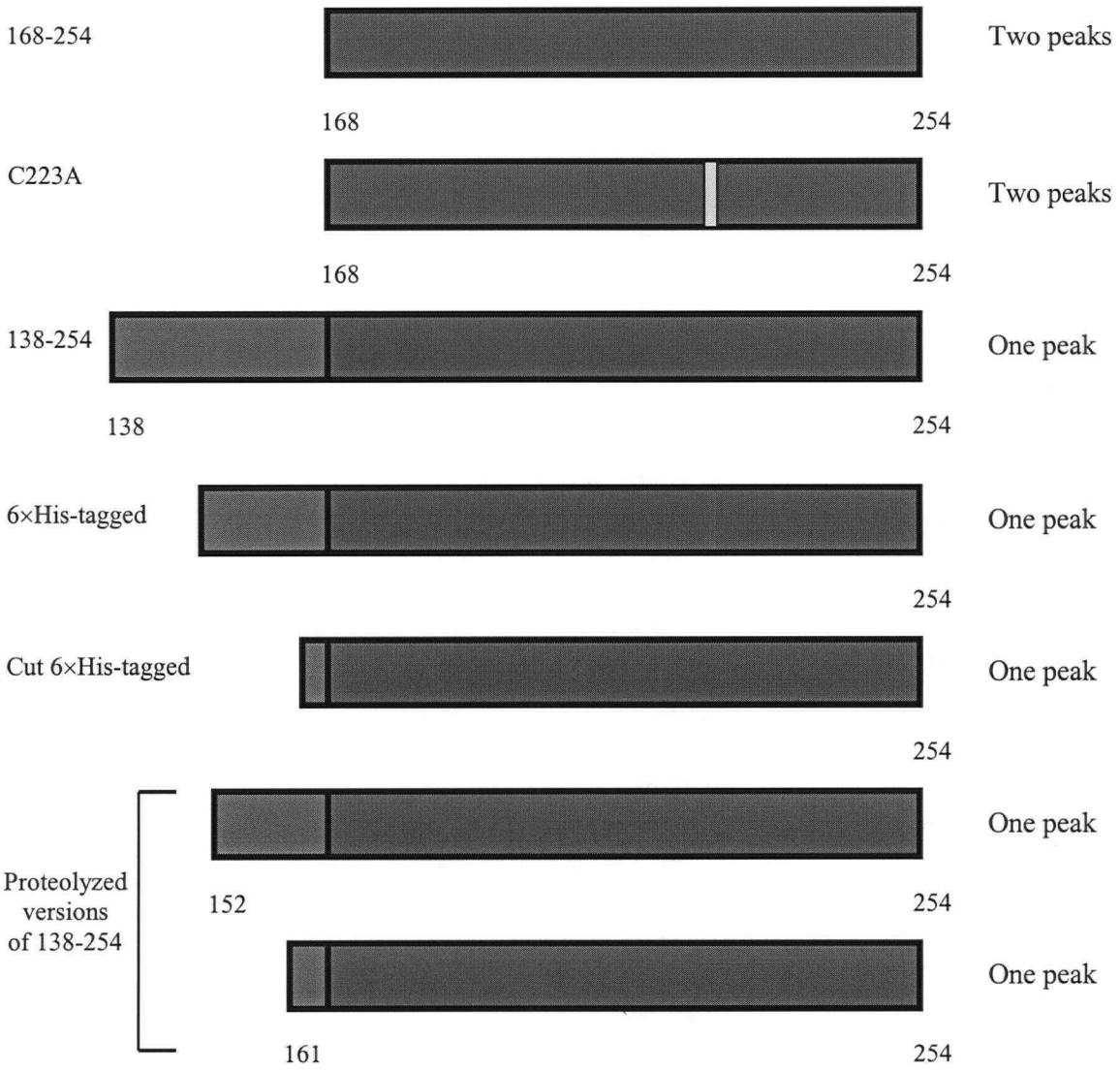


Figure 3.4. Behaviour of various GABPα PNT domain constructs during anion exchange FPLC.

3.4 Summary and conclusions

Purification of minimal-sized GABPα PNT$^{(168-254)}$ by anion exchange FPLC resulted in resolution of two forms of the protein, A and B. Construction of a Cys → Ala mutant confirmed that form B was not due to a disulfide linkage between two GABPα PNT$^{(168-254)}$ monomers. This also eliminated the possibility of more extreme Cys oxidation to cysteic acid being responsible, although the mass spectrometry results described in section 3.3.2 made this possibility extremely unlikely. Glutaraldehyde cross-linking experiments and NMR dilution experiments failed to produce evidence of GABPα PNT$^{(168-254)}$ self-association and relaxation experiments indicate that both forms are monomeric. Observation of the ^{15}N-^{1}H HSQC spectra of both forms over time suggest that these two forms of GABPα PNT$^{(168-254)}$ are in equilibrium and interconvert over the course of several days. Form A consists of GABPα PNT$^{(168-254)}$ in a single conformation while form B is a heterogenous mixture of GABPα PNT$^{(168-254)}$ with two major conformations. One of these conformations has a greater tendency to aggregate and gradually precipitates out of solution. It is surprising that the protein shows such conformational heterogeneity on a slow timescale.

Further elucidation of the details of these processes is largely academic since only the minimal GABPα PNT$^{(168-254)}$ construct results in production of forms A and B. This is an unfortunate artifact and is unlikely to be a biologically relevant process. Although the structure of GABPα PNT$^{(168-254)}$ form A (shown in Figure 1.4) appears to encompass the entire PNT domain, this conformation is apparently unstable. Addition of a few extra amino acids at the amino-terminus is required to abrogate this behaviour. One possible explanation may involve the macrodipole moment of α-helices (Wada, 1976). The partial

positive charge at the amino-terminal end of helix H1 and the full positive charge at the amino-terminus of the protein may be in close proximity in GABPα $PNT^{(168-254)}$ resulting in electrostatic repulsion. Since H1 is only loosely associated with the remainder of the PNT domain, this repulsion may be sufficient to introduce instability into the conformation of the protein without disrupting the entire structure of the domain. Addition of amino acid residues at the amino-terminus of GABPα $PNT^{(168-254)}$ would increase the separation between the charges, resulting in less electrostatic repulsion. This may result in stabilization of the conformation and is consistent with the results obtained above. The work described in this chapter demonstrates that GABPα $PNT^{(168-254)}$ is not a discrete dimer. The crosslinking results may be due to the conformational mobility of this GABPα PNT construct.

Chapter 4

Investigation of Unassigned Resonances in the GABPα PNT NMR Spectrum

Previous work in the McIntosh laboratory, described below in section 4.1, suggested that an unidentified small molecule bound the PNT domain of GABPα. Work described in this chapter was undertaken to identify this molecule and determine if it was binding to the domain.

4.1 Background

During previous work by Lisa Gentile on the structure of GABPα PNT$^{(168\text{-}254)}$ form A, three resonances were present in the ^{15}N-^{1}H HSQC spectrum of the protein that could not be assigned to any amino acid residues in the protein. Similar peaks were observed in the spectra of GABPα PNT$^{(168\text{-}254)}$ form B and GABPα PNT$^{(138\text{-}254)}$. The ^{15}N-^{1}H HSQC spectrum of GABPα PNT$^{(168\text{-}254)}$ is shown in Figure 4.1 with these unassigned resonances boxed in blue. The fact that these resonances are observable in the spectrum indicates that they must originate from an *E. coli* molecule with an amide group since a contaminant from another source would not contain isotopic labels. In addition, ^{15}N relaxation analysis indicated that these resonances were very sharp compared to the protein resonances and had long T_1 and T_2 relaxation times. This suggests that these resonances are from a small molecule that could be weakly binding the domain. Identification of this molecule could reveal a novel function of the PNT domain.

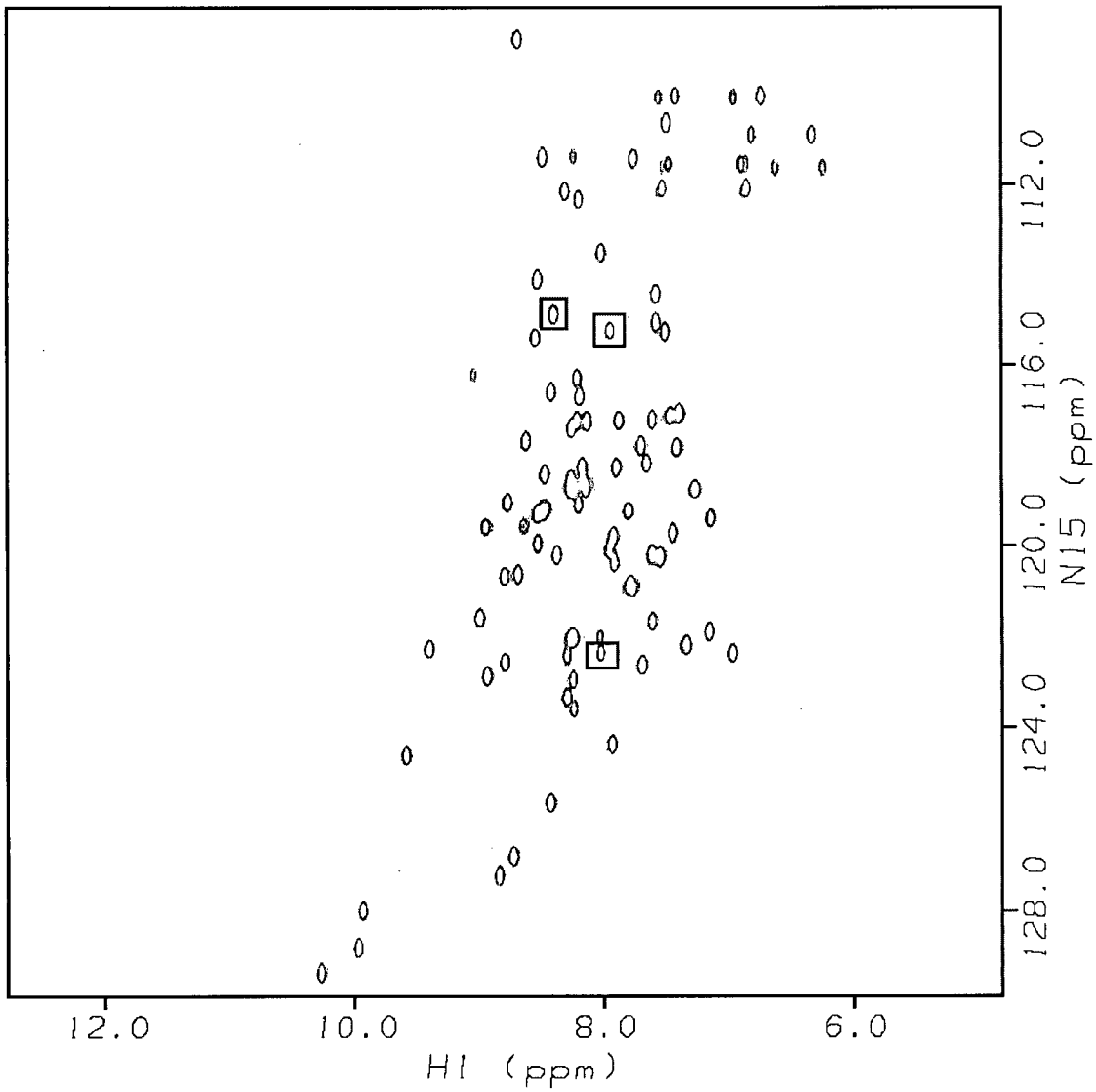


Figure 4.1. ^{15}N-^{1}H HSQC spectra of GABPα PNT$^{(168-254)}$ form A. The three unassignable resonances are boxed.

4.2 Materials and methods

4.2.1 Gel filtration chromatography

GABPα PNT$^{(168-254)}$ in 20 mM Na_2HPO_4, 20 mM NaCl, pH 7.2 was diluted ten-fold into 50 mM Tris-HCl, 100 mM NaCl, pH 7.0 and applied to a Sephacryl S-100 gel filtration column at a rate of 0.5 ml/min. The column was run using 50 mM Tris-HCl, 100 mM NaCl, pH 7.0 at a flow rate of 0.5 ml/min.

4.2.2 RP-HPLC of GABPα PNT$^{(138-254)}$

A 1.5 mM sample of GABPα PNT$^{(138-254)}$ was diluted with five volumes of 0.1% TFA which lowered the pH of the solution to ~2. The sample was then loaded onto a reverse phase 1080 HPLC column and the protein eluted using a linear gradient of acetonitrile. The A_{280} of the eluate was monitored and GABPα PNT$^{(138-254)}$ collected as a single peak, which was lyophilized overnight. The lyophilized protein was resuspended in 10 ml of 20 mM Na_2HPO_4, 20 mM NaCl, pH 7.2 and was concentrated to 400 μM for NMR analysis.

4.2.3 NMR spectroscopy

NMR spectra were collected as described in section 3.2.9.

4.3 Results and discussion

4.3.1 Analysis of NMR spectra

Through the use of HNCACB and CBCA(CO)NH 3D NMR experiments, the backbone resonances of GABPα PNT$^{(168-254)}$ form A were assigned. The results of this

analysis confirmed the previous GABPα PNT$^{(168-254)}$ assignments performed by Lisa Gentile. The same three resonances in the ^{15}N-^{1}H HSQC spectrum were not assignable to any amino acids in the protein. These experiments also established that these three resonances were associated with carbon chemical shifts not expected for amino acids.

4.3.2 Purification GABPα PNT$^{(168-254)}$ at pH 7.0

To determine if the unassigned resonances arise from a co-eluting contaminant, GABPα PNT$^{(168-254)}$ was purified as described previously in section 3.2.2 with the exception that the pH was 7.0 rather then 8.5. Analysis of the ^{15}N-^{1}H HSQC spectrum revealed that the three unassigned resonances were still present. In addition, there were several new unassigned resonances present as well.

GABPα PNT$^{(168-254)}$ purified using anion exchange FPLC at pH 8.5 was then subjected to gel filtration chromatography in an attempt to remove the molecule responsible for the unassigned resonances. Again, analysis of the ^{15}N-^{1}H HSQC spectrum following gel filtration revealed that the unassigned resonances were still present. Thus, the molecule was either binding to the domain or is large enough (greater than ~1 kDa) that it is not removed by gel filtration or by dialysis.

4.3.3 Purification of 6×His-tagged GABPα PNT$^{(168-254)}$

6×His-tagged GABPα PNT$^{(168-254)}$ was expressed and purified as described for 6×His-tagged PBP in sections 2.2.8 and 2.2.9, respectively. The unassigned resonances were not present in the ^{15}N-^{1}H HSQC spectrum of this protein. Although the high salt concentration used during the purification (500 mM NaCl, pH 7.5) could disrupt binding

between the molecule and GABPα PNT$^{(168-254)}$, this result raises the possibility that the molecule could actually be a family of contaminants from the anion exchange purification that were not removed by gel filtration chromatography due to the low resolution of this technique.

4.3.4 Reversed- phase HPLC of GABPα PNT$^{(168-254)}$

In an attempt to separate the molecule responsible for these resonances from GABPα PNT$^{(168-254)}$, reversed-phase HPLC was employed. SDS-PAGE analysis of GABPα PNT$^{(138-254)}$ samples taken before and after RP-HPLC revealed that the protein was highly purified (data not shown). An overlay of the ^{15}N-^{1}H HSQC spectra of GABPα PNT$^{(138-254)}$ before and after purification by HPLC appears in Figure 4.2. Inspection of the overlay reveals that HPLC removed the species contributing to the resonances that were unassignable, which are boxed in blue. Each resonance in the spectrum collected after HPLC overlayed exactly with a resonance in the spectrum collected before HPLC indicating that the protein has the same conformation. A molecule strongly binding to a protein would be expected to perturb the chemical shifts of resonances in the HSQC spectrum of the protein. Thus, if a molecule was strongly bound to GABPα PNT$^{(138-254)}$, we would expect to see changes in the HSQC spectrum when this molecule was removed. This was not observed, therefore it is unlikely that the assignable resonances in the HSQC are due to a tightly bound molecule. This experiment was repeated using GABPα PNT$^{(168-254)}$ form A and similar results were obtained. The unassignable resonances were no longer present but the protein had the same conformation as evidenced by the exact

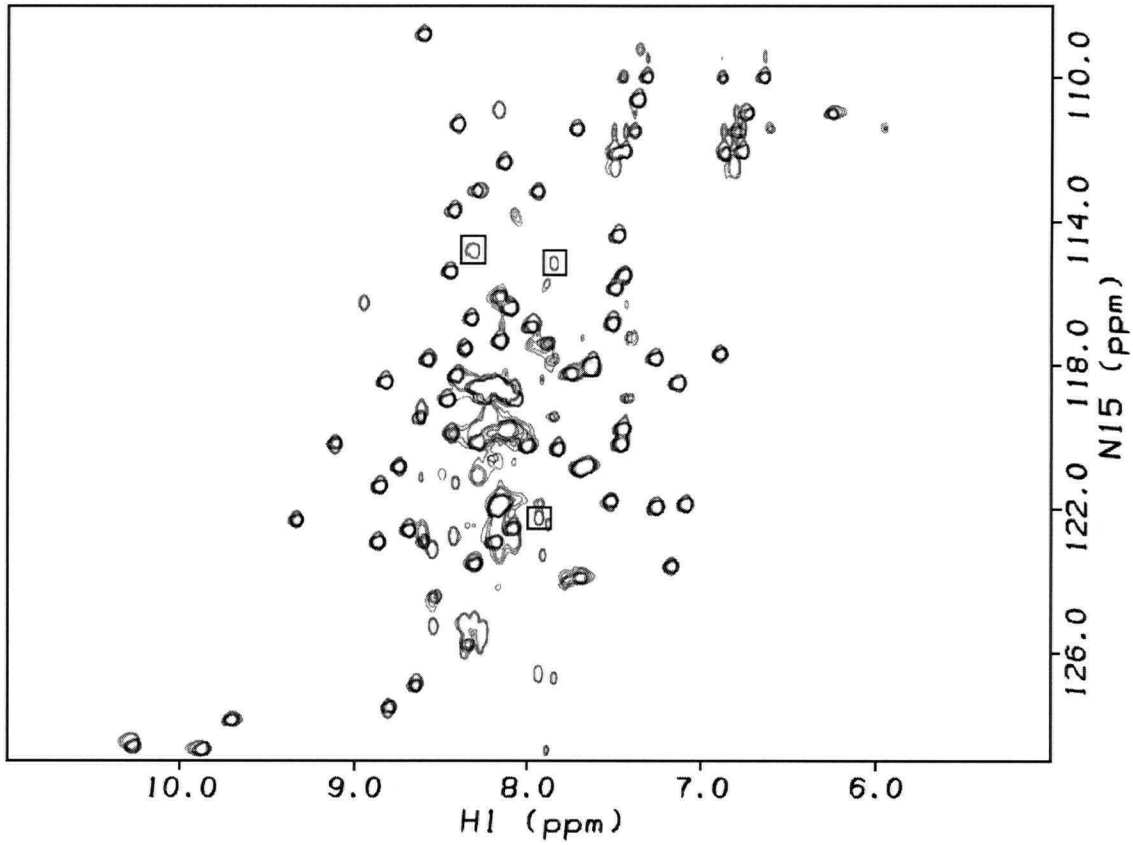


Figure 4.2 Overlay of ^{15}N-^{1}H HSQC spectra of GABPα PNT$^{(138-254)}$ before (grey) and after (black) reversed-phase HPLC purification. Unassigned residues are boxed.

overlap of resonances in the ^{15}N-^{1}H HSQC spectra. These results suggest that these resonances could arise from a small ^{15}N, ^{13}C labelled molecule from *E. coli* that co-purifies with the GABPα PNT domain but does not bind the protein. It is interesting to note that GABPα $PNT^{(168-254)}$ forms A and B and GABPα $PNT^{(138-254)}$ have similar but not identical unassignable resonances. This suggests that rather than being due to the same molecule, these resonances arise from a group of similar molecules, which may each elute under the different conditions corresponding to the different GABPα PNT constructs. This experiment also confirms work described in chapter 3. The multiple resonances found in GABPα $PNT^{(168-254)}$ form B were no longer present following HPLC as the ^{15}N-^{1}H HSQC spectrum was identical to that of GABPα $PNT^{(168-254)}$ form A. This could reflect folding of GABPα $PNT^{(168-254)}$ to a single conformation or removal of alternate conformations.

4.3.5 ESI-MS analysis

Given that the unknown molecule was removed by RP-HPLC, it was reasonable to assume that the desalting column used during mass spectrometry analysis may have also removed this molecule making determination of its mass difficult. To circumvent this problem, GABPα $PNT^{(168-254)}$ was dialyzed into deionized water so that direct injection into the mass spectrometer was possible. Analysis revealed a series of low molecular weight molecules that differed in mass by 203 amu. Interestingly, this is precisely the mass of an N-acetyl hexosamine. This raised the possibility that the unassigned resonances in the ^{15}N-^{1}H HSQC spectra were due to the presence of oligomers of an N-acetyl hexosamine, such as N-acetyl glucosamine.

4.3.6 NMR analysis

An N-acetylated sugar would be expected to possess amide NH shifts and thus would be present in the ^{15}N-^{1}H HSQC spectrum of a protein. However, some carbohydrate carbon resonances would be at chemical shifts distinct from the chemical shifts found in proteins. For example, the anomeric carbon of a hexose is expected to have a chemical shift in the range of 90-105 ppm, whereas unmodified proteins have no carbons with chemical shifts in this range. Examination of the ^{13}C-^{1}H constant time HSQC of GABPα $PNT^{(168-254)}$ form A, shown in Figure 4.3, revealed the presence of three resonances in the 95-105 ppm range, suggesting that a carbohydrate was present. During the backbone assignment of GABPα $PNT^{(168-254)}$ form A, the β carbon of Thr 208 was found to have the highest chemical shift of any aliphatic carbon atom in the protein (70.31 ppm). However, inspection of Figure 4.3 reveals the presence of over a dozen resonances in the 70-80 ppm range. In addition, the phase of these resonances indicates that they have an even number of carbon neighbors. This data is also consistent with the presence of a carbohydrate. Based on this information, the 3D NMR experiments conducted earlier were re-examined to partially assign these carbon shifts. Carbon chemical shifts for the unassigned peaks were compared with published chemical shifts for N-acetyl-glucosamine, an N-acetylated hexosamine present in the peptidoglycan layer of *E. coli*. Results of this comparison for one of the unassigned peaks appear in Table 4.1 and strongly suggest that the molecule present in the GABPα $PNT^{(168-254)}$ form A sample is an N-acetyl hexosamine oligomer.

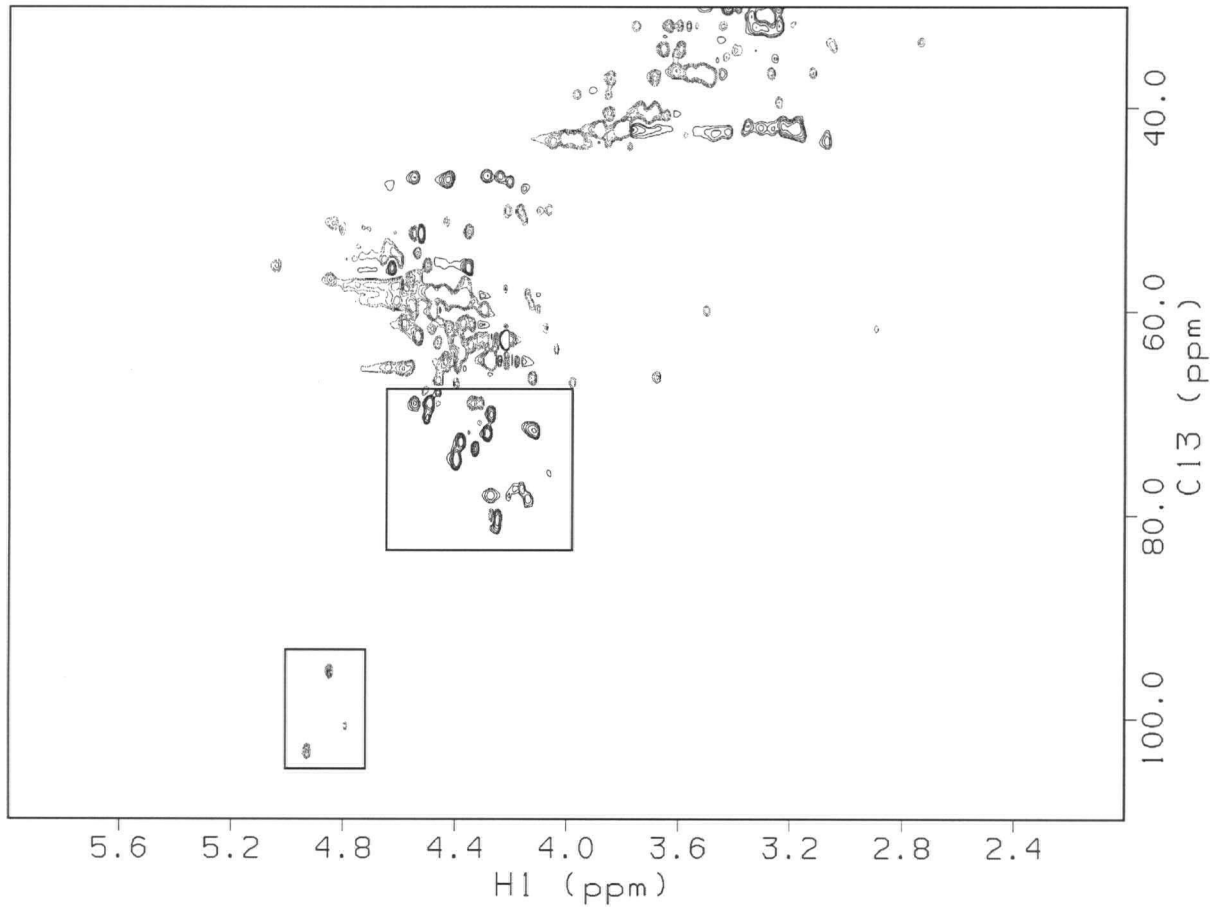


Figure 4.3. The ^{13}C-^{1}H constant time HSQC spectrum of GABPα $PNT^{(168-254)}$ form A contains resonances characteristic of carbohydrates. Resonances characteristic of carbohydrate carbons are boxed.

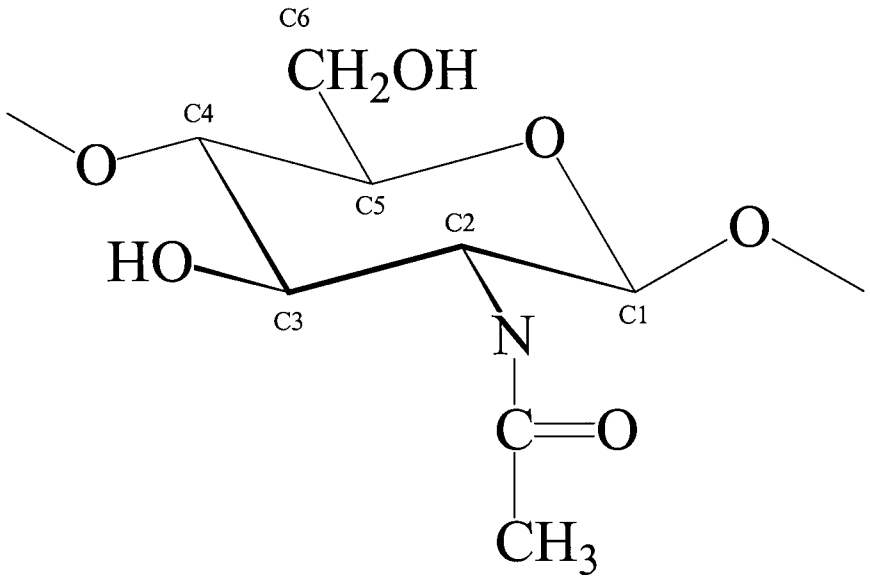


Figure 4.4. Structure of a β(1→4) linked N-acetyl-D-glucosamine residue.

Chemical shift of unassigned peak (ppm)	**Published chemical shift for β(1→4) N-acetyl-D-glucosamine residue (ppm)**
24.7	22.3 - Methyl group carbon
55.5	56.3 - C2
71.8	72.6 - C3
95.3	95.0 - C1

Table 4.1. Comparison of ^{13}C carbon chemical shifts for one unassigned resonance in the ^{15}N-^{1}H HSQC spectrum of GABPα PNT$^{(168-254)}$ form A with published values for a β(1→4) linked N-acetyl-D-glucosamine residue (Saito *et al.*, 1981).

4.3.7 NMR relaxation analysis

NMR relaxation analysis was conducted on the three putative N-acetyl hexosamine resonances present in the ^{15}N-^{1}H HSQC spectrum of GABPα PNT$^{(168-254)}$ form A. The average T_1 and T_2 values and the heteronuclear NOE for GABPα PNT$^{(168-254)}$ form A and for the putative carbohydrate peaks are shown in Table 4.2. The carbohydrate resonances had T_1 and T_2 relaxation times that were much longer than the times measured for the protein. In addition, these resonances have a negative NOE. Both these observations are indicative of fast tumbling in solution. If a molecule were tightly bound to a protein, it would be expected to have T_1 and T_2 relaxation times that were similar to the protein. The fact that the relaxation times for these resonances are longer suggest that the molecule is not tightly bound to GABPα PNT$^{(168-254)}$ form A and is highly mobile in solution.

4.4 Summary and conclusions

The ^{15}N-^{1}H HSQC spectra of GABPα PNT domain constructs contained resonances that were not assignable to any amino acid residues in the protein, raising the possibility that a small molecule could be bound to domain. This was supported by the observation of this species or closely related forms when GABPα PNT$^{(138-254)}$ and GABPα PNT$^{(168-254)}$ were purified under different anion exchange FPLC conditions, and that dialysis and gel filtration did not remove the species. Using a combination of mass spectrometry and NMR analysis, this molecule was tentatively identified as an N-acetyl hexosamine oligomer.

Parameter	GABPα PNT$^{(168-254)}$ form A	Putative carbohydrate resonances
T_1	440 ms	645 ms
T_2	123 ms	390 ms
NOE	0.703	-0.188

Table 4.2. The putative carbohydrate resonances in the NMR spectra of GABPα PNT$^{(168-254)}$ form A arise from a molecule that is small and highly mobile in solution compared to the protein. Average T_1 and T_2 values in ms and the NOE for GABPα PNT$^{(168-254)}$ form A and the carbohydrate resonances are shown.

This result was intriguing because modification of proteins with O-linked β-N-acetyl glucosamine is a common post-translational modification in eukaryotes and is believed to play an important role in signal transduction (Wells *et al.*, 2001). Since GABPα PNT$^{(168-254)}$ was expressed in *E. coli* for these studies, it would not be modified itself. However, in its normal eukaryotic environment, one function of the GABPα PNT domain could be to bind another protein modified with O-linked β-N-acetyl glucosamine. When expressed in *E. coli*, the domain may weakly bind N-acetyl glucosamine from the peptidoglycan layer that is present in the lysate. However, NMR relaxation analysis suggests that the molecule is not tightly bound to GABPα PNT$^{(168-254)}$. Comparison of the ^{15}N-^{1}H HSQC spectra of GABPα PNT$^{(138-254)}$ before and after RP-HPLC is also consistent with this view as no significant resonance shifts were observed. As mentioned above, one possible source of N-acetyl hexosamines is from the peptidoglycan layer of *E. coli*. The glycan chains in this layer consist of alternating N-acetyl glucosamine and N-acetyl muramic acid residues. This latter carbohydrate contains a carboxyl group giving the chains a negative charge. It is possible that during lysis of the *E. coli* and purification of GABPα PNT$^{(168-254)}$, the peptidoglycan layer was fragmented and some lighter weight fragments were loaded unto the anion exchange column. Since these fragments would be negatively charged, they may co-purify with GABPα PNT$^{(168-254)}$. Slightly different NMR resonances present in the NMR spectra of different GABPα PNT domain constructs indicated that different fragments were present and that these are probably oligosaccharides as they are not removed by dialysis or gel filtration. This also explains why these resonances are not present in a sample of 6×His-tagged GABPα PNT$^{(168-254)}$ purified by nickel chelate chromatography.

Chapter 5

Concluding Remarks

As stated in section 1.6, based on evidence presented in sections 1.4 and 1.5, the following hypothesis was formed and provided the basis of this thesis work: **The PNT domain mediates interactions of Ets proteins with either PNT domain containing-proteins or with other proteins involved in signal transduction and transcriptional regulation.**

The three goals of this thesis work were:

i) To use a yeast two-hybrid system to screen for potential protein partners of the PNT domain of GABPα.

ii) A detailed study of previous work on the PNT domain of GABPα to investigate the possibility of GABPα PNT self-association.

iii) A detailed study of previous work on the PNT domain of GABPα to investigate the possibility of a small molecule binding to the GABPα PNT domain.

Goal i was addressed in chapter 2. A yeast two-hybrid screen was used in an attempt to identify proteins that interacted with the minimal PNT domain of GABPα. One potential interacting protein, PBP, was identified. However, attempts to demonstrate a direct interaction between the proteins *in vitro* was unsuccessful. Several possible explanations of this negative result are discussed in section 2.5.

Goal ii was addressed in chapter 3. Although the behaviour of GABPα PNT$^{(168-254)}$ during anion exchange FPLC suggested that this protein may form oligomers, no

evidence of self-association of the protein was obtained using several different techniques. NMR experiments suggested that GABPα $PNT^{(168-254)}$ is heterogenous and existed in two different conformations that interconverted over the course of several days. However, this appears to be unique to this PNT domain construct as additional amino acids at the amino-terminus are sufficient to abrogate this behaviour. As a result, this is likely an unfortunate artifact with no biological relevance.

Goal iii was addressed in chapter 4. Resonances not assignable to any amino acid residues were present in the ^{15}N-^{1}H HSQC NMR spectrum of GABPα $PNT^{(168-254)}$ purified using anion exchange FPLC. This raised the possibility that a small molecule was binding to the domain. This molecule was tentatively identified as an N-acetyl hexosamine oligomer. However, NMR relaxation experiments and removal of the molecule using RP-HPLC failed to provide any evidence that this molecule was binding to the PNT domain of GABPα with appreciable affinity. One possible explanation of the presence of this carbohydrate is that it is a fragment of the *E. coli* peptidoglycan layer that co-purifies with the GABPα PNT domain due to the presence of carboxyl groups that interact with the anion exchange column.

The work described in this thesis has been unsuccessful in demonstrating a role for the GABPα PNT domain in protein interactions. However, one report has suggested that the GABPα PNT domain may be involved in binding to GABPβ (Chinenov *et al.*, 2000) and that this may be involved in regulating heterotetramer formation (see Figure 1.2). This may open a new avenue of investigation for future GABPα PNT domain work.

As mentioned in chapter 1, other work in the McIntosh laboratory using the minimal PNT domains from several Ets transcription factors have failed to reproduce

many of the interactions that have been reported in the literature. One possible explanation is simply that the minimal PNT domain is not sufficient to mediate these interactions. Virtually all of the reported interactions used large protein fragments containing many residues that lie outside of the PNT domain. These residues may be critical for the interactions. Alternatively, several regions on the protein fragments may mediate these interactions with the result that PNT domain interactions alone may be very weak. Additional studies of PNT domains from several Ets proteins using different fragment sizes may clarify the precise role the PNT domain plays in protein interactions.

Bibliography

Baker, D. A., Mille-Baker, B., Wainwright, S. M., Ish-Horowicz, D. & Dibb, N. J. (2001). Mae mediates MAP kinase phosphorylation of Ets transcription factors in Drosophila. *Nature* **411**, 330-334.

Barr, M. M., Tu, H., Van Aelst, L. & Wigler, M. (1996). Identification of Ste4 as a potential regulator of Byr2 in the sexual response pathway of Schizosaccharomyces pombe. *Mol. Cell. Biol.* **16**, 5597-5603.

Basuyaux, J., Ferreira, E., Stehelin, D. & Buttice, G. (1997). The Ets Transcription Factors Interact with Each Other and with the c-Fos/c-Jun Complex via Distinct Protein Domains in a DNA-dependent and -independent Manner. *J. Biol. Chem.* **272**, 26188-26195.

Batchelor, A., Piper, D., de la Brousse, F., McKnight, S. & Wolberger, C. (1998). The Structure of GABP α/β: An ETS Domain-Ankyrin Repeat Heterodimer Bound to DNA. *Science* **279**, 1037-1041.

Berger, B., Wilson, D., Wolf, E., Tonchev, T., Milla, M. & Kim, P. (1995). Predicting Coiled-Coils by Use of Pairwise Residue Correlations. *Proc. Natl. Acad. Sci.* **92**, 8259-8263.

Brent, R. & Ptashne, M. (1985). A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* **43**, 729-736.

Buttice, G., Duterque-Coquillaud, M., Basuyaux, J., Carrere, S., Kurkinen, M. & Stehelin, D. (1996). Erg, an Ets-Family Member, Differentially Regulates Human

Collagenase1 (MMP1) and Stromelysin1 (MMP3) Gene Expression by Physically Interacting with the Fos/Jun Complex. *Oncogene* **13**, 2297-2306.

Carrere, S., Verger, A., Flourens, A., Stehelin, D. & Duterque-Coquillaud, M. (1998). Erg Proteins, Transcription Factors of the Ets Family, Form Homo, Heterodimers and Ternary Complexes via Two Distinct Domains. *Oncogene* **16**, 3261-3268.

Carroll, M., Tomasson, M., Barker, G., Golub, T. & Gilliland, D. (1996). The Tel/Platelet-Derived Growth Factor Beta Receptor (PDGFbR) Fusion in Chronic Myelomonocytic Leukemia is a Transforming Protein that Self-Associates and Activates PDGFbR Kinase-Dependent Signaling Pathways. *Proc. Natl. Acad. Sci.* **93**, 14845-14850.

Cazzaniga, G., Tosi, S., Aloisi, A., Giudici, G., Daniotti, M., Pioltelli, P., Kearney, L. & Biondi, A. (1999). The Tyrosine Kinase Abl-Related Gene ARG Is Fused to ETV6 in an AML-M4Eo Patient With a t(1;12)(q25;p13): Molecular Cloning of Both Reciprocal Transcripts. *Blood* **94**, 4370-4373.

Chakrabarti, S., Sood, R., Ganguly, S., Bohlander, S., Shen, Z. & Nucifora, G. (1999). Modulation of Tel Transcription Activity by Interaction with the Ubiquitin-Conjugating Enzyme UBC9. *Proc. Natl. Acad. Sci.* **96**, 7467-7472.

Chen, C. & Shyu, A., (1994). Selective degradation of early-response-gene mRNAs: functional analyses of sequence features of the AU-rich elements. *Mol. Cell. Biol.* **14**, 8471-8482.

Chi, S. W., Ayed, A. & Arrowsmith, C. H. (1999). Solution structure of a conserved C-terminal domain of p73 with structural homology to the SAM domain. *EMBO J.* **18**, 4438-4445.

Chinenov, Y., Henzl, M. & Martin, M. E. (2000). The alpha and beta subunits of the GA-binding protein form a stable heterodimer in solution. Revised model of heterotetrameric complex assembly. *J. Biol. Chem.* **275**, 7749-7756.

Donaldson, L., Petersen, J., Graves, B. & McIntosh, L. (1996). Solution Structure of the ETS Domain from Murine Ets-1: a Winged Helix-Turn-Helix DNA Binding Motif. *EMBO J.* **15**, 125-134.

Eguchi, M., Eguchi-Ishimae, M., Tojo, A., Morishita, K., Suzuki, K., Sato, Y., Kudoh, S., Tanaka, K., Setoyama, M., Nagamura, F., Asano, S. & Kamada, N. (1999). Fusion of ETV6 to neurotrophin-3 receptor TRKC in acute myeloid leukemia with t(12;15)(p13;q25). *Blood* **93**, 1355-1363.

Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D. & Kay, L. E. (1994). Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* **33**, 5984-6003.

Fields, S. & Song, O. (1989). A Novel Genetic System to Detect Protein-Protein Interactions. *Nature* **340**, 245-246.

Flanary, P., DiBello, P., Estrada, P. & Dohlman, H. (2000). Functional analysis of Plp1 and Plp2, two homologues of phosducin in yeast. *J. Biol. Chem.* **275**, 18462-18469.

Golub, T., Barker, G., Bohlander, S., Hiebert, S., Ward, D., Bray-Ward, P., Morgan, E., Raimondi, S., Rowley, J. & Gilliland, D. (1995). Fusion of the TEL gene on 12p13 to the AML1 gene on 21q22 in acute lymphoblastic leukemia. *Proc. Natl. Acad. Sci.* **92**, 4917-4921.

Golub, T., Barker, G., Lovett, M. & Gilliland, D. (1994). Fusion of PDGF Receptor Beta to a Novel ets-like Gene, tel, in Chronic Myelomonocytic Leukemia with t(5;12) Chromosomal Translocation. *Cell* **77**, 307-316.

Golub, T., Goga, A., Barker, G., Afar, D., McLaughlin, J., Bohlander, S., Rowley, J., Witte, O. & Gilliland, D. (1996). Oligomerization of the ABL Tyrosine Kinase by the Ets Protein Tel in Human Leukemia. *Mol. Cell. Biol.* **16**, 4107-4116.

Graves, B. (1998). Inner Workings of a Transcription Factor Partnership. *Science* **279**, 1000-1002.

Graves, B. & Petersen, J. (1998). Specificity within the ets family of transcription factors. *Adv. Cancer Res.* **75**, 1-55.

Grzesiek, S. & Bax, A. (1992a). Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.* **114**, 6291-6293.

Grzesiek, S. & Bax, A. (1992b). An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J. Magn. Reson.* **99**, 201-207.

Hahn, S. L., Criqui, P. & Wasylyk, B. (1997). Modulation of Ets-1 Transcriptional Activity by hUBC9, a Ubiquitin-Conjugating Enzyme. *Oncogene* **15**, 1489-1495.

Hannon, G., Demetrick, D. & Beach, D. (1993). Isolation of the Rb-related p130 Through its Interaction with CDK2 and Cyclins. *Genes Dev.* **7**, 2378-2391.

Hirel, P. H., Schmitter, M. J., Dessen, P., Fayat, G. & Blanquet, S. (1989). Extent of N-terminal methionine excision from Escherichia coli proteins is governed by the

side-chain length of the penultimate amino acid. *Proc. Natl. Acad. Sci.* **86**, 8247-8251.

Ho, J., B., B., Squire, J., Frank, D. & Barber, D. (1999). Fusion of the Ets Transcription Factor Tel to Jak2 Results in Constitutive Jak-Stat Signaling. *Blood* **93**, 4354-4364.

Hoffman, C. & Winston, F. (1987). A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of Escherichia coli. *Gene* **57**, 267-272.

Holland, P. M. & Cooper, J. A. (1999). Protein modification: docking sites for kinases. *Curr. Biol.* **9**, R329-R331.

Hollenberg, S., Sternglanz, R., Cheng, P. & Weintraub, H. (1995). Identification of a New Family of Tissue-Specific Basic Helix-Loop-Helix Proteins with a Two-Hybrid System. *Mol. Cell. Biol.* **15**, 3813-3822.

Iijima, Y., Ito, T., Oikawa, T., Eguchi, M., Eguchi-Ishimae, M., Kamada, N., Kishi, K., Asano, S., Sakaki, Y. & Sato, Y. (2000). A new ETV6 / TEL partner gene, ARG (ABL-related gene or ABL2), identified in an AML-M3 cell line with a t(1;12)(q25;p13) translocation. *Blood* **95**, 2126-2131.

Jousset, C., Carron, C., Boureux, A., Quang, C., Oury, C., Dusanter-Fourt, I., Charon, M., Levin, J., Bernard, O. & Ghysdael, J. (1997). A Domain of Tel Conserved in a Subset of ETS Proteins Defines a Specific Oligomerization Interface Essential to the Mitogenic Properties of the Tel-PDGFRb Oncoprotein. *EMBO J.* **16**, 69-82.

Keegan, L., Gill, G. & Ptashne, M. (1986). Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. *Science* **231**, 699-704.

Kim, C. A., Phillips, M. L., Kim, W., Gingery, M., Tran, H. H., Robinson, M. A., Faham, S. & Bowie, J. U. (2001). Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. *EMBO J.* **20**, 4173-4182.

Klambt, C. (1993). The Drosophila Gene Pointed Encodes two Ets-like Proteins Which are Involved in the Development of Midline Glial Cells. *Development* **117**, 163-176.

Knezevich, S., McFadden, D., Tao, W., Lim, J. & Sorensen, P. (1998). A novel ETV6-NTRK3 gene fusion in congenital fibrosarcoma. *Nat. Genet.* **18**, 184-187.

Kodandapani, R., Pio, F., Ni, C., Piccialli, G., Klemsz, M., McKercher, S., Maki, R. & Ely, K. (1996). A New Pattern for Helix-Turn-Helix Recognition Revealed by the PU.1 ETS-Domain-DNA Complex. *Nature* **380**, 456-460.

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283-292.

Kuno, Y., Abe, A., Emi, N., Iida, M., Yokozawa, T., Towatari, M., Tanimoto, M. & Saito, H. (2001). Constitutive kinase activation of the TEL-Syk fusion gene in myelodysplastic syndrome with t(9;12)(q22;p12). *Blood* **97**, 1050-1055.

Kwiatkowski, B., Bastian, L., Bauer, T., Tsai, S., Zielinska-Kwiatkowski, A. & Hickstein, D. (1998). The ets Family Member Tel Binds to the Fli-1 Oncoprotein and Inhibits Its Transcriptional Activity. *J. Biol. Chem.* **273**, 17525-17530.

Kwiatkowski, B., Zielinska-Kwiatkowski, A., Bauer, T. & Hickstein, D. (2000). The Ets Family Member Tel Antagonizes the Fli-1 Phenotype in Hematopoietic Cells. *Blood Cells, Molecules, and Diseases* **26**, 84-90.

Kyba, M. & Brock, H. W. (1998). The SAM domain of polyhomeotic, RAE28, and scm mediates specific interactions through conserved residues. *Dev. Genet.* **22**, 74-84.

Lacronique, V., Boureux, A., Valle, V., Poirel, H., Quang, C. T., Mauchauffe, M., Berthou, C., Lessard, M., Berger, R., Ghysdael, J. & Bernard, O. (1997). A Tel-Jak2 Fusion Protein with Constitutive Kinase Activity in Human Leukemia. *Science* **278**, 1309-1312.

Laemmli, U. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Larsen, T., Olson, A. & Goodsell, D. (1998). Morphology of protein-protein interfaces. *Structure* **6**, 421-427.

Lazarov, M., Martin, M., Willardson, B. & Elton, T. (1999). Human phosducin-like protein (hPhLP) messenger RNA stability is regulated by cis-acting instability elements present in the 3'-untranslated region. *Biochim. Biophys. Acta* **1446**, 253-264.

Leprince, D., Gegonne, A., Coll, J., de Taisne, C., Schneeberger, A., Lagrou, C. & Stehelin, D. (1983). A putative second cell derived oncogene of the avian leukemia retrovirus E26. *Nature* **306**, 395-397.

Li, R., Pei, H., Watson, D. K. & Papas, T. S. (2000). EAP1/Daxx interacts with ETS1 and represses transcriptional activation of ETS1 target genes. *Oncogene* **19**, 745-753.

Liang, H., Mao, X., Olejniczak, E., Nettesheim, D., Yu, L., Meadows, R., Thompson, C. & Fesik, S. (1994). Solution Structure of the ets Domain of Fli-1 when Bound to DNA. *Nat. Struct. Biol.* **1**, 871-876.

Lupas, A., van Dyke, M. & Stock, J. (1991). Predicting Coiled-Coils from Protein Sequences. *Science* **252**, 1162-1164.

Ma, J. & Ptashne, M. (1987). Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell* **48**, 847-853.

Mo, Y., Vaessen, B., Johnston, K. & Marmorstein, R. (1998). Structures of SAP-1 Bound to DNA Targets from the E74 and c-fos Promoters: Insights into DNA Sequence Discrimination by ETS Proteins. *Mol. Cell* **2**, 201-212.

Mo, Y., Vaessen, B., Johnston, K. & Marmorstein, R. (2000). Structure of the Elk-1-DNA Complex Reveals how DNA-Distal Residues Affect ETS Domain Recognition of DNA. *Nat. Struct. Biol.* **7**, 292-297.

Nakai, K. (2000). Protein sorting signals and prediction of subcellular localization. *Adv. Protein. Chem.* **54**, 277-344.

Nakai, K. & Kanehisa, M. (1992). A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* **14**, 897-911.

Nunn, M., Seeburg, P., Moscovici, C. & Duesberg, P. (1983). Tripartite structure of the avian erythroblastosis virus E26 transforming gene. *Nature* **306**, 391-395.

Papadopoulos, P., Ridge, S., Boucher, C., Stocking, C. & Wiedemann, L. (1995). The novel activation of ABL by fusion to an ets-related gene, TEL. *Cancer Res.* **55**, 34-38.

Peeters, P., Raynaud, S., Cools, J., Wlodarska, I., Grosgeorge, J., Philip, P., Monpoux, F., Van Rompaey, L., Baens, M., Van der Berghe, H. & Marynen, P. (1997). Fusion of TEL, the ETS-Variant Gene 6 (ETV6), to the Receptor-Associated Kinase JAK2 as a Result of t(9;12) in a Lymphoid and t(9;15;12) in a Myeloid Leukemia. *Blood* **90**, 2535-2540.

Peterson, A. J., Kyba, M., Bornemann, D., Morgan, K., Brock, H. W. & Simon, J. (1997). A domain shared by the Polycomb group proteins Scm and ph mediates heterotypic and homotypic interactions. *Mol. Cell. Biol.* **17**, 6683-6692.

Ponting, C. P. (1995). SAM: a novel motif in yeast sterile and Drosophila polyhomeotic proteins. *Prot. Sci.* **4**, 1928-1930.

Ribi, E., Perrine, T., List, R., Brown, W. & Goode, G. (1959). Use of Pressure Cell to Prepare Cell Walls from Mycobacteria. *Proc. Soc. Exp. Biol. Med.* **100**, 647-649.

Romana, S., Mauchauffe, M., Le Coniat, M., Chumakov, I., Le Paslier, D., Berger, R. & Bernard, O. (1995). The t(12;21) of acute lymphoblastic leukemia results in a tel-AML1 gene fusion. *Blood* **85**, 3662-3670.

Saito, H., Mamizuka, T., Tabeta, R. & Hirano, S. (1981). High resolution ^{13}C NMR spectra of chitin oligomers in aqueous solution. *Chem. Lett.*, 1483-1484.

Salomon-Nguyen, F., Della-Valle, V., Mauchauffé, M., Busson-Le Coniat, M., Ghysdael, J., Berger, R. & Bernard, O. (2000). The t(1;12)(q21;p13) translocation of human acute myeloblastic leukemia results in a TEL-ARNT fusion. *Proc. Natl. Acad. Sci.* **97**, 6757-6762.

Saraste, M., Sibbald, P. & Wittinghofer, A. (1990). The P-loop--a Common Motif in ATP- and GTP-Binding Proteins. *Trends in Bioc. Sci.* **15**, 430-434.

Schiestl, R. & Gietz, R. (1989). High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr. Genet.* **16**, 339-346.

Schultz, J., Ponting, C. P., Hofmann, K. & Bork, P. (1997). SAM as a protein interaction domain involved in developmental regulation. *Prot. Sci.* **6**, 249-253.

Slupsky, C., Gentile, L., Donaldson, L., Mackereth, C., Seidel, J., Graves, B. & McIntosh, L. (1998). Structure of the Ets-1 Pointed Domain and Mitogen-Activated Protein Kinase Phosphorylation Site. *Proc. Natl. Acad. Sci.* **95**, 12129-12134.

Smalla, M., Schmieder, P., Kelly, M., Ter Laak, A., Krause, G., Ball, L., Wahl, M., Bork, P. & Oschkinat, H. (1999). Solution structure of the receptor tyrosine kinase EphB2 SAM domain and identification of two distinct homotypic interaction sites. *Prot. Sci.* **8**, 1954-1961.

Stapleton, D., Balan, I., Pawson, T. & Sicheri, F. (1999). The crystal structure of an Eph receptor SAM domain reveals a mechanism for modular dimerization. *Nat. Struct. Biol.* **6**, 44-49.

Tatusova, T. & Madden, T. (1999). BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol. Lett.* **174**, 247-250.

Thanos, C. D., Faham, S., Goodwill, K. E., Cascio, D., Phillips, M. & Bowie, J. U. (1999a). Monomeric structure of the human EphB2 sterile alpha motif domain. *J. Biol. Chem.* **274**, 37301-37306.

Thanos, C. D., Goodwill, K. E. & Bowie, J. U. (1999b). Oligomeric structure of the human EphB2 receptor SAM domain. *Science* **283**, 833-836.

Thompson, C. C., Brown, T. A. & McKnight, S. L. (1991). Convergence of Ets- and notch-related structural motifs in a heteromeric DNA binding complex. *Science* **253**, 762-768.

Traut, T. (1994). The Functions and Consensus Motifs of Nine Types of Peptide Segments that Form Different Types of Nucleotide-Binding Sites. *Eur. J. Bioc.* **222**, 9-19.

Triezenberg, S. J., LaMarco, K. L. & McKnight, S. L. (1988). Evidence of DNA: protein interactions that mediate HSV-1 immediate early gene activation by VP16. *Genes Dev.* **2**, 730-742.

Tu, H., M Barr, Dong, D. L. & Wigler, M. (1997). Multiple regulatory domains on the Byr2 protein kinase. *Mol. Cell. Biol.* **17**, 5876-5887.

Virbasius, J. V., Virbasius, C. A. & Scarpulla, R. C. (1993). Identity of GABP with NRF-2, a multisubunit activator of cytochrome oxidase expression, reveals a cellular role for an ETS domain activator of viral promoters. *Genes Dev.* **7**, 380-392.

Vojtek, A., Hollenberg, S. & Cooper, J. (1993). Mammalian Ras Interacts Directly with the Serine/Threonine Kinase Raf. *Cell* **74**, 205-214.

Wada, A. (1976). The alpha-helix as an electric macro-dipole. *Adv. Biophys.* **9**, 1-63.

Watanabe, H., Sawada, J., Yano, K., Yamaguchi, K., Goto, M. & Handa, H. (1993). cDNA cloning of transcription factor E4TF1 subunits with Ets and notch motifs. *Mol. Cell. Biol.* **13**, 1385-1391.

Wells, L., Vosseller, K. & Hart, G. W. (2001). Glycosylation of nucleocytoplasmic proteins: signal transduction and O-GlcNAc. *Science* **291**, 2376-2378.

Werner, M., Clore, G., Fisher, C., Fisher, R., Trinh, L., Shiloach, J. & Gronenborn, A. (1997). Correction of the NMR structure of the ETS1/DNA complex. *J. Biomol. NMR* **10**, 317-328.

Wright, H. T. (1991). Nonenzymatic deamidation of asparaginyl and glutaminyl residues in proteins. *Crit. Rev. Biochem. Mol. Biol.* **26**, 1-52.

Wu, C., Leberer, E., Thomas, D. Y. & Whiteway, M. (1999). Functional Characterization of the Interaction of Ste50p with Ste11p MAPKKK in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* **10**, 2425-2440.

Yang, X., Khosravi-Far, R., Chang, H. Y. & Baltimore, D. (1997). Daxx, a novel Fas-binding protein that activates JNK and apoptosis. *Cell* **89**, 1067-1076.

Zhao, J., Hyman, L. & Moore, C. (1999). Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.* **63**, 405-445.

Zhu, X. & Craft, C. M. (2000). Modulation of CRX transactivation activity by phosducin isoforms. *Mol. Cell. Biol.* **20**, 5216-5226.