Functional characterisation of the Polycomblike protein of *Drosophila melanogaster*

A thesis submitted for the degree of Doctor of Philosophy

by

Sinead O'Connell, B.Sc (Hons)

Department of Genetics,
University of Adelaide,
Adelaide, SA, 5005,
Australia

December, 1999

‘Strive to achieve’

Julie ‘Bay Buddy’ Secombe,
Donna ‘Crackster’ Crack,
Michelle ‘MI’ Coulson,
(1995-1999)

Table of Contents

Abstract

The Polycomb Group (PcG) of genes are responsible for the epigenetic silencing of genes during *Drosophila* development. To date 14 members have been identified on the basis of a similar mutant phenotype. Epistasis experiments suggest that the encoded proteins are going to act in a multimeric complex to heritably repress gene transcription. Overlapping polytene binding patterns for some PcG proteins, co-immunoprecipitation experiments from embryonic extracts, and direct interactions between some members of the PcG further support the theory of a multimeric complex. However, it is not yet clear what the mechanism of PcG repression is, and an important first step in resolving this is to characterise the role of each PcG member and identify intra-complex interactions.

Polycomblike (*Pcl*) has been shown to be a key member of the PcG. Molecular characterisation of *Pcl* revealed the presence of two PHD fingers, a Cys_4-His-Cys_3 motif in the protein product. The region surrounding and including the PHD fingers is highly conserved with two mammalian homologues, and is termed the conserved domain (cDOM). Yeast two hybrid analysis has demonstrated an interaction between the cDOM of PCL and Enhancer of zeste (E(Z)), another key member of the PcG. Further characterisation of the interaction between the cDOM of PCL and E(Z) using co-immunoprecipitations from embryonic extracts, confirmed their association *in vivo*. *In vitro* mutagenesis was used to demonstrate that the interaction between PCL and E(Z) is mediated through the PHD finger motifs of PCL.

Using an *in vivo* tethering assay, PCL has been shown to be able to initiate heritable silencing of a reporter gene when tethered to DNA via a GAL4 DNA binding domain. This assay was used to identify functionally important regions of PCL. The amino terminus of PCL was sufficient to initiate heritable repression and the repression conferred by both the GAL-PCL fusion protein and the GAL-Amino fusion protein was dependent on the endogenous PcG. The amino terminus was shown by far western analysis to be capable of homotypic interactions. Yeast two hybrid analysis identified a possible interaction between the amino terminus of PCL and a fragment of PC.

The C-terminus of PCL was shown to interact with PHO and SU(Z)2 in a yeast two hybrid assay.

The work described in this thesis has identified key interactions between PCL and other members of the PcG and suggested a model for the role of PCL within the PcG.

Declaration

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and to the best of my knowledge contains no material previously published or written by another person, except where due reference has been made in the text.

This thesis represents my own work with the exception of the following work described in Chapter 4;

Tory McGrath generated the *hb*-GAL-Amino, *hb*-GAL-Carboxy, *hb*-GAL-PHD and *hb*-GAL-Amino+CarboxyΔPHD constructs and injected them into *Drosophila* embryos. I then isolated transformant flies and determined which chromosome the insert was on.

I give full consent for this copy of my thesis to be made available for loan and photocopying when deposited in the University library.

Sinead O'Connell
December, 1999

Acknowledgments

Well what can I say, a Ph.D. can be hell and parts of it have been. Most of it, however, has been incredibly entertaining. This is due entirely to the people that I have had the pleasure in interacting with in the last few years. I owe a great deal to the people in the Saint lab, both past and present, who have provided me with an interactive learning environment as well as a fun place to work. Specifically I would like to thank the following people.

Firstly Rob, my supervisor. You have provided an environment where people can work and learn independently, and go on extended overseas holiday...oooppss...work related trips:). Thank you.

Julie, bay buddy:) what can I say, you made our little corner of the lab very entertaining. You are a fantastic friend and I miss you terribly, but look forward to catching up with you in Seattle and San Diego and where ever else we feel like holidaying in the US:)

Tory for keeping me entertained both here and in the US and trying to train me in the American ways. Donna for being a great friend both in and out of the lab. Thanks also to Jason for cooking a great barbie:) I am so pumped you guys are going to be parents!!!!

Thanks to Michelle I, Mr Mike and Dan for putting up with all the little questions I continuously asked. Michelle II for making the Saint Lab an entertaining place to work, keep on singing babe you have a great voice:) Tatiana for being a fantastic Russian teacher:) Michael Clarkson for being so funny and for leading an interesting life☺. Thanks also to Tod for allowing me to share his bench! and Deb, the guest 213er! Janet for becoming my new bay buddy!! and also Kym, Louise, Jane and Greg for making our end of the lab entertaining.

Special thanks to Sonia. You are an incredible person and a truly special friend.

Thanks also to Bec for keeping me highly amused both here and in the US!!

Thank you to the various people around the Genetics Department who made the place entertaining; PhD students, honours students and staff. Specifically Paul for his help with everything!! I promise I am not as hopeless with equipment as you think I am, I just know you love to help people☺ and Daniela for allowing me to share her office during the write up.

Thanks to Velta for her friendship throughout the years, and of course for the chocolate (on behalf of everyone in the Saint lab) which we never ate in the lab:) I will be sure to send you a postcard from every place Julie and I visit:)

To all my friends and family around the world, thank you for all the emails and all the places to crash. Specifically Ann, Carol-ann, Mary, Bec, Aunty Breda, Clare, Tory, Steve and Mike.

Thanks also to Hugh Brock, Micahel Kyba, Judy Kassis, Lesley Brown, Rick Jones, Aidan Peterson, Peter Harte, and all the PcG people around the world I was able to interact with during the course of this work and for providing reagents.

Lastly thank you to Mum, Dad and Aishling for everything. I know I couldn't have done it without you.

Chapter 1: Introduction

How can the expression of approximately 100,000 genes in the human genome be regulated spatially and temporally throughout development? This question has intrigued developmental biologists for many years. Although there is as yet no clear answer, significant progress has been made in identifying the regulatory mechanisms through which gene expression is controlled. The regulation of gene expression involves both the activation and repression of genes by transcription factors and an understanding of how these transcription factors interact with the transcription machinery is slowly emerging. Epigenetic mechanisms also exist, whereby the expression state of a gene is maintained in daughter cells after cell division. Less is known about these mechanisms that maintain genes in the 'on' or 'off' state after the initial developmental decisions have been made. This is however an important problem in the development of a multicellular organism as the developmental cues that determine the 'on' or 'off' state of a gene are often only present transiently early in embryogenesis, yet mutational analysis has revealed that the correct expression pattern of developmentally important genes is required throughout development.

Although several examples of epigenetic mechanisms exist, including imprinting (reviewed by Brannan and Bartolomei, 1999), X-inactivation (reviewed by Latham, 1996) and position effect variegation (PEV) (reviewed by Wakimoto, 1998), this discussion will focus specifically on the epigenetic mechanisms that control repression of the homeotic loci of *Drosophila*.

1-1 The regulation of homeotic gene expression

The correct spatial and temporal expression of the homeotic genes is important in specifying the anterior/posterior fate of all developing metazoans. The homeotic genes, collectively referred to as HOX genes, are clustered into two complexes; the Antennapedia (ANTP-C) and Bithorax (BX-C) complexes. The regulation of HOX gene expression has been best characterised in *Drosophila melanogaster* and involves two distinct processes, initiation of the correct pattern of expression early in embryogenesis and maintenance of this expression pattern throughout subsequent development.

A cascade of events leads to the correct initiation of homeotic gene expression. Prior to fertilisation, maternal factors are deposited into the egg and specify the anterior-posterior axis crucial for initiating a gene regulatory cascade after fertilisation. These maternal factors contribute spatial information and after fertilisation act to specify the broad expression domains of the gap genes which then regulate the expression of the pair-rule genes. It is the combination of gap and pair-rule gene products which initiate homeotic gene expression (reviewed by Ingham 1988). Generally, the gap gene products act to repress homeotic

genes, while the pair-rule genes act to initiate homeotic gene expression. Both the gap and pair-rule gene products encode DNA binding proteins and compete for binding sites on the promoter regions of homeotic genes. One of the best studied examples of the regulation of a homeotic gene is that of the *Ultrabithorax* (*Ubx*) gene. The gap gene product, Hunchback (HB), is expressed in the anterior half of the blastoderm embryo and represses *Ultrabithorax* (*Ubx*) expression in this region of early embryos (Qian *et al.*, 1993). The products of the *fushi tarazu*, *engrailed* and *twist* genes act to initiate the expression of the *Ubx* gene in the posterior region of the embryo (Qian *et al.*, 1993).

The early expression of the gap and pair-rule genes ceases at approximately stage 10, yet correct expression of the homeotics needs to be maintained throughout development. Two groups of genes act to maintain the correct expression of the homeotics after the degradation of the gap and pair-rule proteins; the trithorax (trxG) and Polycomb (PcG) Groups. Mutations in trxG members result in reduced homeotic gene expression (Sedkov *et al.*, 1994; Breen and Harte, 1993). This reduction is not seen until mid to late in embryogenesis, consistent with the trxG being involved in the maintenance of expression of the homeotic loci, and not in their initiation. In contrast, mutations in PcG genes result in ectopic expression of the homeotic genes, consistent with a role in the maintenance of repression of these genes throughout embryogenesis (Simon, 1995; Pirrotta, 1997 and references therein). The role of the PcG and the mechanisms through which it heritably represses homeotic gene expression is the focus of the rest of this chapter.

1-2 The Polycomb Group

Polycomb Group genes were initially identified from mutant phenotypes which resembled gain of function homeotic mutations (Lewis, 1978; Jurgens, 1985). The PcG mutant homeotic transformation phenotype results from the derepression of homeotic genes outside of their normal domains of expression (see Pirrotta, 1995 for a review). To date, 14 members have been identified, ten of which have been characterised at the molecular level; *Polycomb (Pc)* (Lewis, 1978; Paro and Hogness, 1991), *Polycomblike (Pcl)* (Duncan, 1982; Jurgens, 1985; Lonie *et al.*, 1994), *Posterior sex combs (Psc)* (Jurgens, 1985; Brunk *et al.*, 1991), *Additional sex combs* (*Asx)* (Jurgens, 1985; Sinclair *et al.*, 1998), *Sex combs on midleg* (*Scm*) (Jurgens, 1985; Bornemann *et al.*, 1996), *Enhancer of zeste (E(z))* (Jones and Gelbart, 1990; Jones and Gelbart, 1993), *extra sex combs* (*esc*) (Struhl, 1981; Simon *et al.*, 1995), *Pleiohomeotic* (*pho*) (Girton and Jeon, 1994; Brown *et al.*, 1998), *polyhomeotic (ph)* (Dura *et al.*, 1985; DeCamillis *et al.*, 1992), *Sex combs extra* (*Sce*) (Breen and Duncan, 1986), *super sex combs (sxe)* (Ingham, 1984), *multi sex combs (mxc)* (Santamaria and Randsholt, 1995) and *cramped (crm)* (Yamamoto *et al.*, 1997). Three additional genes, *Suppressor of zeste two* (*Su(z)2*)) (Adler *et al.*, 1989; Brunk *et al.*, 1991), *dMi-2* (Kehle *et al.*, 1998) and *Enhancer of Polycomb* (*E(Pc)*) (Sato *et al.*, 1984;

Stankunas *et al.*, 1998) have been identified as genetic interactors of members of the PcG. Mutations in these genes do not result in a homeotic phenotype, but do enhance the homeotic phenotype of other PcG members.

1-2.1 Molecular characterisation of the PcG proteins

Drosophila *PcG proteins*

Many of the PcG proteins that have been characterised at the molecular level contain previously characterised motifs. PC contains a chromodomain which is also present in HP1, a protein associated with heterochromatin (Paro and Hogness, 1991). A variety of other proteins thought to be involved in epigenetic mechanisms also contain chromodomains suggestive of the evolution of a domain involved in chromatin remodelling (James and Elgin, 1986; Cavalli and Paro, 1998). Mutations in the chromodomain of PC or HP1 affect their localisation to chromosomes. No DNA binding activity of PC or HP1 has been reported, so this domain is presumably necessary for localisation of these proteins to chromosomes via protein-protein interactions (Messmer *et al.*, 1992; Platero *et al.*, 1995).

The PH protein contains an SPM domain, named after the three founding *Drosophila* proteins containing this domain; SCM, PH, and Lethal (3) Malignant Brain Tumour (L(3)MBT). In addition to the SPM domain, PH contains a Zn finger region, which is also found in SCM (Bornemann *et al.*, 1996). The SPM domain is important in mediating the homotypic interactions of PH and SCM as well as heterotypic interactions between PH and SCM (Peterson *et al.*, 1997; Kyba and Brock, 1998b). No function has yet been attributed to the Zn finger regions of either protein, which do not resemble previously characterised DNA binding domains. The SCM protein also contains a region of homology that is present in l(3)MBT, although the function of this region in either protein currently remains unknown (Bornemann *et al.*, 1996).

PSC and SU(Z)2 contain RING finger domains which exhibit a high level of sequence similarity (Brunk *et al.*, 1991). In addition, the types of amino acids found in the C-terminus of both proteins are similar. E(Z) contains an interesting protein motif called the SET domain which is also found in TRX, a trxG member, SU(VAR)3-9, a suppressor of position effect variegation and a variety of proteins that function in modulating transcription (Jones and Gelbart, 1993; Jenuwein *et al.*, 1998). The presence of the SET domain in both TRX and E(Z) raises the possibility that this domain mediates the antagonistic repressive versus activating functions of the PcG and trxG respectively (Jones and Gelbart, 1993).

PCL contains two PHD finger domains, Cys_4-His-Cys_3 motifs, which are also found in multiple copies in TRX and several other proteins involved in chromatin structure and regulation of gene expression through chromatin-mediated mechanisms (Aasland *et al.*,

1995). It is not known whether the PHD finger motif is responsible for mediating interactions with other proteins or with DNA. It is possible that the PHD finger motifs of PCL and TRX mediate the antagonistic functions of the trxG and PcG (Lonie *et al.*, 1994; Aasland *et al.*, 1995).

ESC contains seven WD-40 motifs which have been shown to be important in mediating an interaction with E(Z) (Simon *et al.*, 1995; Jones *et al.*, 1998; Tie *et al.*, 1998). PHO contains a Zn finger DNA binding domain and has been shown to bind *in vitro* to Polycomb Response Elements (PREs), regulatory sequences known to be required for PcG-mediated repression. To date PHO remains the only cloned member of the PcG that has been shown to bind DNA directly (Brown *et al.*, 1998). Table 1.1 summarises the features of the cloned PcG members.

PcG homologues

The study of mammalian, plant and nematode homologues of the *Drosophila* PcG has revealed putative functionally important domains within many of the PcG proteins. Mammalian homologues have been identified for many of the PcG members including PC, PH, PCL, dMi-2, SCM, PSC, SU(Z)2, PHO, ESC and E(Z) (reviewed in van Lohuizen, 1999). Mutational studies in several of these genes show a homeotic derepression phenotype analogous to that observed in *Drosophila* suggesting a similar evolutionarily conserved mechanism (see van Lohuizen 1998; Weigel, 1997 for reviews). Interestingly, only homologues of ESC and E(Z) have been found in *Caenorhabditis elegans* and *Arabidopsis thaliana* suggestive of a PcG independent function for these two proteins, although the identification of other PcG proteins in these organisms cannot be ruled out (Weigel, 1997; van Lohuizen, 1999). Surprisingly, although a trxG-like complex, which contains previously characterised trxG members, has been identified in yeast, no PcG homologues have been identified The absence of PcG members suggests that a separate complex is responsible for epigenetic silencing in this organism. Table 1.2 is a summary of PcG homologues.

1-2.2 PcG proteins interact to form a multimeric complex

The majority of PcG genes were identified genetically on the basis of their haplo-insufficient gain of function homeotic phenotype. Given that all members are defined by a homeotic phenotype, it was of interest to determine whether members of the PcG act independently of one another, or in the same complex or pathway. Genetic epistasis experiments were performed to determine the relationship of each of the members to each other (Jurgens, 1985; Campbell *et al.*, 1995). Double PcG mutants have a more severe phenotype than the sum of the single mutants, indicating a synergistic interaction between members of the PcG (Jurgens, 1985; Cheng *et al.*, 1994; Campbell *et al.*, 1995). The PcG proteins were

Table 1.1: A summary of the cloned PcG members and the protein motifs they encode. See text for a discussion of these domains.

Cloned Polycomb Group members	Known protein motifs	Function
Additional sex combs	cysteine cluster	unknown
cramped	none	
Enhancer of zeste	SET domain,	unknown
	cysteine rich region (CXC)	unknown
extra sex combs	WD-40 motifs,	mediate interaction with E(Z)
pleiohomeotic	Zn finger DNA binding domain	binds DNA
Polycomb	chromodomain	unknown-presumed protein/protein interaction domain
Polycomblike	PHD fingers	unknown-presumed DNA binding or protein-protein interaction domain
polyhomeotic	Zn finger region	unknown
	SPM domain	mediates homotypic interactions and heterotypic interactions with SCM
Posterior sex combs	RING finger	mediates interaction with PC along with helix-turn-helix region
Sex combs on midleg	SPM domain	mediates homotypic and heterotypic interactions with PH

Table 1.2: A summary of the known homologues of the *Drosophila* PcG genes.

Drosophila gene	mammalian homologue*	other homologues	References
Additional sex combs	human EST	none known	Sinclair *et al.*, 1998
cramped	none known	none known	Yamamoto *et al.*, 1997
Enhancer of zeste	*EZH1/Enx1, EZH2/Enz2*	*MEDEA, Curly Leaf-A. thaliana, mes-2-C.elegans*	Abel *et al.*, 1996; Chen *et al.*, 1996; Houbert *et al.*, 1996; Lible *et al.*, 1997; Goodrich *et al.*, 1997; Grossniklaus *et al.*, 1998; Holdman et. al. 1998
extra sex combs	*hEED/eed*	*mes-6-C.elegans*	Schumacher *et al.*, 1996; Korf *et al.*, 1998
pleiohomeotic	*hYY1/mYY1*	none known	Brown *et al.*, 1998
Polycomb	*HPc/M33, HPc2/MPc2*	*Xpc1X-laevis*	Pearce *et al.*, 1992; Alkema *et al.*, 1997; Satjun *et al.*, 1997; Strouboulis *et al.*; 1999
Polycomblike	*PHF1/mPHF1, hMtf2/Mtf2*	none known	Coulson *et al.*, 1998 and Chapter 3
polyhomeotic	*HPH1/Mph1(rae28), HPH2/Mph2*	none known	Nomura *et al.*, 1994; Gunster *et al.*, 1997
Posterior sex combs	*BMI-1/bmi-1,* *MEL-18/mel-18*	*Xbmi-1- X. laevis*	van Lohuizen *et al.*, 1991; Brunk *et al.*, 1991; Ishida *et al.*, 1993; Reijen *et al.*, 1995
Sex combs on midleg	*SCML1* (human), *SCML2* (human)	none known	van de Vosse *et al.*, 1998; Montini *et al.*, 1999

* human homologues is written first, followed by the murine orthologue.

therefore thought to act either in a similar pathway that leads to repression of target genes or a multimeric protein complex which represses transcription. Several lines of evidence support the latter model. Firstly, immunostaining of salivary gland polytene chromosomes show identical binding patterns for PC, PH and PCL on approximately 100 sites (DeCamillis *et al.,* 1992; Franke *et al.,* 1992; Lonie *et al.,* 1994). PSC is found at 83 sites, 63 of which coincide with PC/PH/PCL (Rastelli *et al.,* 1993). ASX binds to 90 sites on polytene chromosomes, 63 of which overlap with the PC/PH/PCL binding sites (Sinclair *et al.,* 1998). The binding pattern of SCM shares a 90% overlap with that of PC/PH/PCL (Peterson *et al.,* 1997). E(Z) binds to 44 sites, 43 of which overlap with known PcG binding sites (Carrington and Jones, 1996) and shares an identical distribution to that of ESC (Tie *et al.,* 1998). SU(Z)2 and E(PC) also share a significant overlap in polytene binding patterns to that of other PcG members. Using a temperature sensitive allele of *E(z),* the binding of PSC, SU(Z)2 and PH to the majority of their polytene binding sites, was shown to be dependent on E(Z) function (Rastelli *et al.,* 1993). Although the resolution of a binding site on polytene chromosomes is predicted to be ~100kb, the overlap in distribution was predicted to be greater than chance alone supporting the theory that the PcG proteins form a multimeric complex (Rastelli *et al.,* 1993).

More significant than the overlapping polytene distribution of the PcG members is the growing body of evidence that the PcG proteins (both the *Drosophila* proteins and their mammalian counterparts) interact with each other. PC/PH/PSC/SCM are able to be co-immunoprecipitated from nuclear extracts (Franke *et al.,* 1992; Kyba and Brock, 1998a; Shao *et al.,* 1999). Direct *in vitro* interactions have been demonstrated for PC and PSC via the helix-turn-helix (HTH) and RING finger domains of PSC (Kyba and Brock, 1998a). Similarly PC and PSC homologues from *Xenopus laevis*; Xpc and Xbmi-1, have been shown to interact directly with each other (Reijen *et al.,* 1995). PSC also interacts directly with PH via the HTH region of PSC and the H1 domain of PH (Kyba and Brock, 1998a). The interaction between PSC and PH has also been demonstrated in the mammalian homologues; BMI-1 and HPH1 and HPH2 (Gunster *et al.,* 1997). SCM has been shown to interact directly with PH, via the SPM domain of both proteins (Peterson *et al.,* 1997). ESC and E(Z) have also been shown to be associated *in vivo* and to interact directly *in vitro* (Tie *et al.,* 1998; Jones *et al.,* 1998). The mammalian equivalents of E(Z), EZH1 and Ezh2 have been shown to interact with Eed1, the mammalian homologue of ESC (Denisenko *et al.,* 1998; van Lohuizen *et al.,* 1998). This interaction is mediated through the conserved WD-40 repeats of ESC and an N-terminal region of E(Z). Other protein-protein interactions are likely to be identified as more members of the group are cloned and their interactors identified.

In addition to the heterotypic interactions between members of the PcG, there is also *in vitro* evidence to suggest that many PcG proteins form homotypic interactions. SCM has been

shown to form homotypic interactions *in vitro*, via the SPM domain, although it remains to be established whether this self-association is relevant *in vivo* (Peterson *et al.*, 1997). Likewise, the *Drosophila* and mammalian PH proteins have also been shown to self-associate *in vitro* through the SPM domain (Alkema *et al.*, 1997; Peterson *et al.*, 1997).

Co-immunoprecipitation experiments using nuclear extracts have led to the isolation of 2MDa complexes containing PC, PH and 10-15 other proteins (Franke *et al.*, 1992). Recent co-immunoprecipitation experiments have led to the isolation of a similar size complex containing PC/PH/SCM/PSC and several other proteins (Shao *et al.*, 1999). This complex, which has been termed Polycomb Repressive Complex 1 (PRC1), is capable of blocking the remodelling of nucleosomes by the activating SWI/SNF complex (see Hartzof and Winston, 1997 for a review of the SWI/SNF complex). Surprisingly, E(Z), PCL and ESC were not found to be part of this massive protein complex, raising the possibility of multiple PcG complexes *in vivo*, although the possibility that the interaction of these members of the group was not strong enough to survive the purification procedure could not be ruled out.

Further evidence for the formation of multiple PcG complexes comes from the polytene binding patterns of members of the group. Although there are a number of sites where complete overlap of the PcG occurs, there are a larger number of sites where only a subset of PcG proteins are bound. Immunoprecipitation of cross-linked chromatin has also demonstrated the formation of distinct PcG complexes. Strutt and Paro (1997) showed that PSC, PC and PH are associated with identical regulatory elements of the *engrailed* gene, but are differentially distributed on the *engrailed*-related gene *invected*, again raising the possibility that at each target site a distinct complex forms to repress gene transcription. The distribution of other PcG proteins on both the *engrailed* and *invected* regulatory regions was not determined. Isolating genes that are under PcG control and performing similar cross-linking experiments using the full repertoire of PcG antisera available, should shed further light on the formation of multiple PcG complexes at different target loci. The likely role of multiple complexes is yet to be investigated and it is not known whether the complexes use different mechanisms to repress target gene expression.

1-3 Key questions in PcG research

It is now well established that the PcG acts through the formation of multimeric protein complexes. Research in this field is now focussing on two main questions. How are the PcG proteins recruited to target loci and what is the mechanism through which they heritably represses transcription? These questions will be discussed separately.

1-3.1 Recruitment of the PcG to target loci

All characterised PcG members are expressed ubiquitously throughout development (Bornemann *et al.*, 1996; Frank *et al.*, 1992; Jones and Gelbart, 1993; Kehle *et al.*, 1998; Lonie *et al.*, 1994; Martin and Adler, 1993; Sinclair *et al.*, 1998; Stankunas *et al.*, 1998; Yamamoto *et al.*, 1997). Their expression pattern therefore confers no positional information which could be interpreted by the developing embryo. This is in contrast to the homeotic genes which are differentially expressed along the anterior/posterior axis. In any given cell, some homeotic genes are repressed by the PcG whilst others remain active. The PcG must somehow be recruited to different target genes in each cell type.

PcG targeting to the Ubx *gene*

The regulation of the *Ubx* gene is the best studied example of recruitment of the PcG to a target gene. As discussed above, HB is required to repress *Ubx* in the anterior half of the blastoderm embryo. HB binding sites have been identified in the *Ubx* promoter and are required for repression of *Ubx* (Qian *et al.*, 1991). However mid-way through embryogenesis the PcG takes over the role of repressor as HB degrades. The presence of HB at the *Ubx* locus made it a candidate for the recruitment of the PcG complex. Several experiments were performed which supported this. The BXD enhancer is a regulatory region of the *Ubx* gene which directs expression of *Ubx* in a head-to-tail pattern throughout embryogenesis. When this enhancer is linked to HB binding sites or a region of the *Ubx* gene containing HB binding sites, expression from the BXD enhancer becomes restricted to a *Ubx* domain in embryos (Müller and Bienz, 1991; Zhang *et al.*, 1991; Zhang and Bienz, 1992). This pattern is maintained late in embryogenesis and repression in the anterior half of the embryos is dependent on PC function (Zhang and Bienz, 1991). Further evidence for the involvement of HB in recruiting the PcG was provided by Müller (1995). He showed that the requirement for HB in initiating silencing on the BXD enhancer could be bypassed when PC was artificially tethered to DNA via a GAL4-DNA binding domain early in embryogenesis. The silencing conferred by this GAL-PC fusion protein was also heritable and maintained long after the GAL-PC fusion had been degraded. This heritable silencing was shown to be dependent on the function of the endogenous PcG suggesting that GAL-PC was recruiting endogenous PcG proteins to initiate heritable repression. It must be noted though that heritable silencing of the reporter gene was only conferred by the GAL-PC fusion when regulatory sequences from homeotic genes were located adjacent to the GAL4 binding sites, suggesting that the maintenance of repression required specific *cis*-acting sequences. These maintenance *cis*-regulatory elements will be discussed further below.

Given that HB is responsible for recruiting the PcG to the *Ubx* gene, it would be expected that HB either directly interacts with one or more PcG proteins or the interaction between HB and the PcG occurs via intermediary proteins which interact with both HB and PcG proteins.

Currently there is no evidence to suggest that HB interacts directly with any cloned PcG member. In order to identify the 'missing link' between the PcG and HB, HB was used as a 'bait' in a yeast two hybrid assay (Kehle *et al.,* 1998). Six HB interacting proteins (hips) were identified and of these, three interacted exclusively with HB (hip57, hip66 and hip76), and not with three other proteins with which they were tested. Two of the other interactors (hip11 and hip34), interacted not only with HB but also with PC and two other proteins involved in early embryogenesis, Bicoid and Oskar. The third interacted only with HB and PC (hip7). The interaction of hip7 with both HB and PC provides a direct link between HB and the PcG. hip76 displayed the strongest interaction with HB, and showed high sequence similarity to human Mi-2 (Seeling *et al.,* 1995) and was therefore named dMi-2. The region of HB responsible for mediating the interaction with dMi-2 had previously been shown to be important in the repression of BX-C genes.

Surprisingly the expression of *Ubx* and *Abdominal-B (Abd-B)* was completely normal in *dMi-2* mutants. Whether or not this indicates a functional redundancy for the role of dMi-2 is not known, but hip7, which interacts with both PC and HB, may substitute for the function of dMi-2 in recruiting the PcG to the *Ubx* locus. Embryos mutant for both *hb* and *dMi-2* showed a stronger derepression of the anterior boundary of *Ubx* than that of *hb* mutants alone. A synergistic interaction was also observed between *dMi-2* and several PcG mutants. In embryos mutant for *dMi-2* and either *Pc, Pcl* or *Psc,* derepression of both *Ubx* and *Abd-B* was observed. The cuticle phenotypes of these double mutants were also examined and found to resemble that of other double PcG mutant combinations.

These results suggest that dMi-2 functions in PcG repression in a manner similar to that of E(PC) and SU(Z)2, which act to enhance the homeotic derepression phenotype of other PcG mutants, but by themselves show no homeotic phenotype. This suggests that dMi-2 is unlikely to be a key member of the PcG, but instead acts to recruit the PcG to target loci. However this interaction between HB and a genetic interactor of the PcG provides an important functional link between HB repression and the initiation of PcG repression. In addition to this, there has also been some suggestion that dMi-2 may temporarily repress target gene transcription until the PcG complex has been recruited (van Lohuizen, 1999). The mammalian homologue of dMi-2 has been shown to be a member of a complex that has both histone deacetylase (HDA) and nucleosome remodelling activity (Zhang *et al.,* 1998). HDA activity has been shown to mediate repression of gene expression through the deacetylation of histone subunits (Pazin and Kadonaga, 1997).

Considering that HB only functions to recruit the PcG to the *Ubx* gene, how then is the PcG recruited to other target genes? It may be that other gap proteins are responsible for targeting the PcG. Evidence for this comes from studies of the mammalian C-terminal Binding Protein (CtBP), which has been shown to interact with HPC2, a human homologue of

Drosophila PC. dCtBP, which acts as a corepressor at a number of target loci, has been shown to interact with the *Drosophila* pair-rule segmentation protein Hairy and the gap segmentation protein Knirps (Poortinga *et al.,* 1998; Nibu *et al.,* 1998). It is possible that these proteins interact with dCTBP which then recruits the PcG to target loci.

The role of esc *in initiating PcG-mediated gene silencing*

Genetic evidence has long suggested a unique role for *esc* in establishing PcG-mediated silencing (Struhl, 1981; Struhl and Brower, 1982). Embryos lacking maternal *esc* only develop to first instar larvae and show extreme phenotypic transformation of abdominal and thoracic segments, in addition to some head segments, which develop structures characteristic of the eighth abdominal segment (Struhl, 1981). Using a temperature sensitive allele, *esc* has been shown to be important in early embryogenesis and dispensable thereafter (Struhl, 1981; Struhl and Brower, 1982). This is in contrast to other PcG members whose presence is required throughout development. Heatshock experiments further confirm that *esc* is required early in embryogenesis and the function of ESC in initiating silencing at target loci is dependent on its WD-40 repeats (Simon *et al.,* 1995). These repeats have been shown to be important in mediating an interaction with E(Z) (Jones *et al.,* 1997; Tie *et al.,* 1997), suggesting that the interaction between E(Z) and ESC is important in initiating PcG-mediated repression. This possibility is supported by the observation that the *esc* maternal effect phenotype is enhanced when the dose of *E(z)* is either increased or decreased, demonstrating that ESC function is sensitive to the levels of E(Z) protein (Campbell *et al.,* 1995).

It has been proposed that the ESC/E(Z) complex is responsible for initiating PcG-mediated silencing, while another group of PcG members (at least PC, PH and PSC) are responsible for maintaining repression (van Lohuizen *et al.,* 1998). Evidence for this comes from co-immunoprecipitation experiments which were unable to detect an association between the mammalian homologues of ESC/E(Z) and Mph1, Bmi1, Mel-18, M33 and Mpc2 (van Lohuizen *et al.,* 1998). However, this model does not take into account the observation that E(Z) is required throughout development and not just early in embryogenesis when targeting of the PcG is occurring (Campbell *et al.,* 1995; Rastelli *et al.,* 1993; Carrington and Jones, 1996). Perhaps E(Z) is required both early in embryogenesis to target PcG complexes and throughout development to stabilise these complexes. Alternatively, it could be a member of both the PcG initiating (PcGi) and PcG maintenance (PcGm) complexes. Evidence to support its role in stabilising the association of PcG complexes with chromatin comes from experiments using the temperature sensitive allele of *E(z).* At the restrictive temperature, PSC, SU(Z)2 and PH dissociate from salivary gland polytene chromosomes (Rastelli *et al.,* 1993) and the chromosomes appear decondensed, not only supporting a role for E(Z) in binding of the PcG to chromosomes but also suggesting that E(Z) functions to maintain

chromosomal integrity. However the absence of E(Z) from PRC1 in nuclear extracts (Shao *et al.,* 1999) and the absence of any reported interactions between E(Z) and other PcG members, namely PC, PH, PSC or SCM, which are present in PRC1 argue against a role for E(Z) as an integral part of the PcGm complex (Shao *et al.,* 1999).

The role of Pleiohomeotic in recruiting PcG complexes to target loci.

The recent cloning of *pho* has shed further light on how the PcG may be recruited to target genes (Brown *et al.,* 1998). PHO was identified in a screen for proteins bound to a PRE from the *engrailed* locus, a gene known to be regulated by the PcG (Dura and Ingham, 1988; Moazed and O'Farrell 1992). PHO contains four Zn finger motifs, most similar to those of mammalian Yin Yang 1 (YY1), an evolutionary conserved transcription factor (reviewed in Shi *et al.,* 1997). The presence of a DNA binding domain in PHO raises the obvious possibility that the PcG is recruited to DNA or held at particular loci via the binding of PHO. PHO binding sites would therefore be expected to be present in known PREs and mutations of these sites should abolish PcG-mediated silencing. Evidence to support this comes from analysis of a 1.6kb PRE fragment taken from the *Ubx* gene. It had been shown previously that when this fragment is placed upstream of a *lacZ* reporter gene it confers PcG dependant repression in a pattern similar to that of the endogenous *Ubx* gene (Chan *et al.,* 1994). This transgenic PRE also creates a novel polytene binding site for PC *in vivo* (Chan *et al.,* 1994). Sequence analysis of this PRE revealed the presence of six consensus PHO binding sites (Fritsch *et al.,* 1999). Mutation of these sites abolished PHO binding *in vitro* and, importantly, abolished PRE function *in vivo*. Mutations in *pho* show slight misexpression of the endogenous *Ubx* gene and this misexpression is enhanced by a mutation in *Pc,* indicating that the endogenous *Ubx* gene is regulated by PHO (Fritsch *et al.,* 1999).

The recruitment of the PcG to target genes cannot be as simple as PHO recognising and binding to the *Ubx* gene to initiate heritable repression through the recruitment of other PcG proteins. *pho*, like all other identified PcG members is expressed ubiquitously throughout embryogenesis and is therefore present in cells expressing *Ubx* and in cells where *Ubx* is silent (Brown *et al.,* 1998). Therefore PHO must itself be recruited to *Ubx* by some other DNA binding protein (perhaps by HB) or a cofactor that confers specificity. Alternatively, PHO may bind to the *Ubx* gene in cells where *Ubx* is both expressed and repressed but it can only recruit the PcG in cells where *Ubx* is repressed. It has been postulated, but remains to be proven, that the HB/dMi-2 complex deacetylates the surrounding chromatin and this could trigger the binding of PHO to *Ubx*. Alternatively, PHO could be bound and the deacetylation of chromatin is a prerequisite for the recruitment of other PcG members by PHO (Fritsch *et al.,* 1999). Interestingly, YY1 has been shown to mediate transcriptional repression through its interaction with RPD3, the mammalian homologue of a yeast histone

deacetylase (Yang *et al.*, 1996; Rundlett *et al.*, 1996). However, no HDA activity has been detected in PcGi or PcGm complexes from mice, suggesting that there is no intrinsic HDA activity in PcG complexes (unpublished data cited in van Lohuizen, 1999).

1-3.2 Maintenance of PcG silencing

Silencing by the PcG not only depends on recruitment of the PcG to target loci, but also requires the active maintenance of silencing throughout cell division. This maintenance of PcG silencing activity has been shown to be dependant on both a PRE and PcG proteins. In experiments carried out by Busturia *et al.* (1997), a 725bp fragment (MCP725) taken from the *Abd-B* gene was shown to mediate silencing in proliferating cells of the imaginal discs. Silencing by this element was dependent on endogenous *Pc* and *Pcl*. The identification of a silencer element whose ability to silence was dependent on the endogenous PcG, enabled Busturia and colleagues to investigate the requirement of this silencer element throughout development. In an elegant experiment, the FLP/FRT system was used to show that MCP725, the silencer element, is required for silencing throughout larval development, as removal led to derepression of the *lacZ* reporter gene. Silencing by the PcG therefore requires the continued presence of PREs indicating that the PcG must recognise and reassociate with this element after each round of DNA replication.

1-3.3 Mechanism of repression

Despite the enormous progress that has been made in the past few years on the molecular characterisation of the PcG members, relatively little is known about the mechanism through which the PcG heritably silences target loci. Several models have been proposed and are discussed below.

(i) Compaction model

The compaction model of silencing proposes that the PcG mediates the formation of a more compact chromatin, analogous to that of heterochromatin, which has traditionally been viewed as 'silent' chromatin. Heterochromatin is replicated late in S-phase of the cell cycle and remains condensed throughout cell division. Due to its presumed compaction no genes were thought to be present in heterochromatic regions and it was thought to consist mainly of repetitive DNA sequences. Euchromatin on the other hand replicates throughout S-phase, decondenses during interphase and contains genes which are actively transcribed. The compact state of heterochromatin compared with euchromatin led to the general belief that it consists of highly condensed DNA which is inaccessible to transcription factors and other components of the transcription machinery. Much of what is known about the formation of heterochromatin comes from the study of a phenomenon known as Position Effect

Variegation (PEV) in *Drosophila* and telomeric heterochromatin in *Saccharomyces cerevisiae*.

Position effect variegation

PEV describes the inactivation of genes normally found in euchromatin, after translocation to heterochromatic regions. The normally expressed genes become clonally repressed in some cells. The most intensively studied example of this phenomenon is that of an allele of the *white* gene w^{m4}. *white* is essential for the formation of adult eye pigment observed. The w^{m4} allele is an X-chromosome inversion which places the *white* gene adjacent to the heterochromatic centromere. Flies carrying this allele show a mosaic eye colour, with patches of red and white cells. The patches of red cells are indicative of expression of the *white* gene, whereas patches of white represent the cells in which the surrounding heterochromatin is presumed to have 'spread' into the juxtaposed *white* gene, silencing it. Components of heterochromatin have been identified using this w^{m4} allele in a genetic screen to identify modifiers termed *Enhancers of variegation (E(var))* or *Suppressors of variegation (Su(var))*. *Su(var)s* and *E(var)s* act to suppress or enhance the number of cells in which the *white* gene is repressed, leading to more cells (in the case of the *Su(var)s*) or less cells (in the case of the *E(var)s*) with red eye pigment. The cloning of some of these genes has identified proteins thought to be involved in a multimeric protein complex, like that of the PcG, which act on chromatin structure. Analysis of the role of these proteins in heterochromatin formation is yet to be performed. Additionally, no sequence analogous to that of a PRE has been identified that is able to mediate heterochromatin formation. As is the case with the PcG, not much is known about the mechanism of heterochromatin formation. It is, however, thought that the mechanism of PcG-mediated repression may be similar to that of heterochromatin formation. Several lines of evidence support this.

Firstly, some of the proteins involved in both PcG and PEV silencing share domains involved in chromatin-mediated modification of gene expression. The chromodomain is shared by both PC and HP1, a gene known to be localised to heterochromatin (Paro and Hogness, 1991). This observation led researchers to propose that PcG-mediated repression must be similar to heterochromatin formation. However, the chromodomain is also shared by proteins that are known to mediate transcriptional activation. For example the male-specific lethal (MSL)-3 protein, involved in hyperactivation of the X chromosome in male *Drosophila,* also contains a chromodomain (Koonin *et al.,* 1995). Perhaps stronger evidence for PcG-mediated repression being analogous to the of PEV is the observation that mutations in some members of the PcG have been shown to consistently modify PEV; *Asx, Pcl, Psc* and *E(Pc)* (Sinclair *et al.,* 1998; Stankunas *et al.,* 1998). Thirdly, transgenes containing PREs adjacent to a mini-*white* marker gene, used for selection of the plasmid

upon generation of transgenic lines, often show variegated eyes similar to that seen in the w^{m4} eyes (M. Fietz, unpublished observations).

Telomeric silencing

Telomeric silencing in yeast shows many similarities to that of PEV. The most striking of these is observed when an *ADE2* gene is placed near telomeric heterochromatin. Absence of the ADE2 protein causes the accumulation of a precursor that gives the yeast colonies a red appearance. In yeast where the *ADE2* gene has been inserted near the telomere, the colonies appear mottled, that is they have a red/white sectored appearance reminiscent of that seen in eyes of flies showing PEV (Gottschling *et al.*, 1990). Some of the proteins important in yeast telomere silencing have been well characterised. RAP1 encodes a DNA binding protein which recognises specific sequences present in telomeres and targets heterochromatin formation to these regions (Conrad *et al.*, 1990; Moretti *et al.*, 1994) When tethered to DNA via a GAL4 DNA binding domain RAP1 is able to silence adjacent genes at internal loci (Buck and Shore; 1995). The proteins, SIR2, SIR3 and SIR4, also play a role in telomeric silencing, in addition to the histone subunits, H3 and H4, whose involvement provides evidence for the role of nucleosomes in silencing (reviewed in Grunstein, 1997; Hecht *et al.*, 1995). In the current model of telomeric silencing, RAP1 recruits SIR3 and SIR4 to the telomere, SIR3/4 then polymerise, spreading into adjacent chromatin to generate a heterochromatic complex through their interaction with H3 and H4 (see Grunstein, 1997 for a review). This spreading of silencing is similar to the model proposed for heterochromatin-mediated silencing in *Drosophila*. Interestingly, mutations in *Drosophila* histone components suppress PEV (Moore *et al.*, 1979). No such effect has been observed on PcG-mediated silencing, suggesting that it is not going to be a straight forward similarity between the mechanisms involved.

Both PEV and telomeric silencing are thought to invoke packaging of the surrounding DNA into a higher order chromatin structure which in turn prevents access of proteins. This is inconsistent with the existence of at least 40 actively transcribed genes within the heterochromatic regions of *Drosophila* DNA (reviewed in Gatti and Pimpinelli, 1992). Compaction of chromatin would be predicted to reduce the access of restriction enzymes to silenced DNA. Two studies have been performed to investigate this and have generated conflicting results (Schloβherr *et al.*, 1994; Wallrath and Elgin, 1995). Schloβherr and colleagues investigated whether three different restriction endonucleases were able to access heterochromatic DNA. In this experiment they used the w^{m4} allele of the *white* gene. DNA was isolated from adult heads and subject to digestion with restriction endonucleases followed by ligation mediated-PCR to quantitate the accessibility of the *white* gene. Schloβherr *et al.* showed elegantly that the ability of a restriction enzyme to access the gene

was indistinguishable between E(var) and Su(var) lines, apparently ruling out the compaction model.

In a similar set of experiments, Wallrath and Elgin (1995) performed two experiments to test for accessibility of heterochromatic DNA. Firstly they generated a *P*-element transgene carrying the *hsp26* promoter and part of the *hsp26* ORF fused to a barley cDNA. This fusion was placed upstream of a *hsp70-white* reporter gene and transgenic flies were generated. The presence of the white gene enabled isolation of variegating lines, which were all found to be inserted in heterochromatic regions of the genome. If PEV is mediated by compaction of the chromatin, it would be expected that the *hsp26* promoter in the transgene would be less responsive to heat shock. To test this, fly lines exhibiting PEV of *hsp70-white* expression were heat shocked, along with lines which showed no variegation. A northern blot was then performed on isolated total RNA to determine the level of *hsp26* expression. In variegating lines, a significant reduction in the level of *hsp26* expression was observed. This reduction in *hsp26* expression was suppressed in the presence of a *Su(var)2-5* allele. This suggested a reduction in the accessibility to chromatin in regions undergoing heterochromatic silencing. Next they investigated the accessibility of a restriction endonuclease to transgene DNA. All variegating lines showed a reduction in the accessibility of a restriction endonuclease, providing evidence for a higher level of chromatin packaging in heterochromatic DNA, consistent with the compaction model.

Wallrath and Elgin expressed doubt about the results generated by Schloβherr and colleagues on the basis that the *white* gene is not expressed, and therefore not thought to be accessible in the majority of cells of the adult head, the tissue used in Schloβherr and colleagues access experiments. Wallrath and Elgin, on the other hand used the heat shock promoter which is potentially accessible in all tissues, making their results more significant. The compaction model could therefore account for PEV-mediated silencing. Interestingly yeast telomeric DNA has also been shown to be inaccessible to *E. coli* Dam DNA methylase, providing evidence that higher order chromatin structure mediates silencing at yeast telomeres (Gottschling, 1992). It remains to be shown whether a similar mechanism of silencing exists for both PEV and silencing at the yeast telomeres.

To determine whether PcG-mediated silencing was due to compaction of chromatin, similar to that observed in PEV and yeast telomere silencing, Schloβherr and colleagues (1994) performed restriction endonuclease access experiments. They chose to investigate access to the *Abd-B* locus in embryos. This locus is only expressed in parasegments 10-15 in the developing embryo and is therefore repressed by the PcG in the majority of cells. They compared wild-type embryos with *Pc* mutant embryos, where derepression of the *Abd-B* gene has occurred. If chromatin compaction is the method through which target loci are silenced by the PcG, then a mutation in *Pc* should render the DNA more accessible to

restriction endonucleases. They demonstrated that accessibility to the *Abd-B* gene locus by restriction endonucleases was uninhibited. The access of restriction endonucleases was similar in both wild-type and *Pc* mutant embryos. This experiment argues against the compaction model of PcG-mediated repression.

However, access to the DNA by restriction endonucleases does not necessarily imply that a locus is accessible to the transcription machinery. McCall and Bender (1995) examined the accessibility of the GAL4 transcriptional activator and T7 RNA polymerase to DNA known to be under PcG control *in vivo*. GAL4 is a yeast transcription factor, which had previously been shown to activate transcription in *Drosophila* from a promoter containing synthetic GAL4 binding sites (UAS sites) (Brand and Perrimon, 1993; Fischer *et al.,* 1988). McCall and Bender placed UAS sites upstream of a *lacZ* reporter gene and replaced an already existing *P*-element in the *bx* region of the *Ubx* gene with the *UASlacZ P*-element plasmid to generate $bx^{UASlacZ}$. They chose the *bx* region of the *U*bx gene as it had previously been characterised and shown to be under PcG control in the anterior half of the embryo (McCall *et al.,* 1994). They also generated a control line, whereby the *UASlacZ P*-element had been inserted into another region of the genome to generate *UASlacZ*. To determine if GAL4 could access the UAS binding sites in both the $bx^{UASlacZ}$ line and the *UASlacZ* line, GAL4 was introduced into these embryos and the expression pattern of *lacZ* was determined using *in situ* hybridisation. Expression of *lacZ* from the *UASlacZ P*-element indicated that GAL4 was able to access the UAS sites and direct expression of the *lacZ* gene. This is in contrast to the result seen with the $bx^{UASlacZ}$ which showed a segmentally restricted pattern of *lacZ* expression in the anterior half of the embryo. GAL4 was therefore excluded from DNA once it came under PcG-mediated repression. Further evidence for this exclusion of GAL4 comes from work performed by Zink and Paro (1995), who showed a mutually exclusive poytene binding pattern of GAL4 and PC, indicating that GAL4 was excluded from regions of DNA which were under PcG control.

A similar experiment was also performed using T7 RNA polymerase (McCall and Bender, 1995). T7 is a single subunit polymerase which binds to a 23 nucleotide sequence. The T7 binding site was inserted into the *bx* regulatory region (bx^{T7}) and T7 polymerase was introduced. *In situ* analyses revealed no difference in the accessibility of the T7 polymerase to sites found within the *bx* region (bx^{T7}) or other regions of the genome (*T7*). The PcG is therefore not preventing access of the T7 RNA polymerase to its binding site. This is in contrast to the results of the GAL4 access experiments where GAL4 was unable to access DNA and activate transcription. Perhaps GAL4 is able to access DNA but is unable to recruit the transcription machinery it requires to initiate transcription. T7 polymerase on the other hand does not require the presence of the endogenous transcription machinery. The results of both the restriction enzyme experiments and the T7 experiments argue against the compaction model as the favoured model for PcG-mediated repression. Clearly PcG-

mediated repression does not prevent the binding of all trans-acting factors to DNA under its control.

In direct contrast to the results of McCall and Bender (1995), Boivin and Dura (1998) demonstrated an *in vivo* difference in the accessibility of *E.coli. dam* DNA methyltransferase to chromatin under PcG-mediated repression. Little or no endogenous methylation is observed in *Drosophila* and previous work had demonstrated that expression of *E.coli. dam* DNA methyltransferase (Dam) in *Drosophila melanogaster* has no effect on viability. To assess methylase activity, the *Bcl*I enzyme, which is sensitive to the state of methylation, was used. Dam activity, and therefore accessibility of Dam to chromatin, was measured by probing the same sequences in different contexts, after digestion with *Bcl*I. Dam activity was significantly reduced at the *white* locus in variegating *P* element strains located in heterochromatic regions of the chromosome, when compared to euchromatically located *P* elements. Using the same assay, the effect of the PcG on access of the Dam methyltransferase was investigated. Accessibility studies were performed on two fly strains carrying *P* element insertions with a region of the *ph* regulatory region (*P[ph]*). This region of *ph* had previously been shown to induce variegation independent of the site of insertion. One strain contained a *P*[*ph*] inserted at 65F (T3), which is not a site of PcG binding and the other strain contained a *P*[*ph*] located at 86CD (T30), which is a site of PcG binding. Both constructs show variegation that was modified by mutations in both *ph* and *Psc*. The accessibility of both sites was measured in a wild type, *ph* and *Psc* background and the level of Dam activity measured. The level of Dam activity was significantly increased in both *ph* and *Psc* flies for both T3 and T30, indicating that the accessibility of Dam to sites under PcG control is inhibited. The similarity in the accessibility of Dam in PEV and PcG silenced strains is suggestive of a similar mechanism of repression. This is in direct contradiction to the results observed in the T7 polymerase access experiments (McCall and Bender, 1995). Unpublished observations cited in Boivin and Dura (1998) suggest that the ability of T7 polymerase to access DNA under PcG control is due to its small size, as a T7 polymerase-β-galactosidase fusion is unable to access DNA. It therefore seems plausible that the mechanism of action of the PcG group is analogous to that of the mechanism of heterochromatic silencing.

NOTE: Please read underneath page before proceeding to the page 17

mediated repression does not prevent the binding of all trans-acting factors to DNA under its control.

Spreading of heterochromatin

Both the yeast silencing and PEV model of silencing are thought to be mediated through the spreading of multimeric protein complexes into adjacent DNA (see Wakimoto, 1998; Lustig, 1998 for a review) Does PcG-mediated repression also use a 'spreading' mechanism? The insertion of a PRE into chromosomal DNA causes the formation of an additional PC binding site and also causes the silencing of neighbouring genes (Zink *et al.*, 1991; Chan *et al.*, 1994; Zink and Paro, 1995). This suggests that the PcG recognises the PRE sequence, silences it and then 'spreads' into adjacent genes. To determine whether this 'spreading' occurs, cross-linking studies were performed (Orlando and Paro, 1993; Strutt *et al.*, 1997). Early studies performed by Orlando and Paro (1993) demonstrated that PC was associated with the entire region of the BX-C in tissue culture cells. Strutt *et al.* (1997) repeated these cross-linking experiments using a more sensitive method of detection. They observed the binding of PC to discrete regions previously found to contain characterised PREs. They further showed that the association of PC to these regions spread only a few kilobases from the strongest binding point. The extent of this spreading is similar to that seen in telomeric silencing where the repressive SIR complex spreads only several kilobases from the site of nucleation (Lustig, 1998). This argues against long distance spreading as a mechanism of PcG-mediated silencing, however the authors could not rule out a low level of PC binding over the entire silenced region. It is not known how far PEV spreads from its site of nucleation as no sequences have been identified which initiate the formation of heterochromatin.

The evidence discussed above is inconclusive as to whether PcG-mediated repression acts to induce heterochromatin and the formation of compacted chromatin. An alternative mechanism of PcG-mediated repression is that of nuclear compartmentalisation, whereby certain loci are silenced by the PcG by being sequestered into a 'silencing' compartment devoid of transcription factors and other transcriptional machinery from accessing the DNA (Cockell and Gasser, 1999). Evidence exists for a similar mechanism in both PEV and telomeric silencing in yeast.

(ii) Nuclear compartmentalisation model

As mentioned previously, the SIR proteins (SIR3 and 4) and RAP1 are important in telomeric silencing and are thought to be important in both the initiation and maintenance of silencing (see Lustig, 1998 for a review). The localisation of RAP1, SIR3 and SIR4 in the nucleus was determined using immunofluorescence (Palladino *et al.*, 1993). It was shown

that all three proteins are associated with telomeres and are found in discrete foci at the periphery of the nucleus, consistent with a compartmentalisation model.

The nuclear compartmentalisation model predicts that the position of a gene in a chromosome could be important for its expression. For example, translocation of a euchromatic gene into a heterochromatic region would be expected to affect its expression by altering its localisation. The importance of chromosomal positioning has been demonstrated using a variegating allele of the *brown* gene, bw^D. *brown* is a gene required for eye pigmentation and is located on a distal region of chromosome 2, away from the silencing effects of heterochromatin (Lindsay and Zimm; 1992). The variegated phenotype of the bw^D allele is caused by an ~2 megabase insertion of heterochromatic DNA into the *brown* gene. The variegation phenotype is dominant over the wild type *brown* phenotype indicating that the allele must trans-inactivate the wild type gene. This dominant silencing has been shown to be dependent on pairing of the homologues (Dreesen *et al.*, 1991). Several alleles in which the bw^D allele is translocated to other regions of the genome were used to demonstrate that the position of the allele with respect to centromeric heterochromatin was important in the level of silencing. Translocating the bw^D locus closer to centromeric heterochromatin enhanced the silencing effect, suggesting an interaction between the heterochromatin insertion at the bw^D locus and centromeric heterochromatin (Henikoff *et al.*, 1995). Fluorescence *In Situ* Hybridisation (FISH) was performed to confirm the association between centromeric heterochromatin and the bw^D locus *in vivo* (Csink and Henikoff, 1996). Csink and Henikoff used two probes, one which labelled centromeric DNA and another which specifically labelled the *brown* gene. They performed double FISH labelling in two fly strains; a bw^+/bw^+ strain and a bw^+/bw^D strain and compared the localisation of the *bw* locus in each of these strains with the position of centromeric heterochromatin. The centromeric heterochromatin was consistently found to be positioned closer to the *brown* locus in the interphase nuclei of bw^+/bw^D strains than in the wild type strain. These data support the proposal that heterochromatin is compartmentalised in *Drosophila* nuclei.

Several lines of evidence support nuclear compartmentalisation as a model for PcG-mediated silencing. Pairing-sensitive silencing of the mini-*white* gene has been seen in constructs which contain PRE sequences, raising the possibility that these PREs are pairing and sequestering the surrounding region into a compartment which cannot be accessed by transcription factors (Fritsch *et al.*, 1999; M. Fietz, unpublished results). Secondly, localisation of PC in tissue culture cells shows it is present in 5-10 discrete foci, favouring a compartmentalisation model (Messmer *et al.*, 1992). The mammalian homologues of the PcG are also present in discrete foci in tissue culture cell lines (Alkema *et al.*, 1997; Saurin *et al.*, 1998).

Evidence contradicting the nuclear compartmentalisation model arises from another localisation experiment (Buchenau *et al.*, 1998). High-resolution confocal microscopy was used to investigate the nuclear distribution of three PcG proteins; PSC, PH and PC in whole mount embryos. Approximately 100 or more staining foci were observed to be distributed throughout most of the nucleus. This is in agreement with the number of binding sites observed for PcG members on polytene chromosomes and indicates that there is no compartmentalisation of silenced PcG loci in *Drosophila* embryonic nuclei. One explanation for the discrepancy in the results observed by Messmer *et al.* and Buchenau *et al.* is the use of tissue culture cells which are immortal and may overexpress some members of the PcG leading to clumping of PcG proteins into discrete foci, although no such overexpression has been reported. However the evidence argues against compartmentalisation as a model for silencing by the PcG, but does not rule out the association of adjacent PREs forming a stable mini-compartment.

The evidence presented suggests that neither model is capable of explaining the mechanism of PcG-mediated silencing. Silencing is likely to involve some modification of chromatin but the exact nature of this modification and its consequences remains unknown. The role of histone deacetylases in PcG-mediated repression has not been fully investigated, however no HDA activity has been detected in PcG complexes in mice (unpublished observations cited in van Lohuizen 1999 (review)). Clearly further investigation is needed to identify the mechanism through which the PcG is able to silence heritably. Isolation of non-PcG interactors of PcG members may shed light on both the mechanism of PcG silencing and the initiation of PcG-mediated silencing. The isolation of the Polycomb Repressive Complex 1 (PRC1) and investigation of the action of this complex on PREs, is sure to shed light on the mechanism of silencing and also on the competition that occurs between the PcG and the trxG. Furthermore, identifying all members of the PcG and characterising their intra-group interactions will shed light of the formation of the complex, information that is likely to be important in the recruitment of the PcG to their target loci and to the repression mechanism.

1-4 Polycomblike

Polycomblike (Pcl) was identified as a member of the PcG in a mutagenesis screen searching for enhancers of the homeotic phenotype of *Pc* mutants (Duncan, 1984). Since the mutant phenotype of *Pcl* alleles resembled gain of function homeotic mutations, it was predicted to play a role in repressing homeotic gene function (Duncan, 1984). Investigations into the expression pattern of *Ubx* and *Abd-B* in *Pcl* mutants showed derepression of both genes outside of their normal domains of expression, confirming the role of *Pcl* as a negative regulator of homeotic gene expression (Lonie *et al.*, 1994; Soto *et al.*, 1995). Alleles of *Pcl* have also been shown to interact strongly with other members of the PcG suggesting that it is a key member of the group (Jurgens, 1985; Campbell *et al.*, 1994).

Molecular characterisation of the *Pcl* locus has been performed (Lonie *et al.*, 1994). The gene codes for a 2.574kbp open reading frame and the nuclear protein product is present ubiquitously throughout embryogenesis, consistent with its role as a transcriptional regulator of homeotic gene expression. Subsequent characterisation of the PCL protein sequence revealed the presence of two PHD finger motifs (Lonie *et al.*, 1994; Robert 1997). The PHD fingers contain a Cys_4-His-Cys_3 motif thought to mediate interactions with DNA or other proteins. This motif has been identified in a number of proteins thought to be involved in modulating transcription (Aasland *et al.*, 1995). Two mammalian homologues of PCL have been identified; PHF1 (human) and MTF2 (mouse), and both show 34% sequence identity to PCL over a 258aa region spanning the two PHD fingers and a region C-terminal to the fingers, termed the conserved domain (cDOM).

The role of the PCL in the PcG is not well defined although genetic evidence suggests that it is an important member of the group. Immunolocalisation studies on polytene chromosomes demonstrated that the binding pattern of PCL is identical to that of PC, and therefore PH, and must therefore be partially overlapping with the other PcG members examined (Lonie *et al.*, 1994). However less is known about the molecular role of PCL in the group. In order to address this issue, both full length PCL and the cDOM of PCL were used as 'bait' in a yeast two hybrid matrix to identify any potential interactions between PCL and the products of cloned members of the Polycomb group (Robert, 1997). Full length PCL was unable to interact with any tested member of the PcG. The cDOM however, was shown to interact with E(Z), although this interaction was not confirmed *in vivo*. No function was attributed to the rest of the PCL protein, although these regions are likely to be important in the function of PCL.

1-5 Aims of this study

The broad aim of this thesis is to further investigate the role of PCL in the PcG complex. An interaction between the cDOM of PCL and E(Z) is likely to be important in the function of the PcG. Chapter 3 therefore describes experiments which were carried out to confirm the interaction between PCL and E(Z) *in vivo* and to determine the role of the PHD finger motifs in mediating this interaction. Experiments examining whether this interaction is conserved in the mammalian counterparts are also described.

To date the cDOM of PCL is the only region of the protein which has been assigned a function, that of interacting with E(Z). The *in vivo* significance of this interaction is yet to be demonstrated. Likewise a role for the amino and carboxy terminus of PCL is yet to be demonstrated. To address this issue, Chapter 4 describes the results of an *in vivo* tethering assay which was performed to identify functionally important regions in PCL. This assay demonstrated a PcG dependant silencing role for the amino terminus of PCL. The

experiments performed to identify interactions between the amino terminus and other PcG members are described.

Chapters 3 and 4 describe experiments performed to characterise the role of the cDOM and amino terminus of PCL. Chapter 5 describes yeast two hybrid experiments performed to identify interactors with the carboxy terminus of PCL.

The studies described in this thesis aimed to more fully characterise the functionally important regions of PCL and identify interactions between PCL and members of the PcG. The results of these studies will shed light on the role of PCL as a member of the PcG.

Chapter 2: Materials and Methods

2-1 Abbreviations

Abbreviations used are as described in "Instructions to authors", *Biochem. J* (1978) 169, 1-27. In addition;

aa	amino acids
APS	ammonium persulphate
bisacrylamide	N,N'-methylene-bisacrylamide
Blotto	5% skim milk powder in PBT
CIP	alkaline calf intestinal phosphatase
DAB	3, 3' Diaminobenzidine
EDTA	ethylenediaminetetraacetic acid
GAL	galactose
GLU	glucose
HRP	Horse radish peroxidase
IPTG	Isopropyl b-D-thiogalactopryanoside
kb	number of kilobase pairs
kDa	number of kilo daltons
LEU	leucine
MQ	MilliQ
NP-40	non-idet P-40
PAGE	polyacylamide gel electrophoresis
PBS	phosphate buffered saline
RO water	Reverse Osmosis water
r.p.m.	revolutions per minute
RT	room temperature
SDS	sodium dodecyl sulphate
TEMED	N,N,N',N-tetramethylenediamine

2-2 Materials

2-2.1 Enzymes

Enzymes were obtained from the following sources:

Restriction endonucleases	Pharmacia, New England Biolabs, Boehringer Mannheim.
T4 DNA ligase and CIP	Boehringer Mannheim
RNase A and Lysozyme	Sigma

2-2.2 Kits

Qiaex Gel purification kit :Qiagen.
Vectastain ABC kit: Vector labs Inc.
Enhanced Chemiluminescence Kit: Amersham
Bresaspin plasmid Mini Kit: Bresatec
TnT Coupled Reticulysate System:Promega

2-2.3 Antibodies

Primary antibodies

The antibodies used in this thesis were provided by the following people:
anti-E(z) (rabbit): Rick Jones (Southern Methodist University, USA).
anti-LacZ (rabbit) (protein-G purified): Julie Secombe (Saint Lab).
anti-LexA (rabbit): Roger Brent (Harvard, USA).
anti-Pcl (rabbit): Saint Lab.
anti-HA (rat): Boehringer Manneheim

Secondary antibodies

anti-rabbit HRP: Jackson Laboratories
anti-rat HRP: Jackson laboratories
streptavidin-HRP (Vectastain ABC kit)

2-2.4 Radiolabelled compounds

^{35}S-L-methionine was of *in vivo* grade and purchased from Amersham (Cat.No. SJ1015)

2-2.5 Antibiotics

Ampicillin: Sigma.

2-2.6 Molecular weight standards

(i) DNA

λDNA was digested with *Bst* EII and *Sal* I to produce fragments of (in kb) 14.14, 7.24, 4.82, 4.32, 3.68, 3.13, 2.74, 2.32, 1.93, 1.37, 1.26, 0.70, 0.45, 0.22 and 0.11

(ii) Protein

High molecular weight markers (GIBCO BRL) sizes (in kDa) 205, 116, 97, 66, 45 and 29

2-2.7 Bacterial strains

DH5α: F⁻,f80, *lacZ*ΔM15, *rec*A1, *end*A1, *gyr*A96, *thi*-1, *hsd*R17, (r_K-, m_K+), *sup*E44, *rel*A1, *deo*R, Δ(*lac*ZYA-*arg*F) U169 (Hanahan, 1983).

2-2.8 Yeast strains

The following strains were obtained from Roger Brent (Harvard, USA)

EGY48 *MATα, trp1, his3, ura3,* 6ops-*LEU2*.

EGY191 *MATα, trp1, his3, ura3,* 2ops-*LEU2*

2-2.9 *Drosophila* strains

P{ry⁺,BGUZ}/CyO;ry and *P{ry⁺,GBUZ}/CyO;ry* were obtained from Jurg Müller (MPI, Tuebingen, Germany).

PclE^{90}/CyO was obtained from Stanley Robert (Saint Lab).

The following *Drosophila* strains were obtained from Tory McGrath (Saint Lab); w; *P* {mini-*w⁺*,*hb-GAL-amino*}, *w*; *P* {mini-*w⁺*,*hb-GAL-amino+carboxy*},*w*, *P* {mini-*w⁺*, *hb-GAL-carboxy*}, and *w*, *P* {mini-*w⁺*, *hb-GAL-PHD*}.

E(z)^{61} e^{11}/TM3Sb was obtained from Rick Jones (Southern Methodist University, USA).

w^{1118} and standard "Balancer" stocks were obtained from the laboratory stocks and are as described in Lindsley and Zimm (1992).

2-2.10 Buffers and solutions

Agarose gel loading buffer:	50% (w/v) glycerol, 50 mM EDTA, 0.1% (w/v) bromophenol blue
PBS:	7.5mM Na_2HPO_4, 2.5 mM NaH_2PO_4, 145mM NaCl
PBT:	1 x PBS, 0.1% Tween 20 or Triton X-100
Protein gel transfer buffer:	48mM Tris-base, 39mM Glycine, 0.037% (w/v) SDS, 20% methanol
3X sample buffer:	10% glycerol, 2% SDS, 5% b-mercaptoethanol, 0.05% bromophenol blue12.5% 0.5M Tris-HCl pH6.8
Protein gel running buffer:	1.5% Tris-base, 7.2% Glycine, 0.5% SDS

STET:	50mM Tris-HCl pH8.0, 50mM EDTA, 8%w/v sucrose and 0.05% Triton X-100
TAE:	40mM Tris-acetate, 20mM sodium acetate, 1mM EDTA, pH 8.2
TBS:	20mM Tris-HCl pH 7.5, 150mM NaCl

2-2.11 Media

a) Bacterial media

All media were prepared with distilled and deionised water and sterilised by autoclaving, excepts heat labile reagents, which were filter sterilised. Antibiotics were added from sterile stock solutions after the media had been autoclaved.

L-Broth:	1% (w/v) amine A, 0.5% yeast extract, 1% NaCl, pH 7.0.
SOC:	2% bactotyptone, 0.5% yeast extract, 10mM NaCl, 2.5 mM KCl,10mM $MgCl_2$, 10 mM $MgSO_4$, 20mM glucose.
Plates:	L-Broth with 1.5% (w/v) bactoagar supplemented with ampicillin (50mg/ml) where appropriate.

b) Yeast media

All media were prepared with distilled and deionised water and sterilised by autoclaving, excepts heat labile reagents, which were filter sterilised.

Yeast minimal media:	0.17% yeast nitrogen base, 0.5% ammonium sulphate, 2%carbon source (amino acids were added as required).
YPD:	0.01% Yeast extract, 0.02% Bactopeptone, 0.0005% K_2HPO_4, 0.0005% KH_2PO_4, and 2% carbon source.
Plates:	Liquid broth with 2% bactoagar.

Amino Acids:

100X Tryptophan	0.5% (filter sterilised, wrapped in alfoil, and kept at 4°C)
100X Histidine	0.2%
100X Leucine	0.5%
100X Uracil	0.2%

c) Drosophila media

10% Treacle, 20% yeast, 1% agar, 10% polenta, 2.5% tegosept and 1.5% propionic acid.

2-2.12 Plasmids

a) Cloning and expression vectors

pBSKS+ (Statagene)

pEG202 (Gyuris *et al.*, 1993)

pGEX-3X and 1 (Smith and Johnson, 1988)

pJG4-5 (Gyuris *et al.*, 1993)

pGilda (Origene Technologies).

b) Constructs generated by others but used in this thesis.

The following constructs were generously provided by Michael Kyba and Hugh Brock (University of British Columbia, Canada); JG-AsxA, JG-Asx-Q, JG-AsxZn, JG-E(Pc)Ybox, pact-esc, JG-Pc, JG-Pcdelchromo, JG-Pcdel3', JG-ph, JG-phD, JG-phdelN, JG-phdelS, JG-Psc, JG-PscdelB, JG-Scm, JG-Su(z)2, and JG-Su(z)2delB.

The following constructs were provided by Stanley Robert; pBS-Pcl Sac/Not, pGEX3-Pcl, pGEX3-C430S (Saint Lab).

Michelle Coulson provided me with the pBS-PHF1 construct (Saint lab).

Judith Kassis provided me with pET-pho (FDA/NIH, USA.)

Rick Jones (Southern Methodist University, USA) provided me with JG-E(z), pGEX-EZH1, pBS-EZH2, and pBSKS-E(z)e32.

I.M.A.G.E. consortium provided me with 589332 (a cDNA containing the PHD fingers of hM96)(Lennon et al., 1996)

Tory McGrath provided me with pBS-Bam-Pcl-Bam (Saint lab).

c) The following constructs were generated by me in the following manner:

Pcl constructs

All constructs were generated by PCR using pBS-Pcl Sac/Not as a template unless otherwise stated.

i) Yeast constructs

pGilda-PCL-amino was generated using the 5'PCL *BamH*I primer and the 3' AminoΔPHD (*BamH*I) primer.

pGilda-PCL$_{132-423}$was generated using the 5' AminoCarboxy primer and the 3' AminoΔPHD (*BamH*I) primer

LexA-C430A-cDOM was generated using the 5'PHD finger *BamH*I primer and the 3' cDOM primer (*BamH*I) and pBS-C430A as a template.

LexA-C430A-PHDF1 was generated using the 5'PHD finger *BamH*I primer and the 3' PHDF1 (*BamH*I) primer and pBS-C430A as a template.

LexA-C430S-cDOM was generated using the 5'PHD finger *BamH*I primer and the 3' cDOM primer (*BamH*I) and pGEX-C430S as a template.

LexA-C430S-PHDF1 was generated using the 5'PHD finger *BamH*I primer and the 3' PHDF1 (*BamH*I) primer and pGEX3-C430S as a template.

LexA-C518A-cDOM was generated using the 5'PHD finger *BamH*I primer and the 3' cDOM primer (*BamH*I) and pBS-C518A as a template.

LexA-C518A-PHDF2 was generated using 5'PHDF2 *EcoR*I primer and the 3' PCL PHD primer (*BamH*I) and pBS-C518A as a template.

LexA-C518S-cDOM was generated using the 5'PHD finger *BamH*I primer and the 3' cDOM primer (*BamH*I) and pBS-C518S as a template.

LexA-C518S-PHDF2 was generated using 5'PHDF2 *EcoR*I primer and the 3' PCL PHD primer (*BamH*I) and pBS-C518S as a template.

LexA-PCL$_{567\text{-}700}$ was generated using the 5'Carboxy PCL *BamH*I primer and the 3' 2100 PCL *Nco*I primer.

LexA-PCL-Carboxy was generated using the 5'Carboxy PCL *BamH*I primer and the 3'PCL *BamH*I primer.

LexA-PCL$_{700\text{-}857}$ was generated using the 5'2100 PCL *BamH*I primer and the 3'PCL *BamH*I primer.

LexA-PCL$_{700\text{-}800}$ was generated using the 5'2100 PCL *BamH*I primer and the 3' 2400PCL *Nco*I primer.

LexA-PCL$_{800\text{-}857}$ was generated using the PCL 5' CCC primer (*BamH*I) and the 3'PCL *BamH*I primer.

LexA-cDOM was generated using the 5'PHD finger *BamH*I primer and the 3' cDOM primer (*BamH*I).

LexA-PCL$_{567\text{-}700}$ was generated using the 5' 1900 *BamH*I PCL primer and the 3' 2400PCL *Nco*I primer.

LexA-ΔEH was generated using the 5'PHD finger *BamH*I primer and the 3' PCL PHD primer (*BamH*I).

LexA-PCL$_{1\text{-}200}$ was generated using the 5'PCL *BamH*I primer and the 3' *EcoR*I Δ3PCL primer.

LexA-PHD+C was generated using the 5'PHD finger *BamH*I primer and the 3'PCL *BamH*I primer.

LexA-PHDF1 was generated using the 5'PHD finger *BamH*I primer and the 3' PHDF1 (*BamH*I) primer.

LexA-PHDF1-QQQA-cDOM was generated using the 5'PHD finger *BamH*I primer and the 3' cDOM primer (*BamH*I) and pBS-*PHDF1-QQQA* as a template.

LexA-PHDF1-QQQA-PHDF1 was generated using the 5'PHD finger *BamH*I primer and the 3' PHDF1 (*BamH*I) primer and pBS-*PHDF1-QQQA* as a template.

LexA-PHDF2 was generated using the 5'PHDF2 *EcoR*I primer and the 3' PCL PHD primer (*Xho*I).

LexA-PHDF2-QQQA-cDOM was generated using the 5'PHD finger *BamH*I primer and the 3' cDOM primer (*BamH*I) and pBS-*PHDF2-QQQA* as a template.
LexA-PHDF2-QQQA-PHDF2 was generated using the 5'PHDF2 *EcoR*I primer and the 3' PCL PHD primer (*BamH*I) and pBS-*PHDF2-QQQA* as a template.
pAD-PHDF2 was generated using 5'PHDF2 *EcoR*I primer and the 3' PCL PHD primer (*BamH*I).

All these constructs were sequenced to ensure that no PCR induced errors occured and tested for the induction of the correct size protein using Western blotting with the anti-LexA antibody (1/500) (when cloned into pGilda or pEG202) or an anti-HA antibody (1/500) (when cloned into pJG4-5).

ii) Bacterial expression constructs
pGEX-amino was generated using the was generated using the 5'PCL *BamH*I primer and the 3' AminoΔPHD (*BamH*I) primer.
pGEX-PCL$_{1\text{-}200}$ was generated using the 5'PCL *Bam*HI primer and the 3'*EcoR*I Δ3Pcl primer.

Su(z)2 constructs
AD-HR was generated by PCR using 5'HR Su(z)2*EcoR*I primer and 3'HR Su(z)2 *Xho*I primer. The template was JG-Su(z).
AD-SAACR was generated by PCR using 5'SAACR Su(z)2 *EcoR*I primer and 3'SAACR Su(z)2 *Xho*I primer. The template was JG-Su(z)2.

Pleiohomeotic constructs
pBS-PHO was generated by PCR using 5'PHO *EcoR*I primer and 3' PHO *EcoR*I primer. pET-PHO was used as a template.
AD-PHO was generated by PCR using 5'PHO *EcoR*I primer and 3' PHO *EcoR*I primer. pET-PHO was used as a template.
AD-PHO-Amino was generated generated by digesting the *JG-PHO* construct with *Xho*I which dropped out the carboxy terminus of PHO.
AD-PHO-Carboxy was generated by PCR using the 5' PHO Carboxy *EcoR*I primer and the 3' PHO *EcoR*I primer. The template was pET-PHO.

Miscellaneous constucts.
AD-esc was generated by PCR using the 5' ESC *Xho*I primer and 3' ESC *Xho*I primer. pact-esc was used as the template.

AD-EZH1 was generated by dropping out the EZH1 fragment (*Xho*I) from pBS-EZH1 and cloning in frame into pJG4-5.

AD-EZH2 was generated by PCR using the primers 5' EZH2 *Xho*I primer and 3'EZH2 *Xho*I primer and pBS-EZH2 as a template.

AD-hM96 was generated by dropping out an *EcoRI* fragment from LexA-hM96 and cloning into the *EcoR*I site of pJG4-5.

pBS-C430S was generated by dropping out a *BamH*I fragment from pGex-C430S and cloning into the *BamH*I site of pBSKS+.

pBS-PC was generated using the 5' Pc *EcoR*I primer and the 3' *Xho*I primer and JG-PC as the template.

LexA-EZH1 was generated by digesting pBS-EZH1 with *Xho*I to isolate the insert and then cloning into the *Xho*I site of pJG4-5.

LexA-EZH1 was generated by diegesting AD-EZH2 with *Xho*I to drop out the insert and then cloned into the *Xho*I site of pEG202.

LexA-hM96 was generated using the 5'hM96 *EcoR*I primer and the 3' hM96 *EcoR*I primer and 589332 (cDNA containing hM96) as a template.

LexA-PHF1 was generated using5' PHF1 primer (*EcoR*I) and 3' PHF1 primer (*BamH*I) and pBS-PHF1 as a template.

LexA-TRX-PHDFs was generated using the 5' TRX-PHDF primer (*EcoR*I) and the 3' TRX-PHDF primer (*EcoR*I).

2-2.13 Oligonucleotides

Position of primer in the ORF is indicated in brackets.

esc oligonucleotides (used for cloning into pJG4-5)

5' ESC *Xho*I primer (1-18)	5'-CCGCTCGAGATGAGCAGTGATAAAGTG
3' ESC *Xho*I primer (1260-1278)	5'-CCGCTCGAGTCAGATGGAAGTTGTTTG

EZH2 oligonucleotides (used for cloning into pJG4-5)

5' EZH2 *Xho*I primer (1-18)	5'-CCGCTCGAGATGGGCCAGACTGGGAAG
3' EZH2 *Xho*I primer (2224-2241)	5'-CCGCTCGAGTCAAGGGATTTCCATTTC

hM96 oligonucleotides (used for cloning into pEG202)

5' hM96 *EcoR*I	5'-GGAATTCATGGTCTGTACAATATGT
3' hM96 *EcoR*I	5'-GGAATTCTCCAGAACTGCAGACAGC

(unable to give primer locations because the sequence of the primers were based on a potentially incomplete cDNA clone from I.M.A.G.E.)

Pc oligonucleotides (used for cloning into pBS)

5' Pc *EcoR*I (1-18)	5'-GGAATTCATGACTGGTCGAGGCAAG
3' Pc *Xho*I (1156-1174)	5'-GGGCTCGAGTCAAGCTACTGGCGACGA

Pcl oligonucleotides

i) Oligonucleotides used for cloning into pEG202

*5' PCL *BamH*I primer (1-18)	5'-GGAGGATCCTGATGAACAACCATT
*5' PHD finger *BamH*I primer (1269-1287)	5'-CGGGATCCGCGGACCCATGTGCGTGGCC
5' PHDF2 *EcoR*I primer (1534-1552)	5'-GGAATTCCAGATCTACTGCTACTGC
*5' Carboxy PCL *BamH*I primer (1700-1718)	5'-GCGGATCCGGATAGAATTCGTTCGTCGT
5' 1900 *BamH*I PCL primer (1900-1918)	5'-GCGGATCCAGCAAACCCTTAAGGATTAC
5' 2100 PCL *BamH*I primer (2100-2118)	5'-GCGGATCCACGAGCTTCCGAAAAGAGTT
PCL 5' CCC primer (*BamH*I) (2400-2418)	5'-GCGGATCCGTCACTGTGATCTATCATCC
*3'*EcoR*I Δ3Pcl primer (588-606)	5'-GGAATTCATAGTGGTTATTGATGCA
*3' *EcoR*I Δ2 PCL primer (1182-1200)	5'-GGATTCACCCAGCTTGCGCAGCTT
3' PHDF1 (*BamH*I) primer (1398-1416)	5'-GGGGATCCGGGTTTAGCGCAGCGTTT
*3' PCL PHD primer (*BamH*I) (1681-1699)	5'-GCGGATCCTTAGCCATTGTTGCAGACCGT
3' PCL PHDF2 primer (*Xho*I) (1681-1699)	5'GGGGAGCTCTTAGCCATTGTTGCAGACCGT
3' cDOM primer (*BamH*I) (1797-1815)	5'-AAGGATCCGGGCCAAATGTCATTTAGCAG
3' 2100 PCL *Nco*I primer (2082-2100)	5'-CATGCCATGGGTTGGCACCTTCATCAGG
3' 2400 PCL *Nco*I primer (2382-2400)	5'-CATGCCATGGATTGCGGCTGTTGTCATA
*3' PCL *BamH*I primer (2556-2574)	5'-GCGGATCCTTACTCCGACTCCAGTTC

*indicates that these primers were also used in the generation of pGex3-PCL constructs.

ii) Oligonucleotides used for cloning into pJG4-5

5' *Sal*I PCL primer (1-18) 5'-GCGTCGACATGATGAACAACCATTTT

3' *Sal*I PCL primer (2556-2574) 5'-GCGTCGACTTACTCCGACTCCAGTTC

iii)Oligonucleotides used in the mutagenesis of the PHD fingers

C518S forw (1531-1570) 5'-GAGCAGATCTACTGCTACAGTGGCAAACCGGGAAAA TTCG

C518S back (1531-1570) 5'-CGAATTTTCCCGGTTTGCCACTGTAGCAGTAGATCTG CTC

C430A forw (1273-1299) 5'-CCCATGTGCGTGGCCGCCAAGCGATCGGATATCG

C430A back (1273-1299) 5'-CGATATCCGATCGCTTGGCGGCCACGCACATGGG

C518A forw (1531-1570) 5'-GAGCAGATCTACTGCTACGCCGGCAAACCGGGAAAA TTCG

C518A back (1531-1570) 5'-CGAATTTTCCCGGTTTGCCGGCGTAGCAGTAGATCT GCTC

Forw*PHDF1-QQQA* (1288-1351) 5'-TGCAAGCGATCGGATATCGAAGACGTGCAGCAGC AGGTGAGCGCTGC GGACGTGGCTATCATCGT

Back*PHDF1-QQQA* (1288-1351) 5'-ACGAGATAGCCACGTCCGCAGCGCTCAGCCTGC TGCTGCACGTCTTCGATATCCGATCGCTTGCA

Forw*PHDF2-QQQA* (1552-1611) 5'-GGCAAACCGGGAAAATTCGATCACAATCAGCAG CAGGCATGCAAATGTCGGAACTGGTTCC

Back*PHDF2-QQQA* (1552-1611) 5'-GGAACCAGTTCCGACATTTGCATGCCTGCTGCTGATTG TGATCGAATTTTCCCGGTTTGCC

see Robert (1997) for a description of the generation of the C430S mutation

<u>PHF1 oligonucleotides</u> (for cloning into pEG202)

5' PHF1 primer (*EcoR*I) (259-276) 5'-GGAATTCGAACTCCTCTGTTGTGTC

3' PHF1 primer (*BamH*I) (703-720) 5'-GGGGATCGCCCCCGCGCACCACACA

<u>pho oligonucleotides</u> (for cloning into pJG4-5)

*5' pho *EcoR*I primer (1-18) 5'-GGAATTCATGGCATACGAACGTTTT

*3' pho *EcoR*I primer (1547-1563) 5'-GGAATTCTCAGTCTGCATATACCAC

5' pho carboxy primer (*EcoR*I) (832-849)	5'GCGAATTCGAAACTGCCGCAATGACA

*indicates that these primers were also used for cloning into pEG202 and pBSKS+

Sequencing oligonucleotides

5' GAL4 seq. primer (373-390)	5'-CATAGAATAAGTGCGACA
5' pEG202 seq. primer (unknown)	5'-CGTCAGCAGAGCTTCACCATT
5' pGEX seq. primer (869-891)	5'-GGGCTGGCAAGCCACGTTTGGTG
5' pJG4-5 seq. primer (271-288)	5'-CTGAGTGGAGATGCCTCC

Su(z)2 oligonucleotides (for cloning into pJG4-5)

5'HR Su(z)2*EcoR*I primer (1-18)	5'-GGAATTCATGCATCTGCAAAACACG
3'HR Su(z)2 *Xho*I primer (691-708)	5'-GGGCTCGAGTTGCTCGTTGTCGTAAAG
5'SAACR Su(z)2 *EcoR*I primer (499-516)	5'-GGAATTCGCCGGATTACCTGTGGAG
3'SAACR Su(z)2 *Xho*I primer (1417-1434)	5'-GGGCTCGAGGGATCCATTGGTCGAACT

2-3 Methods

Most of the common methods used in the generation of data for this thesis can be found in Ausubel *et al.* (1995).

2-3.1 Generation and transformation of recombinant plasmids

When generating recombinant plasmids, both the vector (which had been dephosphorylated) and the insert were gel purified using Qiagen's Qiaquick spin colums, prior to the ligation being set up using T4 DNA ligase. The ligations were incubated overnight at 18°C, phenol/chloroform extracted and ethanol precipitated (using glycogen), resuspended in 10µl of sterile MQ water and half was transformed into electrocompetent DH5α.

2-3.2 Transformation procedure for recombinant plasmids into E.coli

A 500ml flask of L-broth was innoculated with 5ml of an overnight culture of *E. coli* DH5α cells and grown to an OD_{A600} of 0.5-0.6. The culture was then chilled in an ice slurry for 15 to 30 minutes and the cells harvested by centrifugation at 4000g for 15 minutes. The cells were then resuspended in 500ml of ice-cold MQH_2O, pelleted at 5000g, resuspended in 250ml of ice-cold MQH_2O, pelleted at 4000g, resuspended in 10ml of ice-cold 10% glycerol, repelleted at 3000g and finally resuspended in 1ml of ice-cold 10% glycerol. The

competent cells were then snap frozen in liquid nitrogen and stored as 45µl aliquots at -80°C. For transformation, cells were thawed at RT, added to a portion of ligation reaction mixture and incubated on ice for at least 30 seconds. Cells were then transferred to an ice-cold 2mm electroporation cuvette and electroporated in a Bio-Rad Gene Pulser at 2500V. The cuvette was immediately washed out with 1ml of SOC, and the suspension incubated at 37°C for 35 minutes. Cells were then pelleted by centrifugation in a bench centrifuge for 8 seconds at 15000g. 800µl of the supernatant was removed, and the cells gently resuspended in the remaining SOC. The cell suspension was plated onto L-agar plates supplemented with 50µg/ml ampicillin and incubated at 37°C overnight.

2-3.3 PCR amplification of DNA

Statagenes *Pfu* polymerase was used in all of the reactions according to the manufacturers instructions. The annealing temperatures ranged from 50°C to 60°C depending on the primers used.

2-3.4 Isolation of plasmid DNA

a) Small scale preparation- rapid boiling preps

A 2ml culture supplemented with the appropriate antibiotics, was incubated overnight at 37°C, with shaking. Cells were harvested by centrifugation at 15000g in a microcentifuge for 15 seconds. The bacterial pellet was then resuspended in 200µl of STET, followed by addition of 10µl of 10mg/ml lysozyme. The suspension was heated at 100°C for 45 seconds and centrifuged at 15000g for 15 minutes. The pellet was removed with a toothpick. Plasmid DNA was then precipitated from the supernatant with 240µl of isopropanol, followed by washing in 70% ethanol, dried and resuspended in 15µl H_2O.

b) Large scale preparation

A single colony was used to innoculate 50ml of L-broth supplemented with the appropriate antibiotics, which was then incubated overnight at 37°C with shaking. Cells were harvested by centrifugation at 5000g for 5 minutes and the bacterial pellet resuspended in 4ml of P1 (50mM Tris-HCl pH7.5, 10mM EDTA pH 8.0 and10mg/ml of RNase A, when RNase treatment was required). Cells were lysed by the addition of 4ml of freshly prepared 0.2M NaOH/1%SDS and incubation at RT for 5 minutes after gentle mixing. Bacterial debris was precipitated by addition of 4ml of ice-cold P3 (3M KAc pH5.5) and incubation on ice for 10 minutes followed by centrifugation at 15000g for 15 minutes. The supernatant was transferrred into a fresh tube and spun again at 15000 for 15 minutes. The mixture was extracted with an equal volume of phenol/chloroform and the aqueous phase separated by centrifugation at 8000g for 2 minutes. DNA was precipitated from the supernatant by addition of 1/10th volume of 3M NaAc pH 4.6 and 2.5 volumes of 95% ethanol, and centrifugation at 15000g for 30 minutes. The DNA pellet was washed in 70% ethanol, dried and resuspended in 200µl H_2O.

c) preparation of "clean" DNA

When ultraclean DNA was required (for use in *in vitro* transcriptions and translations) a CsCl preparation was performed.

2-3.5 Automoated sequencing

The DNA to be sequenced was prepared using the Bresaspin Plasmid Minikit according to the manufacturers protocol. DNA was sequenced using the ABI Prism™ Dye Teminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer), essentially as descibed in the manufacturer's protocol with the modication of using half the descibed amount of reaction mix. Reactions were cycled through 25 cycles 96°C for 10 seconds, 50°C for 5 seconds and 60°C for 4 minutes with a temperature ramp setting of 2 in a Corbett Research capillary thermal cycler. Extension products were purified by ethanol precipitation with 2µL of 3M NaOAc pH 4.6 and 50 µL 95% ethanol and chilling at -20°C for 20 minutes. The precipitate was pelleted by centrifugation at 15000g for 20 minutes at 4°C, then washed with 250µL of 70% ethanol and dried in a 65°C heating block . The sequencing reactions were then processed at the Department of Molecular Pathology, IMVS, Adelaide.

2-3.6 In vitro mutagenesis

The design of the primers used in the site directed mutagenesis followed the instructions provided in Statagene's QuikChange Site-Directed Mutagenesis kit. The reaction was carried out in a 20µl volume. To each reaction the following was added

2µl of 10x reaction buffer
50ng of dsDNA template (pBS-Pcl Sac/Not)
125ng of the forward primer
125ng of the backward primer
1µl of 10mM dNTP mix (2.5mM each dNTP)
Sterile ddwater to a final volume of 19µl
1µl of *Pfu* Turbo DNA polymerase (2.5Uµl) was then added

The reaction was then cycled according to the manufacturers instructions in a capillary PCR machine. The reaction was then cooled to <37°C, transferred from the capillary tip to an eppendorf and 1µl of *Dpn*I (10U/µl) was added to the reaction and mixed and allowed to incubate at 37°C for 1 hour to digest the parental (nonmutated) dsDNA. The reaction was then phenol/$CHCl_3$ extracted, and ethanol precipitated and resuspended in 10µl of water and 2µl was transformed into DH5α using electroporation. Restriction analysis was used to identify mutated clones where possible but all mutations were confirmed by sequence analysis before use in subsequent experiments.

2-3.7 Expression of bacterial fusion proteins

Clones in the pGEX plasmid were transformed into bacterial strains DH5α. A single colony was transferred into a flask containing LB-broth and 100μg/mL ampicillin and grown overnight at 37°C. A 1:10 dilution of this culture was made into a flask containing LB-broth and 100μg/mL of ampicillin and incubated at 37°C until the O.D.A600 reached 0.6-0.8. IPTG was added to a final concentration of 1mM. The culture was then incubated at 37°C for a further 3-6 hrs to allow the accumulation of expressed protein. The bacterial cells were then pelleted at 4800rpm and the medium discarded. The pellet was then resuspended in 1/20th the original culture volume in 3X sample buffer and stored at -20°C.

2-3.8 Protein gel electrophoresis

The Bio-rad Mini-Protean II gel electrophoresis system was used protein gel electrophoresis. The gels were run at 180-200V until the bromophenol blue in the sample buffer had reached the bottom of the gel.

2-3.9 Western blotting

Western blotting of proteins onto nitrocellulose membrane was performed as described in Harlow and Lane (1988). Nitrocellulose blots were washed thoroughly with PBT and then blocked for 1 hour in 5% Blotto. Primary and secondary antibody incubations were carried out overnight at 4°C and for 45 minutes at room temperature respectively, with the appropriate dilutions of antibody in blocking solution. The secondary antibodies were always horseradish peroxidase conjugated (Jackson) and therefore detection was either by the Enhanced Chemiluminescence (Amersham) or by colorimetric detection using nickel enhanced DAB staining (Harlow and Lane, 1988)

2-3.10 Radiolabelling proteins

^{35}S-labelled proteins were prepared using the TnT coupled *in vitro* transcription/translation kit (Promega) according to the manufacturers instructions. If the TnT reaction was to be used for a far western blot, the protein was seperated from the unincorporated label by passing the protein through a Sepharose G-25 spin column (Pharmacia) equilibrated in TEN (10mM Tris-HCl pH 7.5, 0.1mMEDTA, 150mM NaCl). The column was constructed by placing a small amount of sterile glass wool to the bottom of a 1ml syringe and approximately 1ml of beads were added, this was then spun at 2,000rpm for 5 minutes to remove of any excess TEN. The column was topped up with more G-25 until a column volume of 500-600μl was achieved. The TnT reaction was then loaded onto the column and spun at 2,000rpm for 5 minutes, and then rinsed through with 2X50μl of TEN. This entire reaction was then used in the far western blot. If the protein needed storing a final concentration of 15% glycerol was added and the mix was stored at -20°C (short term only).

2-3.11 Far-western protocol

A western blot of the desired proteins were prepared as described in 2-3.9. The filter was washed thoroughly in TBST (with shaking at RT) (50mM Tris-HCl pH 7.5, 150mMNaCl, 0.05% Tween-20) and then treated with 20ml of cold HBB (20mM HEPES-KOH pH 7.6, 1mM KCl, 5mM $MgCl_2$, 1mM$ZnSO_4$, 10mM β-mercaptoethanol) containing 6M guanidine-HCl for 1 hour at 4°C with shaking. Renaturation of the proteins occured by adding 20ml of HBB every 30 minutes for 2 hours to the already existing solution with shaking at 4°C. The solution was then washed extensively in HBB alone, with shaking (for at least an hour, changing washes several times), followed by blocking in 5% Blotto (in HBB) for no longer than 1hr at 4°C. A single *in vitro* TnT reaction (in approximately 3mls of 1%Blotto in HBB) was then added to the blot in a plastic bag that had been tripled sealed to prevent leaking. This was allowed to hybridise overnight at 4°C with nutation. The membrane was then washed at room temperature 3 times for 10 minutes each in PBS+0.2% Triton-X-100, dried completely and then exposed to a phosphoimager plate overnight .

2-3.12 Yeast Transformation protocol

A 5ml culture was set up in the appropriate medium (either YPD or dropout media) and incubated overnight at 30°C with shaking. The next morning another 5ml of media was innoculated to an OD_{A600} of 0.3 and allowed to grow for 3 hours if in YPD and for 5 hours if in dropout media at 30°C with shaking. The yeast were then pelleted for 5 minutes at 2,800rpm in a benchtop centrifuge and the pellet was resuspended in 1ml of 0.9MLiOAc/TE and transfered to an eppendorf tube. The yeast were then spun at 8000rpm for 30 seconds and the supernantant was removed and the yeast were resuspended in 100µl of 0.9MLiOAc/TE for every 0.2 OD units/5ml. For each transformation 8µl of DNA from a rapid boiling prep that had been resuspended in 20µl, was added to an eppendorf along with 12µl of the yeast/LiOAC/TE mix and 45µl of sterile 50% PEG 3350. The tubes were then placed in a 30°C shaking incubator for 1 hour and heat shocked for 5 minutes in a water bath at 42°C. The yeast were then plated out onto selective media and placed at 30°C for 3-4 days to allow growth of transformants.

2-3.13 Yeast protein extraction protocol

A 50 ml culture of yeast was set up in the appropriate selective media and allowed to grow overnight at 30°C with shaking to an OD600 of approximately 0.8. The yeast were then pelleted in a sterile falcon tube at 2,800 for 5 minutes and the supernatant was removed. The yeast were then resuspended in between 200-400µl of ice cold lysis RIPA buffer (50mM Tris-HCl pH 7.4, 150mM NaCl, 5mM EDTA, 0.1% NP-40 and 5mM PMSF) depending on the amount of yeast. The yeast were then transferred to a cold screw cap eppendorf and half the volume of 425-625µm glass beads were added. The eppendorfs were then placed into a

"bead beater" at 4°C (Dalton Lab, Dept. of Biochemistry, University of Adelaide, Australia) and beat 3X 20sec, with a 20 sec break in between. The tubes were then spun at 14K for 15 minutes at 4°C. The supernantant was then transferred to a clean cold eppendorf and stored at -80°C for long term storage or -20°C for short term storage.

2-3.14 Yeast interaction screening procedure

In order to determine if two proteins were interacting using the yeast two hybrid assay, 3 colonies, containing the plasmids encoding the proteins of interest, were picked and first streaked onto GLU plates, followed by GAL plates, then GAL/LEU and finally GLU/LEU. This was done three times. The plates were then allowed to grow for 4 days at 30°C. Colonies of approximately 1-2mm indicated a strong interaction.

2-3.15 Co-immunoprecipitations in vivo

The beads

In order to prepare the protein A CL-4B beads for coupling to the antibody of interest, an appropriate amount of beads were swelled in sterile MQ-water and several washes were performed over a couple of hours. The beads were then washed 3X in 25mM HEPES pH 7.6/150mM NaCl. The appropriate amount of antibody was then added to 20µl beads in screw capped eppendorfs and placed at 4°C overnight with nutation.

The amount of antibody used in the co-IPs performed in this thesis is as follows:

E(z) 20µl of affinity purified anti E(z) antisera generated in rabbit and generously provided by Richard Jones, Southern Methodist University, Texas, USA.

PCL 30µl of affinity purified anti-PCL antisera generated in rabbit.

The embryos

A 14 hour collect of w^{1118} embryos (at 25°C) were aged for 6 hours at 25°C and collected into a "basket" and dechorionated in 50% bleach for 3 minutes. The bleach was constantly poured through onto the embryos to enable proper dechorionation. The embryos were then washed thoroughly in RO water. The embryos were then transferred into an eppendorf tube using a paint brush and rinsed in PBS+0.1%Tween20. The excess liquid was removed and the embryos were snap frozen in liquid Nitrogen and stored at -80°C until needed. On the day that the embryos were going to be used, the embryos were thawed on ice and an equal volume of HoB buffer (25mM HEPES pH 7.4, 150mM NaCl, 5mM NaF, 1mM EDTA, 1mM DTT, 0.1mM Na_3VO_4, 0.1%Triton-X-100 and 1 Complete Mini-EDTA free Protease Inhibitor cocktail tablet (from Boehringer Mannheim). The embryos were then homogenised in the 4°C room using a plastic homogeniser that fitted into the bottom of the eppendorf tube and spun at 14K for 15minutes. The supernatanat was then removed and into a fresh cold tube and spun again at 14K for 15minutes.

The co-IP

400µl of embryonic extract was added to the antibody/protein A bead complex and incubated overnight at 4°C with nutation. The next day the tubes were spun for 10 secs in a centrifuge at 4°C to pellet the bead/antibody/protein complex and the supernatant was carefully removed. Washes were then performed twice in 250µl of HoB buffer by flicking the tube to resuspend the bead/antibody/protein complex and spinning for 10 secs in between washes. Two more washes were then performed by adding 250µl of 25mM HEPES pH7.4/1mMDTT. The second of these washes was allowed to incubate with nutation for 15 minutes at 4°C. 30µl of 3X sample buffer was then added to the beads and boiled at 100°C for two minutes followed by a 15 sec vortex to ensure that the antibody and co-ip'd proteins had dissociated from the beads. The tubes were then spun for 5 minutes to pellet the beads and half of the supernatant was loaded onto duplicate protein gels and western transfers were performed.

2-3.16 Collection and fixation of *Drosophila* embryos

Embryos were collected on grape juice agar plates (25% grape juice,3% sucrose, 3% tegosept and 3% agar) smeared with yeast. They were then harvested and washed thoroughly in a basket using 0.7%NaCl, 0.15%Triton X-100. The sieve was then transferred into a container with 50% commercially available bleach (2% sodium hypochlorite) for 3 minutes to de-chorionate the embryos. The embryos were once again washed in the basket thoroughly using 0.7%NaCl, 0.15%Triton X-100. They were then transferred to a glass scintillation vial containing a two-phase mix of 4 mL of 4% formaldehyde in PBS (made fresh by boiling paraformaldehyde in PBS) and 4 mL of heptane. The vial was then shaken on an orbiting platform such that the interface between the liquid phases was disrupted and the embryos were bathing in an emulsion, for 20 minutes to 'fix' the embryos. The bottom phase (aqueous) was drawn off and replaced with 4 mL of methanol and the vial was shaken vigorously for 30 seconds to de-vitellinise the embryos. De-vitellinised embryos sink from the interface and were collected from the bottom phase (methanol). Embryos were rinsed thoroughly in methanol at which point they were either processed for whole mount immuno-staining or storage at -20°C in methanol.

2-3.17 Whole mount immuno-staining of *Drosophila* embryos

The methanol was removed from embryos in a microfuge tube and replaced with 50% Methanol/PBT. Several rinses were then done using PBT followed by a single wash for 30 minutes. The embryos were then 'blocked' in 1 ml of PBT containing 5% Blotto for at least 1 hour. The blocking solution was removed and primary antibody diluted in fresh blocking solution was added (usually 200µL). The embryos were routinely incubated with gentle nutation at 4°C overnight. The next day, the antibody solution was removed and the embryos were washed extensively in PBT (several changes of buffer over a 2 hour time

period). The embryos were then incubated with secondary antibody diluted in fresh blocking solution for at least 2 hours at room temperature with gentle nutation. Following a period of washing as for the primary antibody, the embryos were then incubated with a tertiary complex (streptavidin-HRP)(Vectastain ABC kit) for two hours and then washed extensively for two hours in PBT. To detect the antibody the embryos were incubated in a solution of 0.5 mg/mL DAB, 0.045% H_2O_2, 0.064% $NiCl_2$ until the staining had developed (as assayed on a dissecting microscope), and then rinsed thoroughly using PBT prior to mounting in PBS/80% Glycerol. The embryos were then placed at 4°C overnight and then mounted onto a slide under a coverslip supported by two pieces of double sided tape and coverslips were sealed to the glass using comercially available clear nail varnish. The embryos were then viewed using a Ziess Axiophot microscope with Normarski optics. Embryos were photographed using Kodak Ektachrome 160T colour slide film. Slides were digitised using a Kodak RFS 2035 Film Scanner and image manipulations were performed using Adobe Photoshop 4.0 or 5.0.

2-3.18 Fly maintenace

All flies were maintained in either 18°C or 25°C constant temperature rooms, which were humidified.

2-3.19 P-element mediated transformation of *Drosophila*

a) micro-injection of embryos

High purity DNA for injection was prepared using the Qiagen DNA preparation kit described above. The construct DNA at 700ng/µl and the transposase activity plasmid, pπ25.7*wc*(Δ2-3), at 300ng/µl were combined in injection buffer. w^{1118} embryos, staged between 30 and 45 minutes AED at 18°C, were aligned on a strip of non-toxic rubber cement (Earth), in a humidified room to prevent excessive desiccation, and then covered with a drop of light paraffin oil. The posterior end of each embryo was then micro-injected with the above DNA mixture and the embryos were left at 18°C in a humidified chamber to hatch and crawl into a yeast paste encircling them.

b) screening for transformants

Adults that developed from injected embryos were individually crossed to w^{1118} virgins or males allowing transformed lines to be identified amongst the progeny by the w^+ eye colour marker. Eye colours obtained varied from pale yellow to strong orange but were consistent within each sex for each independent event.

c) generating stably transformed lines

Independent transformants were crossed to the doubly balanced stock, w^{1118};+/CyO;*Df(3R)ro*^{XB3}/TM6b, *Hu*. Male transformant flies carrying the CyO and TM6b chromosomes were selected and back crossed to w^{1118} virgins in the next generation. The progeny of this cross were scored to determine whether the P-element insert was segregating

from either the second chromosome by the absence of w^+ Cy progeny, or the third chromosome by the absence of w^+ *Hu* progeny. P-element insertions on the X chromosome were identified by the absence of w^+ male progeny. Any lines not segregating with one of these three chromosomes were assigned to the fourth chromosome and discarded as long as at least three other insertion events not on the fourth chromosome were identified. Once the chromosome of insertion was determined, stable lines were generated by homozygosing the P-element insert, or if this was lethal, maintaining the insertion over a balancer chromosome such as CyO or TM6b.

2-3.20 - Regulatory considerations

All manipulations involving recombinant DNA were carried out in accordance with the regulations and approval of the Genetic Manipulation Advisory Committee and the University Council of the University of Adelaide.

All manipulations involving animals were carried out in accordance with the regulations and approval of the Animal Ethics Committee and the University Council of the University of Adelaide.

Chapter 3: Characterisation of the interaction between PCL and E(Z).

3-1 Introduction

To examine the mechanism of PcG action, interactions between members of the PcG need to be characterised. The aim of the work presented in this thesis is to identify interactions between PCL and members of the PcG and to characterise the regions of PCL responsible for mediating these interactions, with the hope of shedding light on the role of PCL in the PcG complex. Previous work by Robert (1997) identified an interaction between PCL and E(Z) using the LexA yeast two hybrid assay (Gyuris *et al.*, 1993). This interaction was found to be mediated through a conserved domain (cDOM) of PCL (aa 424-605), which is defined on the basis of sequence homology with two potential mammalian PCL homologues, MTF2 and PHF1 (Coulson *et al.*, 1998). This region of similarity spans two PHD fingers and a region C-terminal to the fingers, which was been named the Polycomblike extended homology domain (PEH domain) (Figure 3.1). The function of the PHD finger is not yet known although it has been suggested that it may act as a protein/protein interaction motif or a DNA binding domain (Aasland *et al.*, 1995). This chapter describes further characterisation of the interaction between PCL and E(Z) and shows that this interaction is mediated through the PHD fingers of PCL and not the region of extended homology. This chapter also describes the conservation of this interaction between human homologues of PCL and E(Z).

3-2 PCL and E(Z) associate *in vivo*.

To confirm that the interaction between PCL and E(Z) is real and not an artefact of the yeast two hybrid assay, co-immunoprecipitation from embryonic protein extracts was performed. Affinity purified anti-PCL antiserum was used to precipitate PCL and associated proteins from extracts of 6-18 hour embryos. Anti-E(Z) was used in a similar manner to precipitate E(Z). The precipitated protein samples were electrophoresed on a SDS-page gel. In a separate lane, embryonic extract was also electrophoresed, providing a positive control for the size and presence of the approximately 90kDa E(Z) protein (Figure 3.2 Lane 1). The gel was immunoblotted and probed with anti-E(Z) antibody (Figure 3.2). As a control, a mock immunoprecipitation was performed using protein A alone with no antibody (Figure 3.2 Lane 2). No E(Z) was detected in this lane. As a positive control anti-E(Z) was able to immunoprecipitate E(Z) protein (Figure 3.2 Lane 3). Anti-PCL was able to co-immunopreciptiate E(Z) (Figure 3.2, Lane 4) indicating an *in vivo* association between PCL and E(Z). A co-immunoprecipitation experiment was also performed in the opposite direction but anti-E(Z) was unable to precipitate PCL (data not shown).

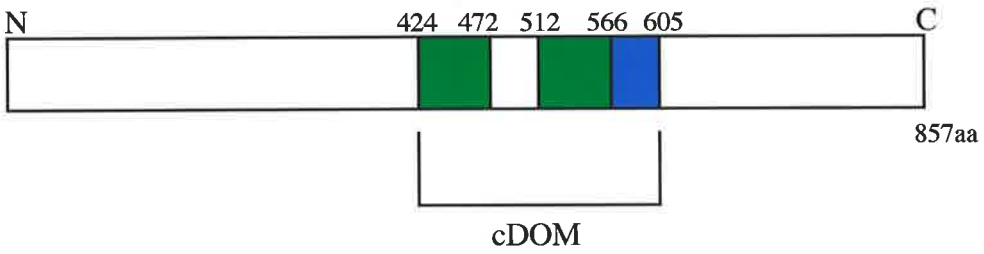


Figure 3.1: A schematic representation of PCL and the conserved domain (cDOM), which was defined by similarity between PCL and two potential mammalian homologues. The cDOM is responsible for mediating the interaction between PCL and E(Z) (Roberts, 1997) and encompasses PHDF1, PHDF2 (both shown in green) and the region of extended homology (PEH domain) C-terminal to PHDF2 (shown in blue). The position of the domains within the PCL protein are indicated in amino acids.

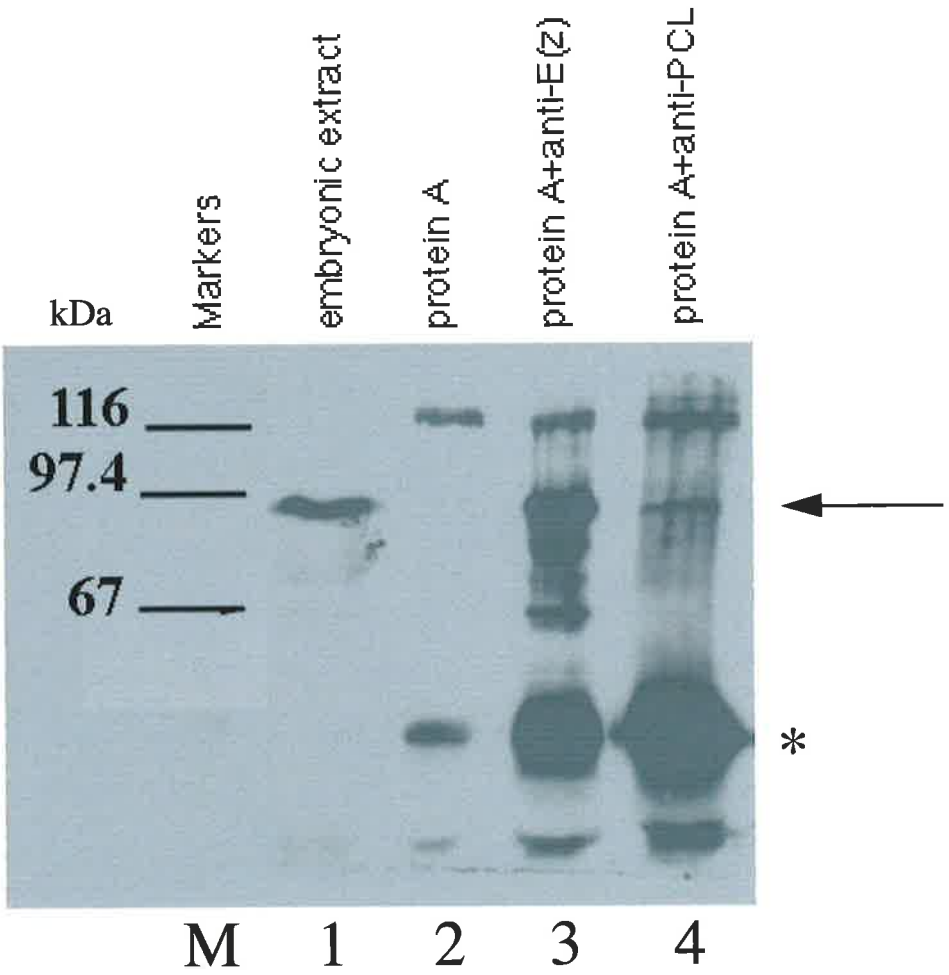


Figure 3.2:

Co-immunoprecipitation of embryonic extracts showing that PCL and E(Z) are associated *in vivo*. Affinity purified anti-PCL antiserum or anti-E(Z) antiserum were used to precipitate 6-18 hour embryonic extracts. The precipitated protein samples were electrophoresed in a 10% SDS gel, immunoblotted and probed with anti-E(Z) antiserum. Embryonic extract was included as a positive control for the presence and size of E(Z), which is 90kDa (Lane 1). Lane 2: precipitate of embryonic extract with protein-A beads alone. Lane 3: embryonic extract immunoprecipitated with anti-E(Z) antiserum. Lane 4: embryonic extract immunoprecipitated with anti-PCL antiserum. The bands corresponding to E(Z) are indicated by an arrow. Signals marked with an asterisk are due to cross-reactivity of the secondary antibody with the primary antibody used in the immunoprecipitation. The markers are in kilodaltons.

3-3 The interaction between PCL and E(Z) is mediated through the PHD fingers of PCL.

To determine whether the interaction between PCL and E(Z) is mediated through the PHD finger region or the region of extended homology (PEH domain) within PCL (Figure 3.1), the LexA yeast two hybrid assay was employed. Briefly, this involved cloning fragments of the PCL cDOM into the LexA DNA binding domain-encoding vector pEG202 to generate LexA-PCL fusions and testing for an interaction with a galactose inducible activation domain-E(Z) fusion construct (AD-E(Z)), in the *S. cerevisiae* EGY48 strain. EGY48 contains 6 LexA binding sites upstream of the endogenous *LEU2* gene and an interaction between the LexA-PCL fusion and the AD-E(Z) fusion can be identified by the ability to grow in the absence of exogenously supplied leucine.

The generation of all the LexA-PCL fusion constructs is described in section 2-2.12. All constructs were generated using high fidelity PCR, and the LexA-PCL junction was sequenced to ensure that an in-frame fusion protein was generated. The constructs were transformed into EGY48 and western analysis was performed using anti-LexA antisera to ensure that the constructs were being expressed. It had previously been shown that AD-E(Z) was being expressed (R. Jones, pers. comm.). All constructs were also tested for self-activation of the *LEU2* reporter gene and were unable to activate unless otherwise stated.

To determine if the PEH region is important in the interaction between PCL and E(Z), LexA-ΔEH was generated by PCR (see 2-2.12). This construct contains a truncated cDOM which removes the region of extended homology (residues 566-605)(Figure 3.1). When LexA-ΔEH was tested for its ability to interact with AD-E(Z), growth was observed on leucine deficient medium, indicating that the region of extended homology is not required for the interaction between E(Z) and PCL (Figure 3.3).

To confirm the role of the PHD fingers in mediating the interaction between PCL and E(Z), PHD finger mutants were generated using *in vitro* site-directed mutagenesis (Figure 3.4). As previously discussed in Section 1-2.1, the PHD finger contains a Cys_4-His-Cys_3 motif which is thought to co-ordinate Zn^{2+} (Aasland *et al.*, 1995). No structural studies have been performed on this recently defined domain and it is not known whether the conserved spacing of the Cys and His residues is important. Whether the Cys residues form hydrogen bonds to co-ordinate Zn^{2+} or covalently bind to the Zn^{2+} is also not known. Presumably, these Cys and His residues have a structural role. Mutating one of these highly conserved residues would therefore be expected to destroy the ability to co-ordinate Zn^{2+} and prevent the interaction between PCL and E(Z).

Although it could be presumed that mutating one of the conserved residues would destroy the function of the PHD finger, three different sets of mutagenesis were designed to ensure

that this was the case. The first mutagenesis converted the second conserved Cys in each PHD finger (residue 430 in PHDF1 and residue 518 in PHDF2), to a Ser which is likely to destroy the ability of that cysteine to covalently co-ordinate Zn^{2+}, but may not destroy the ability to form a hydrogen bond (Figure 3.4). These constructs were named LexA-C430S-cDOM and LexA-C518S-cDOM for mutations in PHDF1 and PHDF2 respectively. The second mutagenesis converted the second conserved Cys to an Ala in both PHDF1 and PHDF2 (residue 430 and 518 respectively). These constructs were called LexA-C430A-cDOM and LexA-C518A-cDOM respectively (Figure 3.4). The conversion of a conserved Cys to an Ala would be expected to destroy the ability of the PHD finger to chelate Zn^{2+} and the ability to form hydrogen bonds.

Apart from the conserved Cys and His residues, Aasland *et al.* (1995) also noticed a region of highly conserved hydrophobicity adjacent to the third conserved Cys. This region corresponds to residues 439-442 in PHDF1 and residues 527-530 in PHDF2 (Figure 3.4). The significance of this stretch of residues is not known and in the PHD fingers of PCL, only two of the three residues are hyrdrophobic. But the conservation of this stretch of amino acids in many PHD finger-containing proteins suggests that it is of importance. The third mutagenesis was therefore designed to convert this stretch of hydrophobic residues to hydrophilic Gln, residues which would be expected to destroy the function of this stretch of amino acids. Gln was chosen as it is a hydrophilic amino acid but is not highly charged. In addition to altering the hydrophobic residues to Gln, the adjacent Cys was changed to an Ala. The hydrophobic cores of PHDF1 and PHDF2 were individually mutated to Gln, along with a conversion of the third Cys to an Ala, to generate LexA-*PHDF1-QQQA*-cDOM and LexA-*PHDF2-QQQA*-cDOM respectively.

All six constructs, LexA-C430S-cDOM, LexA-C518S-cDOM, LexA-C430A-cDOM, LexA-C518A-cDOM, LexA-*PHDF1-QQQA*-cDOM and LexA-*PHDF2-QQQA*-cDOM were tested individually for their ability to interact with AD-E(Z). All PHD finger mutants failed to interact with AD-E(Z) indicating that the PHD fingers are important in mediating the interaction between PCL and E(Z) and that both PHDF1 and PHDF2 are required for this interaction (Figure 3.5 (A) and 3.6 (A)). To ensure that the lack of growth on GAL plates was not due to the absence of yeast being streaked onto the plate, a GLU LEU control plate was always streaked after the GAL plate (Figure 3.5 and 3.6). The lack of growth on the GAL plate is therefore due to the inability of the LexA-PCL fusion/AD-E(Z) to activate the *LEU2* reporter gene and not due to the absence of yeast being streaked onto the plate. Western analysis was performed on each of the cDOM mutants using anti-LexA antisera to confirm that the LexA fusions were being expressed and that the inability of these constructs to interact with E(Z) was due to the mutation in the PHD fingers and not due to differences in the stability of the mutant proteins (Figure 3.5 (B) and 3.6 (B)).

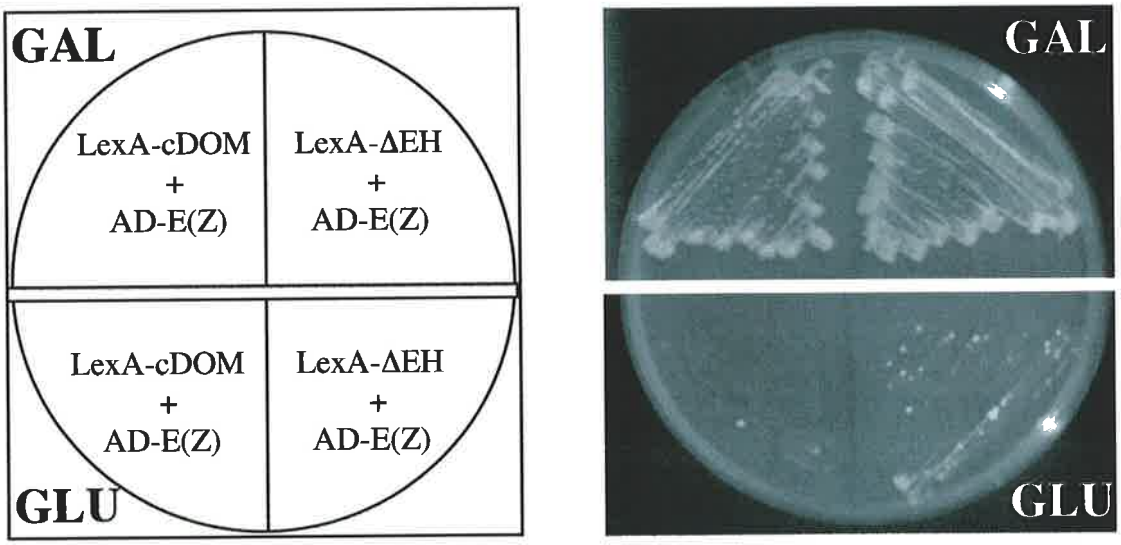


Figure 3.3: The extended homology domain is not required for the interaction between PCL and E(Z). LexA-ΔEH is able to interact with AD-E(Z) leading to activation of the *LEU2* reporter gene. This activation is observed as growth on media lacking leucine but containing galactose (GAL). Interactions are tested on both GAL and GLU plates to ensure that the interaction is dependent on the induction of the AD-E(Z) fusion which is under the control of a galactose inducible promoter.There is a small level of background growth on the LexA-ΔEH + AD-E(Z) GLU plate. This is insignificant compared to the growth observed on the GAL plate and is due to LexA-ΔEH being able to weakly activate *LEU2* expression in the absence of AD-E(Z) expression. As a positive control for interaction the full length LexA-cDOM construct was tested against AD-E(Z).

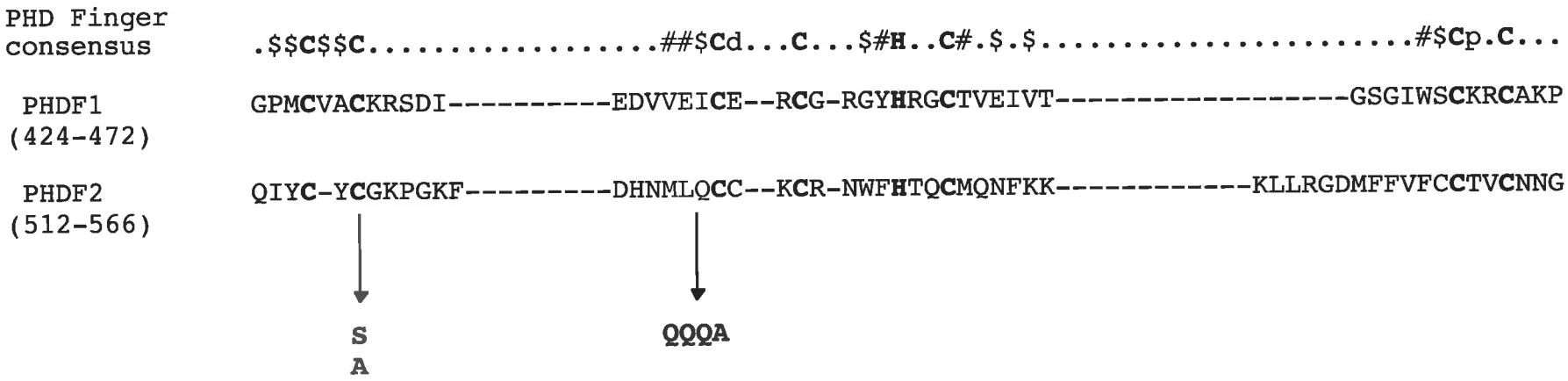
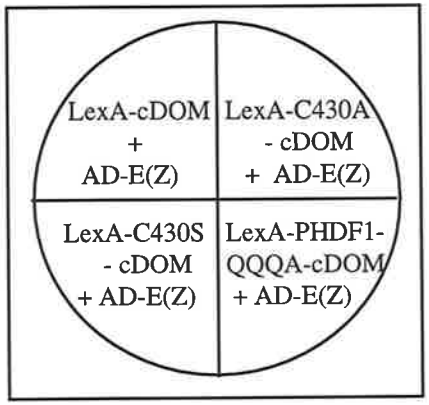


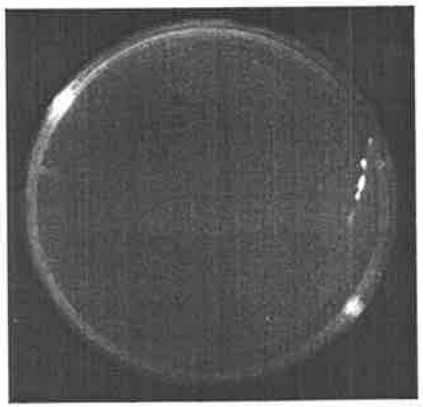
Figure 3.4: An alignment of PHD fingers 1 and 2 with the consensus PHD finger motif (Aasland *et al.,* 1995) and a schematic of the *in vitro* mutagenesis strategy used to analyse the interaction between PCL and E(Z). The conserved Cys and His residues are shown in bold and the highly conserved hydrophobic residues, including the adjacent Cys are shown in green. The second conserved Cys, shown in blue, was mutated to a Ser and an Ala (shown in red) to generate C430S, C430A, C518S and C518A. The three hydrophobic amino acids and the third conserved Cys were mutated to a QQQA respectively and are shown in magenta to generate PHDF-1QQQA and PHDF-2QQQA. The amino acids encoding the PHD finger motifs are in brackets. The consensus symbols are: # strongly conserved hydrophobicity and $ highly conserved hydrophobicity.

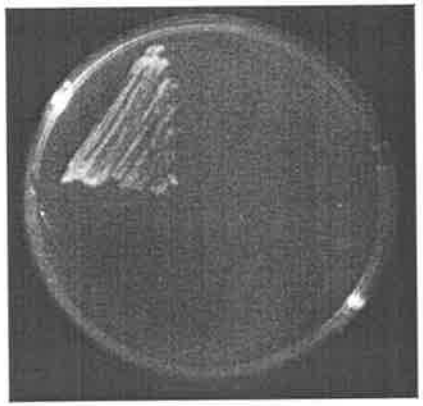
Figure 3.5: Mutations in the first PHD finger of the cDOM abolish the interaction with E(Z).

Three PHDF1 cDOM mutants were generated (for a description of the mutants see text) and cloned into pEG202 to generate LexA-C430A-cDOM, LexA-C430S-CDOM and LexA-PHDF1-QQQA-cDOM. These modified proteins were then tested for their ability to interact with AD-E(Z). Three independant transformations were performed and three colonies were picked and streaked onto yeast plates. (A) All three PHDF1 cDOM mutants were unable to interact with AD-E(Z) as indicated by the lack of growth on the GAL plate. The GLU LEU plates were streaked last to ensure that yeast were being streaked onto each plate. LexA-cDOM was used as a positive control as it had previously been shown to interact with AD-E(Z). (B) Western blot of protein extracts from yeast containing each of the three PHDF1-cDOM mutants, wildtype LexA-cDOM and LexA-PHDF2, electrophoresed in a 12.5% SDS-polyacrylamide gel and probed with anti-LexA antiserum. An equal amount of the cDOM proteins were added in each lane. Lane 1 and Lane 2 contain LexA-PHDF2 (size is estimated to be 32kDa) and LexA-cDOM (42kDa) respectively, which act as positive control proteins. Each cDOM mutant was expressed at a level equivalent to non-mutant protein levels (Lanes 3-5). The size of the markers are in kilodaltons.

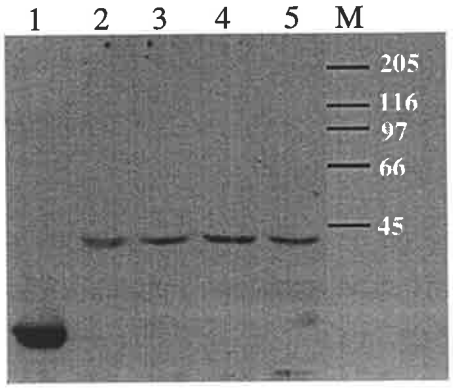
A







B



Lane 1: LexA-PHDF2 as a positive control

2: LexA- cDOM

3: LexA-C430A-cDOM

4: LexA-C430S-cDOM

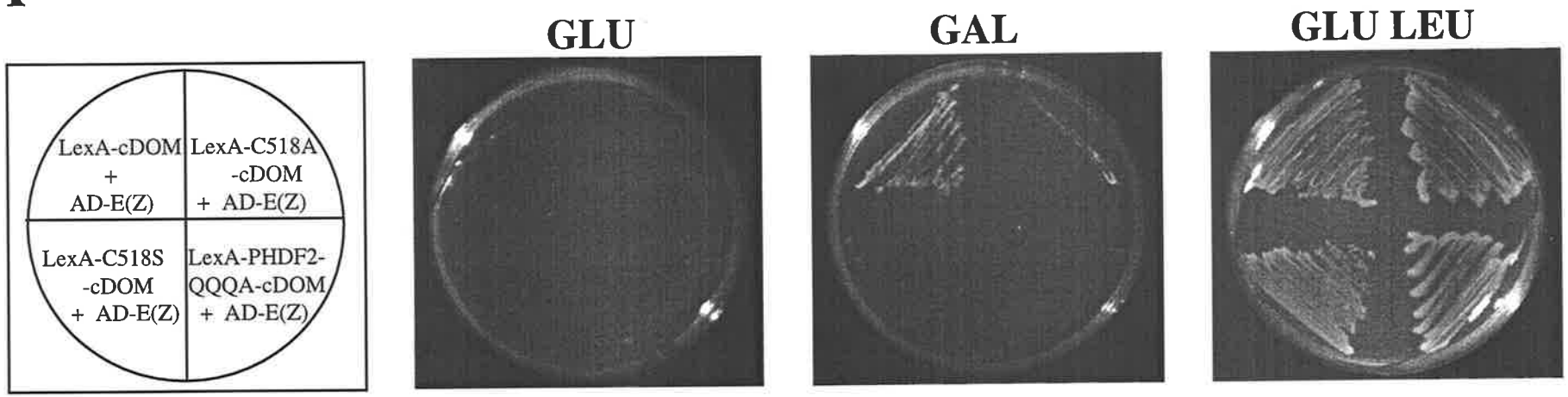
5: LexA-PHDF1-QQQA-cDOM

M: Molecular weight markers in kDa.

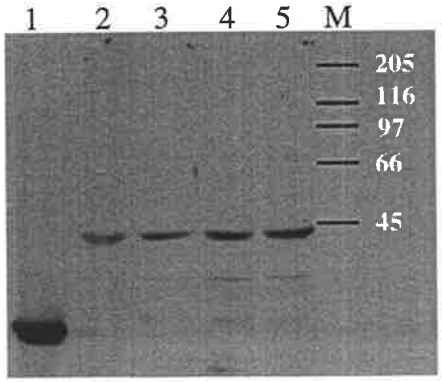
Figure 3.6: Mutations in the second PHD finger of the cDOM abolish the interaction with E(Z).

Three PHDF2 cDOM mutants were generated (for a description of the mutants see text), cloned into pEG202 to generate LexA-C518A-cDOM, LexA-C518S-cDOM and LexA-PHDF2-QQQA-cDOM and the resultant mutant proteins tested for their ability to interact with AD-E(Z). Three independent transformations were performed and three colonies from each were picked and streaked onto yeast plates. (A) All three PHDF2 cDOM mutants were unable to interact with AD-E(Z) as indicated by the lack of growth on the GAL plates. The GLU LEU plate was streaked last to ensure that yeast were being streaked onto each plate. LexA-cDOM was used as a positive control as it had previously been shown to interact with AD-E(Z). (B) Western blot of protein extracts from yeast containing each of the three PHDF2-cDOM mutants, non-mutant cDOM and non-mutant PHDF2 probed with LexA antibody. An equal amount of protein extracts was added for each of the cDOM lanes (both normal and mutant). The estimated size of the LexA-cDOM fusions is 42kDa. Lane 1 and 2 contain LexA-PHDF2 and LexA-cDOM respectively, which serve as positive control proteins. Each cDOM mutant was expressed at a level equivalent to the non-mutant protein (lanes 3-5) and therefore the inability of the PHDF2-cDOM mutants to interact with AD-E(Z) is due to the engineered mutation in PHDF2. The size of the molecular weight markers (M) is shown in kilodaltons.

A



B



Lane 1: LexA-PHDF2 as a positive control

2: LexA-cDOM

3: LexA-C518A-cDOM

4: LexA-C518S-cDOM

5: LexA-PHDF2-QQQA-cDOM

M: Molecular weight markers in kDa.

3-4 A single PHD finger is able to interact with E(Z).

Given that a mutation in either PHD finger was able to destroy the interaction between the cDOM of PCL and E(Z) it was of interest to determine whether a single PHD finger was able to mediate an interaction with E(Z). To test this, the region encoding the minimal PHD finger (Aasland *et al.*, 1995), for both PHDF1 and PHDF2 was cloned into pEG202 to generate LexA-PHDF1 and LexA-PHDF2 (residues 424-472 and 512-566 respectively). These fusions were then tested for their ability to interact with AD-E(Z) by assaying for growth on leucine deficient medium. Both LexA-PHDF1 and LexA-PHDF2 were able to grow on leucine deficient media indicating that they can interact with AD-E(Z) (Figure 3.7). The interaction between LexA-PHDF1 and AD-E(Z) appeared to be considerably stronger than the interaction between LexA-PHDF2 and AD-E(Z) in this assay (Figure 3.7). Whether or not this reflects a quantitative difference in their ability to interact with E(Z) is not known, although it has been noted previously there appears to be a general correlation between the strength of an interaction identified in the yeast two hybrid and that of an *in vitro* interaction (Estojak *et al.*, 1995).

To ensure that the PHD fingers are responsible for mediating the interaction between PCL and E(Z), individual mutant PHD finger constructs were generated using the mutant forms described in section 3-3 as a template. LexA-C430S-PHDF1, LexA-C518S-PHDF2, LexA-C430A-PHDF1, LexA-C518A, LexA-PHDF1-QQQA-PHDF1 and LexA-PHDF2-QQQA-PHDF2 were then tested for their ability to interact with AD-E(Z). As shown in Figure 3.8 and Figure 3.9, the results were complex. Surprisingly, no difference was seen in the interaction between AD-E(Z) and LexA-PHDF1, LexA-C430S-PHDF1 or LexA-C430A-PHDF1 (Figure 3.8). The interaction between LexA-PHDF1-QQQA-PHDF1 and AD-E(Z) is dramatically reduced in comparison to LexA-PHDF1 and AD-E(Z) but growth was still observed on GAL plates. A slight reduction in the interaction between LexA-C518A-PHDF2 and AD-E(Z) was observed when compared to the interaction between LexA-PHDF2 and AD-E(Z) (Figure 3.9). The interaction between LexA-C430S-PHDF2 and AD-E(Z) was weaker than the LexA-C518A-PHDF2 and AD-E(Z) interaction (Figure 3.9). No interaction was observed between LexA-PHDF2-QQQA-PHDF2 and AD-E(Z) (Figure 3.9). All interactions with the PHD finger mutants were tested by three independent transformations, with at least 10 colonies being tested for each transformation. Western analysis using LexA antisera was performed on each of the mutants to confirm that the LexA fusions were being expressed and that the inability of the mutant proteins to interact with E(Z) was due to the mutation in the PHD fingers and not to a difference in their level of expression compared to the non-mutant protein level (Figure 3.8 (B) and Figure 3.9 (B)).

Although no structural information has yet been published on the PHD finger motif, the mutageneses that were carried out were expected to destroy the interaction between PCL and

E(Z). Given that some mutants retained there ability to interact with E(Z), it was possible that the mutants generated were 'sticky' and able to interact non-specifically with proteins. To test this, LexA-PHDF1 and LexA-C430S-PHDF1 were tested for non-specific interactions with four AD-fusion constructs; AD-esc, AD-PscΔB, AD-Pc and AD-Su(Z)2. As shown in Table 3.1 no interactions were observed between either LexA-PHDF1 or LexA-C430S-PHDF1 and the AD-PcG constructs tested, ruling out the possibility that the interaction between the mutant PHD fingers and AD-E(Z) is due to general stickiness of the protein. It therefore remains unclear why the mutant PHD fingers continue to interact with E(Z).

	LexA-PHDF1	LexA-C430S-PHDF1
AD-ESC	-	-
AD-PSCdelB	-	-
AD-SU(Z)2	-	-
AD-PC	-	-

Table 3.1: A summary of the interactions between LexA-PHDF1, LexA-C430S-PHDF1 and four AD-PcG members. '-' indicates no growth was observed on leucine deficient media containing GAL.

3-5 The interaction between PCL and E(Z) is conserved in their human homologues.

Recent work has identified a mammalian PcG complex (reviewed in van Lohuizen, 1999). Many interactions that have been identified between *Drosophila* PcG members have also been observed between the mammalian counterparts (see section 1-1.2). It was therefore of interest to determine whether the interaction between PCL and E(Z) was conserved in their human counterparts. To date, two E(Z) homologues have been identified in humans; EZH1 and EZH2. Both show a high level of sequence similarity in four domains spanning the open reading frame (Labile *et al.*, 1997). Two potential PCL homologues have also been identified; PHF1 (Coulson *et al.*, 1998) and hMTF2 (M. Coulson pers. comm.). PHF1 is highly conserved (34% identity) over the region encoding the cDOM (Coulson *et al.*, 1998). The level of conservation of hMTF2 is yet to be determined because the clone containing hMTF2 has only been end sequenced and full length sequence is not yet available (clone z040a07 obtained from Research Genetics). However these regions are highly conserved with the mouse MTF2 which displays 34% identity to the cDOM region of PCL (Coulson *et al.*, 1998). To determine whether the interaction between PCL and E(Z) is conserved in their mammalian counterparts, the region spanning the two PHDFs of PHF1 and hMTF2 were

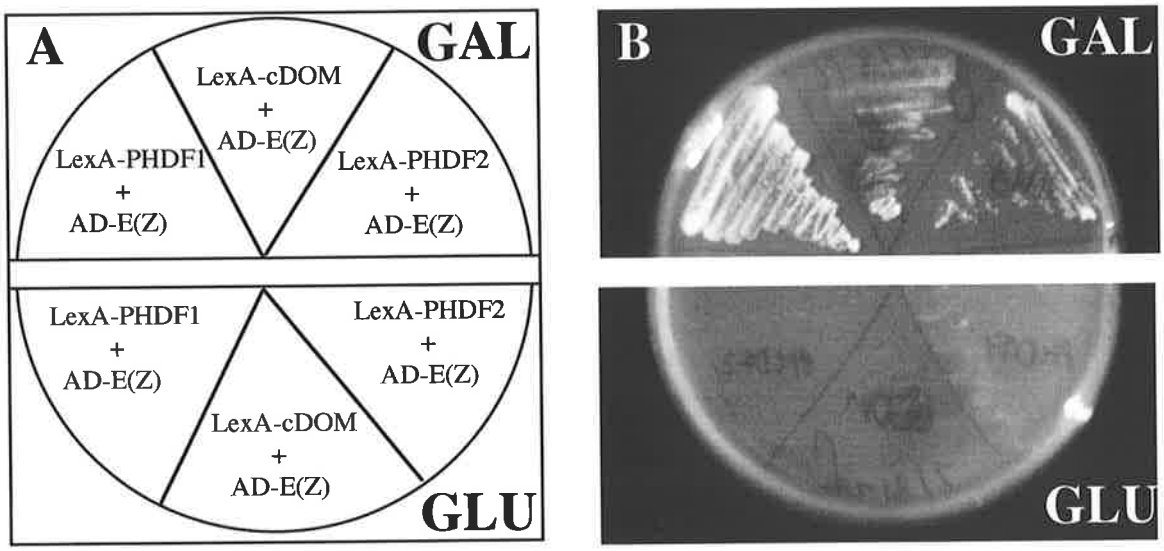


Figure 3.7: Both PHDF1 and PHDF2 can interact with E(Z).

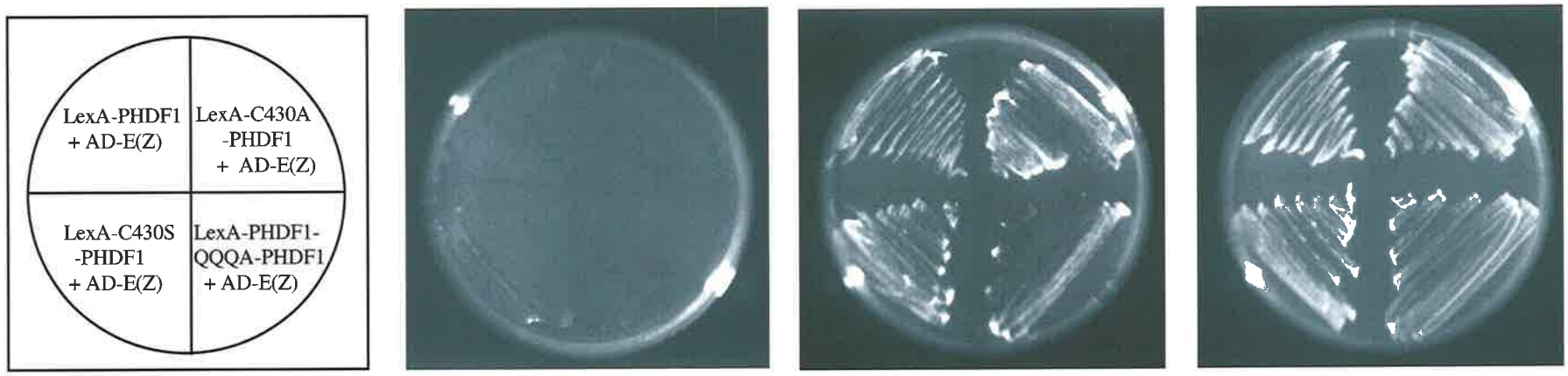
PHDF1 and PHDF2 were cloned into the yeast two hybrid vector pEG202 to generate LexA-PHDF1 and LexA-PHDF2 respectively. The resulting fusion proteins were then tested for an interaction with AD-E(Z). (A) is a schematic diagram of the arrangement on (B). The interaction was tested on GAL vs GLU plates to ensure that growth was dependent on the presence of AD-E(Z). (B) Both LexA-PHDF1 and LexA-PHDF2 can interact with AD-E(Z). This interaction is dependent on induction of the AD-E(Z) fusion, which is under galactose inducible control. Yeast containing both LexA-PHDF1 and AD-E(Z) consistently grew faster than yeast containing LexA-cDOM or LexA-PHDF1 and AD-E(Z).

Figure 3.8: Mutant PHDF1 proteins are still able to interact with AD-E(Z).

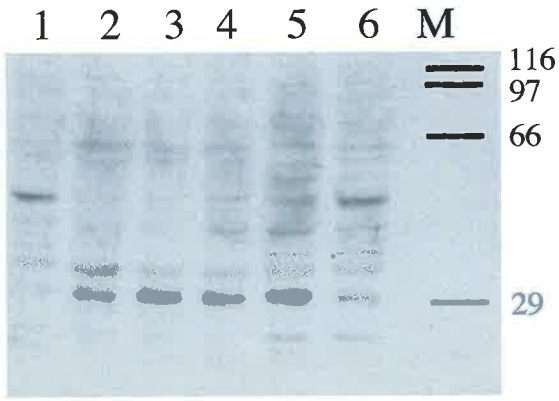
To ensure that the interaction between PCL and E(Z) is mediated through the PHD fingers of PCL, individual mutant PHDF1 constructs were generated. Using PHDF1 cDOM mutants as a template, PHDF1 was PCR amplified and cloned into pEG202 to generate LexA-C430A-PHDF1, LexA-C430S-PHDF1 and LexA-PHDF1-QQQA-PHDF1. The fusion proteins encoded by these constructs were then tested for their ability to interact with AD-E(Z) in the yeast two hybrid assay. (A) LexA-C430A-PHDF1 and LexA-C430S-PHDF1 interact with AD-E(Z) to a level comparable to that of the non-mutated PHDF1. LexA-PHDF1-QQQA-PHDF1 show a reduction in the level of interaction with AD-E(Z), when compared to the interaction with LexA-PHDF1.

(B) Western blot of protein extracts from each of the mutant PHDF1 yeast strains, LexA-PHDF1 and LexA-cDOM probed with LexA antiserum. Lane 1 and 6 contain LexA-cDOM which acts as a positive control for the antisera. Lane 2 contains LexA-PHDF1 which allows comparison of protein levels with the mutant LexA-PHDF1 proteins. Lanes 3-5 contain protein extracts from each of the three PHDF1 mutant yeast strains. The expected size of the LexA-PHDF1 and the mutant LexA-PHDF1 constructs is 32kDa. No significant difference in the level of protein expression is observed between the mutant PHDF1 constructs and that of the non-mutant construct. The molecular weight markers are in kilodaltons (M).

A



B



1. LexA-cDOM
2. LexA-PHDF1
3. LexA-C430A-PHDF1
4. LexA-C430S-PHDF1
5. LexA-PHDF1-QQQA-PHDF1
6. LexA-cDOM

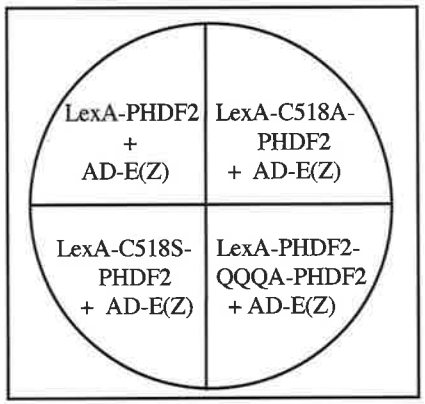
M. Markers in kilodaltons

Figure 3.9: The interaction between LexA-PHDF2 and AD-E(Z) constructs is reduced or abolished in the mutant LexA-PHDF2 constructs.

To ensure that the interaction between PCL and E(Z) is mediated through the PHD fingers of PCL, individual mutant PHDF2 constructs were generated. Using the PHDF2-cDOM mutants as a template, PHDF2 was PCR amplified and cloned into pEG202 to generate LexA-C518A-PHDF2, LexA-C518S-PHDF2 and LexA-PHDF2-QQQA-PHDF2. The fusion proteins produced by these constructs were then assayed for their ability to interact with AD-E(Z) in the yeast two hybrid. (A) LexA-C518A-PHDF2 showed a slightly reduced interaction with AD-E(Z). LexA-C518S-PHDF2 was further reduced in its ability to interact with AD-E(Z) when compared to the interaction of LexA-PHDF2. The interaction between LexA-PHDF2-QQQA-PHDF2 and AD-E(Z) was abolished entirely as no growth was observed on GAL plates.

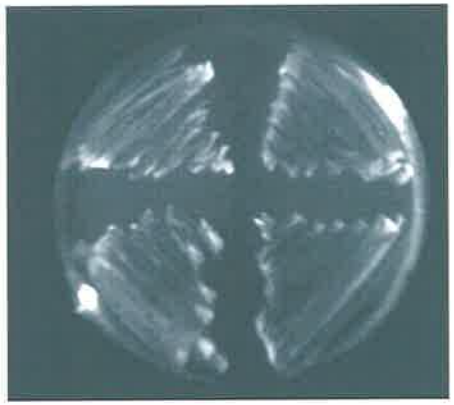
(B) Western blot of protein extracts from each of the mutant yeast strains and the LexA-PHDF2 yeast strain probed with LexA antisera. Lane 1 contains extract from a strain containing LexA-cDOM and acts as a positive control for the LexA antisera. Lane 2 contains LexA-PHDF2 protein extract and enables a comparison to be made between the levels of the wild type protein and that of the mutant proteins in Lanes 3-5. The estimated size of the LexA-PHDF2 fusion proteins is 32kDa. No difference in the expression of the mutant LexA fusions compared with that of the LexA-PHDF2 is seen.

A



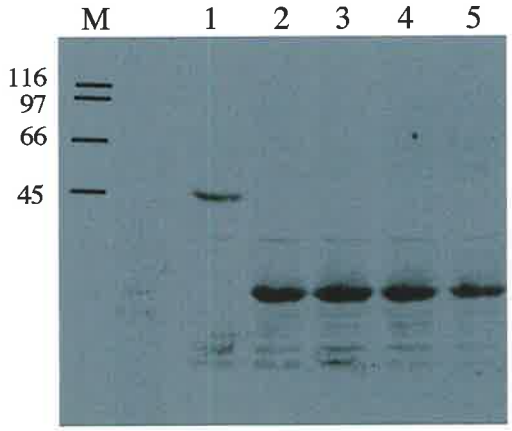
GLU

GAL



GLU LEU

B



M. Molecular weight markers in kilodaltons

1. LexA-cDOM
2. LexA- PHDF2
3. LexA-C518A-PHDF2
4. LexA-C518S-PHDF2
5. LexA-PHDF2-QQQA-PHDF2

cloned into pEG202 to generate LexA-PHF1 and LexA-hMTF2 respectively. The sequence information obtained from the Human Genome Project was sufficient to generate primers to enable PCR amplification of hMTF2. Full length EZH1 and EZH2 were cloned into the yeast two hybrid vector pJG4-5 to generate AD-EZH1 and AD-EZH2 respectively. Both LexA-PHF1 and LexA-hMTF2 were transformed into EGY48 and tested for their ability to self-activate the *LEU2* reporter gene. LexA-PHF1 was unable to activate the reporter gene, whereas LexA-hMTF2 was able to strongly activate it. Self-activation of LexA-hMTF2 also occurred in EGY191, a yeast strain that only contains two LexA binding sites upstream of the *LEU2* reporter gene.

To overcome the problem of activation of the reporter gene, EZH1 and EZH2 were cloned into pEG202 to generate LexA-EZH1 and LexA-EZH2 respectively. Both LexA-EZH1 and LexA-EZH2 were tested for self-activation in EGY48. LexA-EZH1 did not self-activate the reporter gene, whereas LexA-EZH2 did and was therefore transformed into EGY191. LexA-EZH2 did not activate transcription in this yeast strain The PHD finger region of hMTF2 was cloned into pJG4-5 to generate AD-hMTF2.

Because LexA-PHF1 did not self-activate, it was tested for an interaction with AD-EZH1 and AD-EZH1, while AD-hMTF2 was tested for an interaction with LexA-EZH1 and LexA-EZH2. An interaction was observed between LexA-PHF1 and AD-EZH2 (Figure 3.10). No interaction was observed between LexA-PHF1 and AD-EZH1. Likewise, no interaction was observed between AD-hMTF2 and LexA-EZH1 or LexA-EZH2 (Figure 3.10). Western analysis using anti-LexA- antiserum showed that LexA-EZH1 and LexA-EZH2 were being expressed (Figure 3.10). Western analysis using anti-HA antisera showed that AD-EZH1 and AD-hMTF2 were being expressed (Figure 3.10).

3-6 The PHD fingers of Trithorax do not interact with E(Z).

As discussed below, E(Z) displays characteristics of both a PcG and trxG member, depending on the circumstances. It is therefore possible that E(Z) interacts with a motif that is present in both a PcG member and a trxG member. One such motif that is found in a member from each group is the PHD finger motif. PCL has two and TRX has three PHD fingers (Aasland *et al.*, 1995). Evidence presented above shows that the PHD finger motifs of PCL are responsible for mediating the interaction with E(Z), raising the possibility that the PHD fingers of TRX mediate an interaction between TRX and E(Z). The interaction of PCL with E(Z) could lead to recruitment of E(Z) to PcG complexes where E(Z) would play a role in repressing gene transcription, whereas the interaction between the PHD fingers of TRX and E(Z) would lead to recruitment of E(Z) to complexes that activate gene transcription. To test this hypothesis the three PHD fingers of TRX (aa1266-4442) were cloned into the yeast two hybrid vector pEG202 to generate a LexA-TRX-PHDF construct. This construct was

then transformed into EGY48 and tested for the ability to self-activate the *LEU2* reporter gene. LexA-TRX-PHDF was able to activate reporter gene expression and was therefore transformed into EGY191, containing fewer LexA binding sites, and when tested for self-activation was unable to do so. AD-E(Z) was subsequently transformed into yeast and an interaction between LexA-TRX-PHDF and AD-E(Z) assayed. As shown in Figure 3.11, no interaction was observed between TRX and E(Z). The absence of an interaction between the TRX PHD fingers and E(Z) further demonstrates the specificity of the interaction between the PHD fingers of PCL and E(Z).

3-7 Discussion

In order to identify the functional role of PCL within the PcG complex it is first necessary to identify and characterise its molecular interactions with other members of the group. Robert (1997) identified an interaction between the conserved domain of PCL (cDOM) and Enhancer of zeste (E(Z)) using the yeast two hybrid assay. Although the yeast two hybrid assay is an excellent tool which allows identification of protein interactors, interactors identified using this assay must be tested in other assays to ensure that the interactions reflect the *in vivo* behaviour.

The co-immunoprecipitation experiment described in this chapter confirmed that PCL and E(Z) associate *in vivo*. The co-immunoprecipitations were performed in both directions. The PCL antiserum was able to co-precipitate E(Z), but the E(Z) antiserum was unable to precipitate PCL. The reason for this is not apparent as both antibodies were polyclonal and functional. The PCL antiserum was able to precipitate E(Z), and the E(Z) antibody was able to precipitate E(Z) but not PCL. Co-immunoprecipitation of embryonic extracts only confirms that the proteins of interest are present in the same protein complex. As PCL and E(Z) are thought to be part of a huge protein complex *in vivo*, a GST-pulldown has been used to confirm that the interaction between PCL and E(Z) is direct and not mediated through an intermediate protein (R. Jones, pers. comm.). These three results; the yeast two hybrid, the co-immunoprecipitation and the 'GST-pulldown' confirm that the interaction between PCL and E(Z) is direct and occurs *in vivo*.

3-7.1 E(Z) and PCL interact through the PHD fingers of PCL

Once the interaction between PCL and E(Z) was confirmed, a further set of experiments were performed to determine which region of the cDOM was responsible for the interaction with E(Z). The cDOM of PCL consists of two PHD fingers and a region of extended homology termed the Polycomblike Extended Homology (PEH) domain. The PHD fingers are a recently identified protein motif (Aasland *et al.*, 1995). At the start of this work, it was not known whether they mediated protein/protein interactions or protein/DNA interactions. Yeast two hybrid experiments presented in this chapter demonstrate that they are responsible

Figure 3.10: PHF1 interacts specifically with EZH2.

To determine whether the interaction between PCL and E(Z) is conserved in their human homolougues a yeast two hybrid analysis was performed. The PHD finger region of PHF1, EZH1 and EZH2 were cloned into pEG202 to generate LexA-PHF1, LexA-EZH1 and LexA-EZH2. The PHD finger region of hMTF2, EZH1 and EZH2 were cloned into pJG4-5 to generate AD-hMTF2, AD-EZH1 and AD-EZH2. (A) LexA-PHF1 was tested for an interaction with AD-EZH1 and AD-EZH2. AD-hMTF2 was tested for an interaction with LexA-EZH1 and LexA-EZH2. An interaction was observed between LexA-PHF1 and AD-EZH2 when comparing growth on GAL versus GLU medium. No other interaction was observed.

(B) Western blot analysis of protein extracts from the AD-EZH1 and AD-hMTF2 yeast strains probed with HA antiserum. Both proteins are being expressed. AD-EZH1 is slightly larger than the expected size of 94kDa, which is most likely due to post-translational modification. AD-hMTF2 is approximately 30kDa, however the estimated size of this protein is difficult to predict due to the cDNA not being fully sequenced.

(C) Western blot analysis of protein extracts from LexA-EZH1 and LexA-EZH2 yeast strains probed with LexA antiserum. Both proteins are being expressed, and are slightly larger than the expected sizes of 108kDa. This is most likely due to post-translational modification.

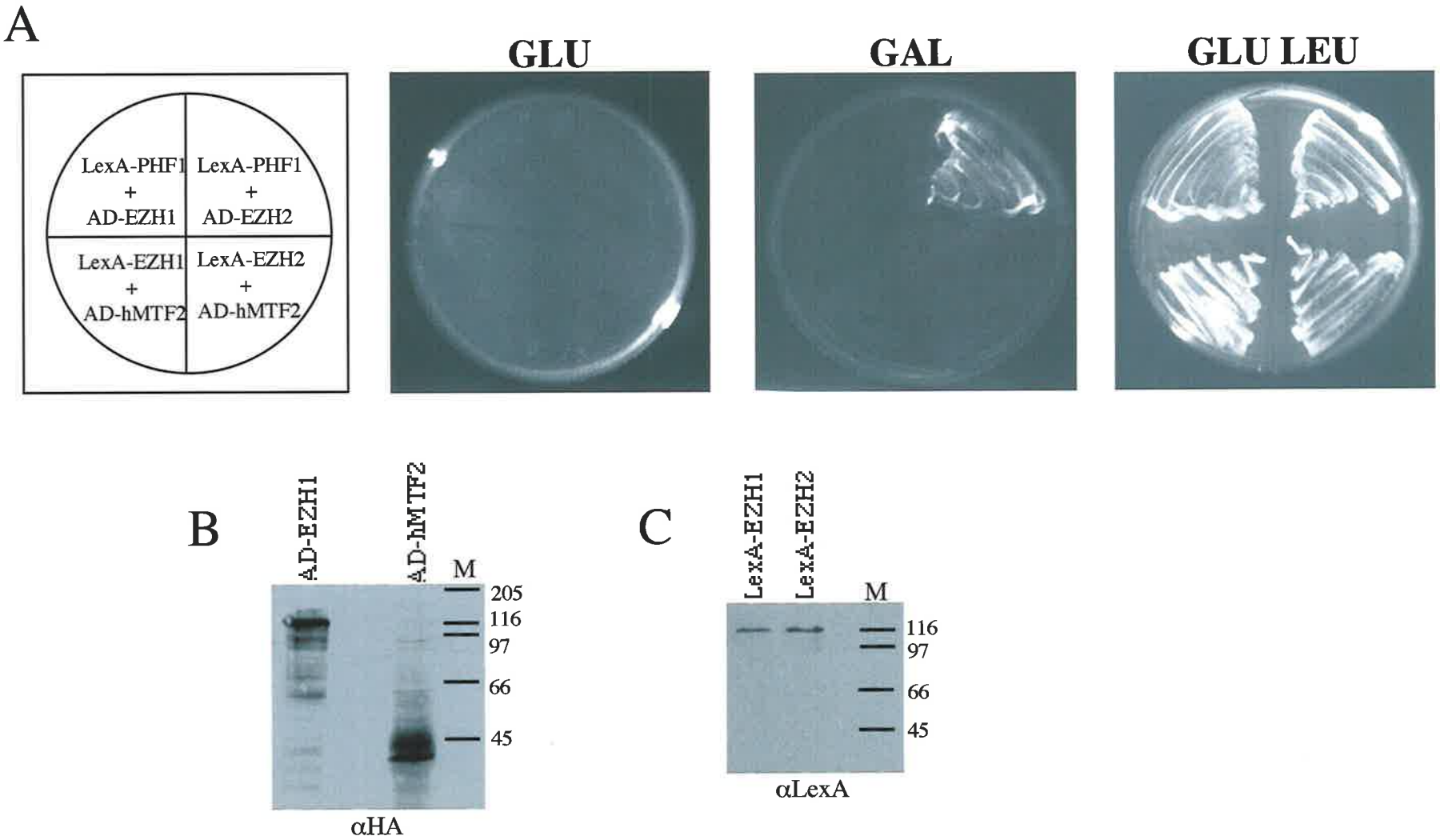

A
GLU
GAL
GLU LEU
LexA-PHF1 + AD-EZH1
LexA-PHF1 + AD-EZH2
LexA-EZH1 + AD-hMTF2
LexA-EZH2 + AD-hMTF2
B
AD-EZH1
AD-hMTF2
M
205
116
97
66
45
αHA
C
LexA-EZH1
LexA-EZH2
M
116
97
66
45
αLexA

Figure 3.11: The PHD fingers of Trithorax do not interact with E(Z).
To determine whether the PHD fingers of TRX interact with the PHD fingers of E(Z), the three PHD fingers of Trithorax were cloned into the yeast two hybrid vector, pEG202, to generate LexA-TRX-PHDFs. This construct was then tested for an interaction with AD-E(Z) in the yeast two hybrid. (A) a GLU control plate showing that LexA-TRX-PHDFs does not self-activate the expression of the *LEU2* reporter gene. (B) LexA-TRX-PHDFs are unable to interact with AD-E(Z). (C) a positive control growth plate streaked after (A) and (B), showing that yeast was streaked onto the GLU and GAL plates.

for mediating protein-protein interactions. Specifically, in PCL, they mediate the interaction with E(Z). The PEH domain is not required for the interaction with E(Z). Both PHDF1 and PHDF2 are able to mediate an interaction with E(Z), however the interaction between PHDF1 and E(Z) appears stronger than the interaction between PHDF2 and E(Z). Whether this represents a real difference in the ability of these PHDFs to interact with E(Z) is not known, but the strength of an interaction in the two hybrid assay generally correlates with the *in vitro* strength of an interaction (Estojak *et al.*, 1995).

Further confirmation of the role of the PHD finger motifs in the interaction with E(Z) came from *in vitro* mutagenesis of both PHDF1 and PHDF2. Mutations generated in the conserved residues of both PHDF1 and PHDF2 in the context of the entire cDOM abolished the interaction with E(Z). Although this demonstrates that the PHD fingers are important in mediating the interaction with E(Z), it also raises an interesting question given that both PHD fingers are able to interact independently with E(Z). Why can't the non mutated finger in the cDOM mediate an interaction with E(Z)? Perhaps the structure of the mutated PHD finger interferes with the structure of the non-mutated PHD finger. This can only be resolved with structural studies of the entire cDOM.

The mutations in the PHD fingers alone (not in the context of the entire cDOM) also generated conflicting results. LexA-C430A-PHDF1 and LexA-C430S-PHDF1 were not reduced in their ability to interact with AD-E(Z). This is in contrast to the results obtained with the equivalent PHDF2 mutants. The level of interaction between LexA-C518A-PHDF2 was reduced when compared to wild type and interestingly the level of interaction between LexA-C518S-PHDF2 was weaker than LexA-C518A-PHDF2. Why there should be a difference in the strength of an interaction between these two PHDF2 mutants when no difference was observed in the equivalent PHDF1 mutants is not known. Perhaps PHDF2 is more sensitive to mutations due to its weaker interaction with E(Z) when compared to PHDF1 (Figure 3.7). Consistent with this is the interaction of AD-E(Z) with LexA-PHDF2-QQQA-PHDF2, which was abolished while the interaction with LexA-PHDF1-QQQA-PHDF1 was only reduced when compared to the normal PHD fingers. Why mutations in the single PHD fingers did not abolish the interaction with E(Z), when the mutations in the context of the entire cDOM did, is not clear. Single PHD fingers are not simply non-specifically sticky as PHDF1 failed to interact with four other proteins tested. Although the PHD fingers have been taken out of context of the whole protein in the individual PHD finger constructs, it is expected that a mutation that abolishes the interaction in the context of a larger domain, would also abolish the interaction in a smaller domain. Until more is known about the structure of the PHD fingers, and the residues important in its structure and function are identified, these results cannot be interpreted. Although the results of the individual PHD finger constructs are difficult to interpret, the results of the mutant cDOM

constructs clearly demonstrate that the PHD fingers are responsible for mediating the interaction with E(Z), providing the first evidence that the PHD finge are responsible for mediating protein/protein interactions.

3-7.2 Interaction between human homologues of PCL and E(Z).

Experiments in the last few years have identified a mammalian PcG which appears to function in an analogous way to that of the *Drosophila* complex (reviewed in van Lohuizen, 1999). Many interactions that have been identified in *Drosophila* can be observed in their mammalian counterparts. Both human homologues of PCL; PHF1 and hMTF2, were tested for an interaction with the two human homologues of E(Z); EZH1 and EZH2, using the yeast two hybrid assay. An interaction was identified between PHF1 and EZH2 but not EZH1. No interaction was identified between hMTF2 and either EZH1 or EZH2 in this assay. There are three potential reasons why hMTF2 does not interact with either EZH1 or EZH2 when it might be expected to do so. Firstly, an interaction is not able to be detected in the yeast two hybrid assay due to improper folding of the fusion proteins. GST-pulldown analyses could be performed in an attempt to overcome this problem. Secondly, there may exist a level of specificity in the protein-protein interactions between members of the mammalian PcG, and hMTF2 may not interact *in vivo* with either EZH1 or EZH2, but with some other as yet unidentified E(Z) homologue. Some unpublished data exists to support specificity within the mammalian PcG complex (M. van Lohuizen pers. comm.). One theory on the evolution of the mammalian genome suggests that there were two genome duplication events in the vertebrate lineage so that for every one *Drosophila* gene there are four vertebrate counterparts (Sidow, 1996). There is evidence to suggest that not all *Drosophila* genes have four mammalian counterparts. For example, only two mammalian homologues of all of the *Drosophila* PcG genes exist so far. However only the complete sequence of the human genome can resolve this. It is possible that the specific interaction between PHF1 and EZH2 and not EZH1, is due to specificity within the mammalian PcG complex, in which PHF1 associates with EZH2 but no other mammalian homologue. Identification of all potential E(Z) homologues together with *in vivo* and *in vitro* experiments confirming specific interactions, should be performed to test this theory. The third possible reason that hMTF2 may not interact with either EZH1 or EZH2 in the yeast two hybrid assay is that hMTF2 may have a PcG independent function that does not involve an interaction with EZH1 or EZH2. To date there is no evidence to support this.

The interaction between PHF1 and EZH2 suggests that the PCL-E(Z) interaction has been conserved in the evolution of the arthropod and chordate lineages, indicating that it is an important part of the function of both the mammalian and *Drosophila* PcG complexes.

3-7.3 The PHD fingers of PCL and TRX do not mediate the function of E(Z) as a PcG or trxG member.

There are several lines of evidence which suggests that E(Z) is a member of both the PcG and trxG. Evidence for E(Z) being a member of the PcG includes the observation that male flies homozygous for a hypomorphic allele of *E(z)* display extra sex combs on the second and third legs, which is a characteristic phenotype of PcG members (Shearn *et al.,* 1978). Secondly, flies heterozygous for mutations in both *E(z)* and *ph* display a more severe extra sex combs phenotype when compared to single mutants alone (Cheng *et al.,* 1994). Thirdly, embryos homozygous for *E(z)* alleles show derepression of homeotic genes (Jones and Gelbart, 1990). The interaction described above between PCL and E(Z) is further evidence that E(Z) acts as a PcG gene. However, LaJeunesse and Shearn (1996) provided evidence that E(Z) is also able to act as a trithorax Group (trxG) member. Double heterozygous combinations of recessive loss of function *E(z)* and *ash1* (a member of the trxG) alleles gave a homeotic transformation phenotype similar to that observed in double heterozygous combinations of recessive loss of function alleles in *trx* and *ash1* (LaJeunesse and Shearn, 1996; Shearn, 1989). They also observed a loss in the expression of the homeotic genes SCR, ANTP and UBX in the thoracic imaginal discs of larvae hemizygous for a null allele of *E(z)*. This phenotype is also seen in the imaginal discs of larvae mutant for null alleles of *ash1* (LaJeunesse and Shearn, 1995). The ability to classify E(Z) as a member of the trxG or PcG appears to depend on the homeotic gene locus as well as on spatial and temporal cues. This dual role of E(Z) is most likely to be defined by protein-protein interactions at the site of action. The presence of PHD finger motifs in both TRX and PCL, along with the key role of the PHD fingers in mediating the interaction with E(Z), raised the possibility that the interaction of E(Z) with the PHD fingers of PCL or TRX determines its function as a PcG or trxG member respectively.

This hypothesis was tested using the yeast two hybrid assay. No interaction was observed between the three PHD fingers of TRX and E(Z), ruling out the PHD fingers as mediators of E(Z) function. TRX has a fourth imperfect PHD finger, which was not included in the LexA-TRX-PHDF construct, so it is possible that this imperfect finger is important in the interaction with E(Z). Although the PHD fingers of PCL and TRX do not appear to mediate the function of E(Z) as a member of the PcG or trxG, it is likely that the association of E(Z) with PCL mediates its role as a member of the PcG, at least in part. An interaction between E(Z) and an as yet unidentified trxG member may mediates its trxG-dependent role. It is also possible that the interaction between ESC and E(Z) mediates the function of E(Z) as a member of the PcG (Tie *et al.,* 1998; Jones *et al.,* 1998). ESC/E(Z) are thought to be part of a PcG initiating (PcGi) complex which is responsible for initiating PcG-mediated repression. ESC is only required transiently in early embryogenesis for the initiation of

PcG-mediated repression. E(Z) however, like most other members of the PcG, is required throughout development for PcG-mediated repression (Jones and Gelbart, 1990). It therefore appears as though E(Z) has a dual role as both a member of both the PcGi and PcGm complexes. Perhaps the interaction of E(Z) with ESC recruits E(Z) as a member of the PcGi and its interaction with PCL recruits E(Z) as a member of the PcGm complex. Identification of the composition of all PcG complexes *in vivo* is needed to clarify this issue.

Secondary to this, the absence of an interaction between the PHD fingers of TRX and E(Z) rules out the possibility that the interaction between the PHD fingers and E(Z) is an artefact of the yeast two hybrid assay and further strengthens the argument that the interaction between PCL and E(Z) is real.

The results presented in this chapter confirm the interaction between the cDOM of PCL and E(Z) identified by Robert (1997). They also provide conclusive evidence that the PHD finger motif of PCL is involved in mediating protein-protein interactions.

Chapter 4: *In vivo* characterisation of the functional domains of PCL.

4-1 Introduction

As discussed in Chapter 1, there is increasing evidence to suggest that the PcG acts to repress gene transcription through the formation of a multimeric protein complex. In order to gain a clear understanding of how the PcG proteins carry out their function, the composition and structure of the protein complexes they form need to be elucidated. To achieve this, interactions between members of the PcG need to be identified. With this in mind, *in vivo* domain analysis was performed on PCL to identify regions of PCL that are important in its function. A key region of PCL was identified using this assay. Yeast two hybrid and far western analysis was performed to identify PcG members which interact with this region. This chapter describes the outcomes of these experiments.

4-2 The tethering assay

To identify the regions of PCL important in its function, an *in vivo* tethering assay was initiated. This assay was first described by Jurg Müller (1995) and involves reporter and effector constructs. The reporter construct (BGUZ) contains a BXD enhancer, taken from the *Ubx* gene, and synthetic GAL4 binding sites upstream of a *lacZ* reporter gene which expresses in a head to tail pattern throughout embryogenesis (Figure 4.1 and Figure 4.2 A-C). The effector construct termed *hb-GAL-PCL*, contains a GAL4 DNA binding domain ($GAL4_{1-147}$), fused to *Pcl,* under the control of a *hb* promoter (Figure 4.1). The *hb* promoter drives expression of the GAL-PCL fusion in the anterior half of the blastoderm embryo, up until stage 7. When transformant flies containing the effector and reporter constructs are mated, analysis of *lacZ* expression in the progeny reveals whether the GAL-PCL fusion is able to repress expression of the reporter gene in the anterior region of the embryo, where the GAL-PCL fusion is being expressed (see Figure 4.1 for a schematic diagram). Staining of stage 10-16 embryos reveals whether the GAL-PCL fusion is able to initiate heritable repression of the reporter gene throughout development, because by this stage the GAL-PCL fusion protein is no longer being expressed.

Full length PCL was tested for its ability to repress *lacZ* expression in this assay. Figure 4.2 (D), shows that GAL-PCL is able to repress gene transcription anterior to PS6 in stage 10 embryos (T. McGrath, unpublished data). The repression mediated by GAL-PCL is maintained throughout embryogenesis (Figure 4.2, (E) and (F)). Repression is dependant on tethering of PCL to DNA via the GAL4 DNA binding domain, as *hb-PCL* does not repress reporter gene transcription (Figure 4.2, (G)-(I)). *hb-GAL* is unable to repress *lacZ* expression

(Müller, 1995) indicating that the repression conferred by the GAL-PCL fusion protein is due to the presence of tethered PCL.

4-3 The amino terminus of PCL is responsible for mediating repression of the *lacZ* reporter gene in the tethering assay

The ability of full length PCL to mediate heritable repression of the *lacZ* reporter gene provided an assay which could be used to identify functionally important regions of PCL. The PHD fingers of PCL have been shown to be important in mediating the interaction between PCL and E(Z) and, so far, are the only identified protein motif in PCL (Chapter 3). To determine whether the PHD fingers were sufficient for the function of GAL-PCL, two constructs were generated (by T. McGrath); *hb-GAL-PHD* and *hb-GAL-Amino+CarboxyΔPHD* (see Figure 4.3 for a schematic diagram). The *hb-GAL-PHD* construct encoded the PHD finger region of PCL (aa424-566) fused to the GAL4 DNA binding domain. The *hb-GAL-Amino+CarboxyΔPHD* construct encoded a fusion of the amino and carboxy termini minus the PHD finger region of PCL (aa1-423 and aa567-857). The junction of the *GAL* and *Pcl* fusions were sequenced to ensure that an in frame protein was generated. Transgenic lines were generated for both constructs and the chromosomal location determined. Flies homozygous for the *hb-GAL-PHD* or *hb-GAL-Amino+CarboxyΔPHD* constructs were mated to flies carrying the *BGUZ* reporter construct and anti-β-galactosidase staining was performed on their progeny (stage 10-16 embryos). Three independent transformant lines were tested for each construct. As shown in Figure 4.4, neither the *hb-GAL-PHD* ((A)-(C)) nor the *hb-GAL-Amino+CarboxyΔPHD* ((D)-(F)) were able to initiate or maintain heritable repression. This suggests that the PHD fingers are not necessary in the function of PCL in this assay and another region of PCL is responsible for mediating the heritable repression observed with full length PCL. However if another region of the protein was responsible why wasn't the *hb-GAL-Amino+CarboxyΔPHD* able to repress reporter gene expression? Perhaps the entire protein is important in the function of PCL or alternatively the PHD fingers are important in the folding of PCL and removing them leads to improper folding. To overcome this possible folding problem a further two constructs were generated; *hb-GAL-Amino* and *hb-GAL-Carboxy*. The *hb-GAL-Amino* construct encoded the amino terminus of PCL up to the start of the PHD fingers (aa1-423) (see Figure 4.3 for a schematic diagram) and the *hb-GAL-Carboxy* construct encoded the region of PCL C-terminal to the PHD fingers (aa567-857) (see Figure 4.3 for a schematic diagram). The junction of the *GAL* and *Pcl* fusion were sequenced to ensure that an in frame fusion protein was generated. Transgenic lines were generated for both constructs and the chromosomal localisation determined. Homozygous flies containing the *hb-GAL-Amino* or *hb-GAL-Carboxy* constructs were mated to flies carrying *BGUZ* and the progeny were stained using anti-β-galactosidase (stage 10-16). Three independent transformant lines were tested for each construct. Figure 4.5 ((A)-(C)) shows that the *hb-*

Figure 4.1: A schematic diagram of the tethering assay developed by Jurg Müller.

The tethering assay requires two constructs; the effector construct (A) and the reporter construct (B). The effector construct contains a GAL4 DNA binding domain (yellow) fused to the protein of interest, in this case PCL (green). Expression of the fusion protein is controlled by a minimal *hunchback* promoter, which drives expression in the anterior half of the blastoderm embryo. The reporter construct contains a *lacZ* reporter gene downstream of a BXD enhancer (blue) and GAL4 binding sites (red). The *BXD* enhancer is taken from the *Ultrabithorax* gene and drives expression in a head to tail pattern throughout embryogenesis. *Ultrabithorax* is known to be under the control of the PcG of proteins. When the GAL-PCL fusion protein is expressed in the anterior half of the embryo, it binds to the GAL4 binding sites. Antibodies to β-galactosidase are used to to detect *lacZ* expression (C). The absence of β-galactosidase protein in the anterior half of the embryo would indicate that the GAL-PCL fusion is able to repress reporter gene expression. The presence of staining in the anterior half of the embryo would indicate that the GAL-PCL fusion is unable to repress reporter gene expression. To determine whether the repression observed is heritable, stage 16 embryos are examined because by this stage no GAL-PCL fusion would be present.

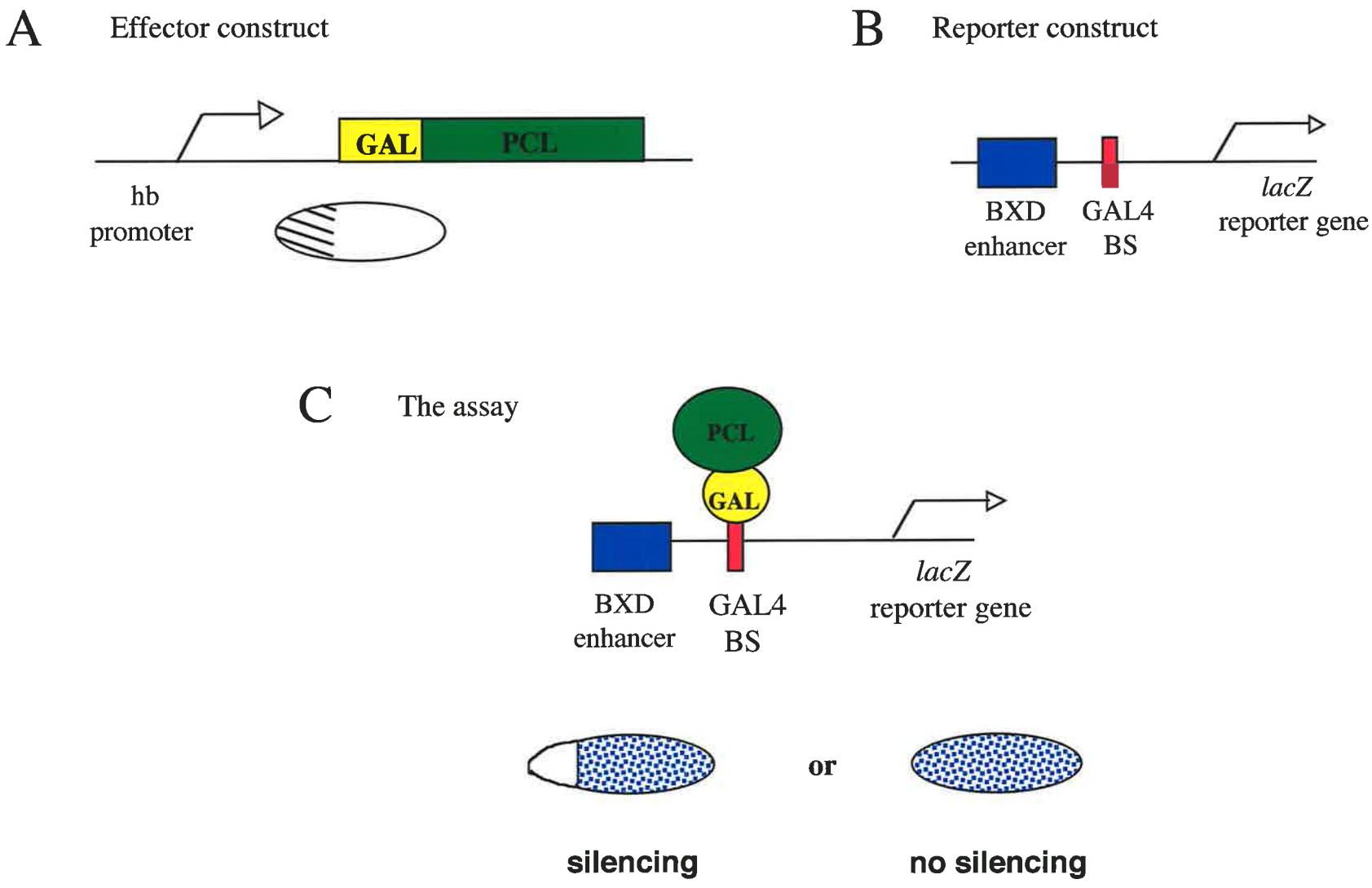

A
Effector construct
GAL
PCL
hb
promoter
B
Reporter construct
BXD
enhancer
GAL4
BS
lacZ
reporter gene
C
The assay
PCL
GAL
BXD
enhancer
GAL4
BS
lacZ
reporter gene
or
silencing
no silencing

Figure 4.2: GAL-PCL is able to initiate heritable repression of the *lacZ* reporter gene.

To determine whether PCL is capable of repressing a *lacZ* reporter gene *in vivo* when tethered to DNA via a GAL4 DNA binding domain, flies containing the BGUZ reporter construct were crossed to flies containing either the BGUZ reporter ((A)-(C)), *hb-GAL-PCL* ((D)-(F)) or *hb -PCL* ((G)-(I)). The progeny of these crosses were stained for *lacZ* expression at three time points; stages 10, 12-14 and 16. The later staged embryos were used to determine whether repression, once initiated, was heritable. (A)-(C) show the head to tail expression of the BGUZ reporter gene in embryos at stage 10 (A), 13 (B) and 16 (C). *hb*-GAL-PCL is able to repress *lacZ* reporter gene expression in the anterior half of stage 10 embryos (compare (D) and (A)) and this repression is maintained throughout embryogenesis (compare (E) with (B) and (F) with (C)). *hb*-PCL is unable to repress reporter gene expression ((G)-(I)) at any stage of embryogenesis, showing that repression of the *lacZ* reporter gene ((D)-(F)) is dependent on tethering to the DNA via the GAL4 DNA binding domain. The GAL4 DNA BD alone is unable to repress gene expression (Müller, 1995). Arrowhead indicates the anterior boundary of parasegment 6.

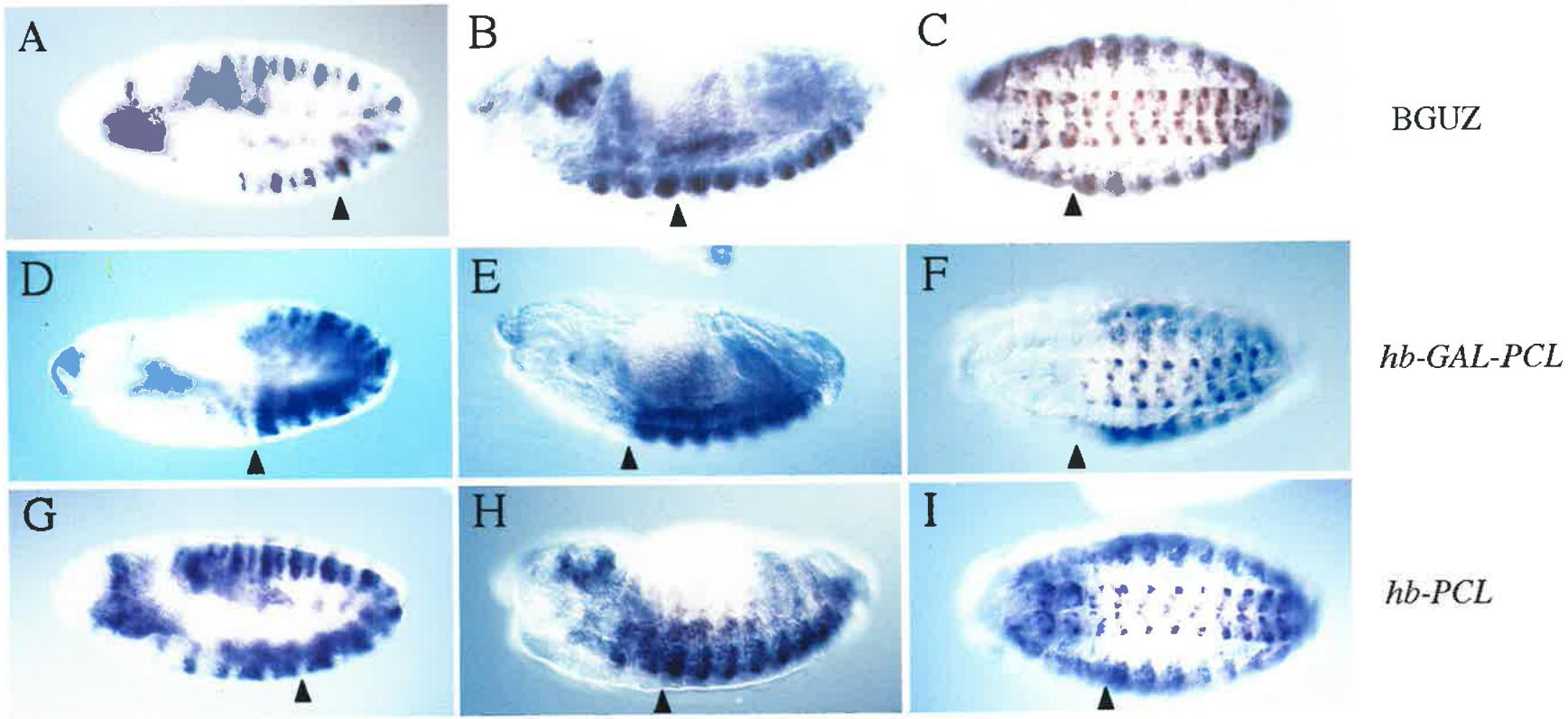

A
B
C
BGUZ
D
E
F
hb-GAL-PCL
G
H
I
hb-PCL

Figure 4.3: A schematic diagram of the effector constructs used to establish the region(s) of PCL that are responsible for initiating heritable repression of the *lacZ* reporter gene.

In order to determine which region of PCL was responsible for initiating heritable repression, four effector constructs were generated each containing different fragments of the PCL open reading frame. A schematic representation of the PCL open reading frame is shown with the PHD fingers represented by green boxes. *hb-GAL-PHD* encodes the region spanning both PHD fingers (aa424-566). *hb-GAL-Amino+CarboxyΔPHD* encodes aa1-423 fused to aa567-857. *hb-GAL-Amino* encodes aa1-423 of PCL. *hb-GAL-Carboxy* encodes aa567-857. The numbers represent amino acid position of each fragment.

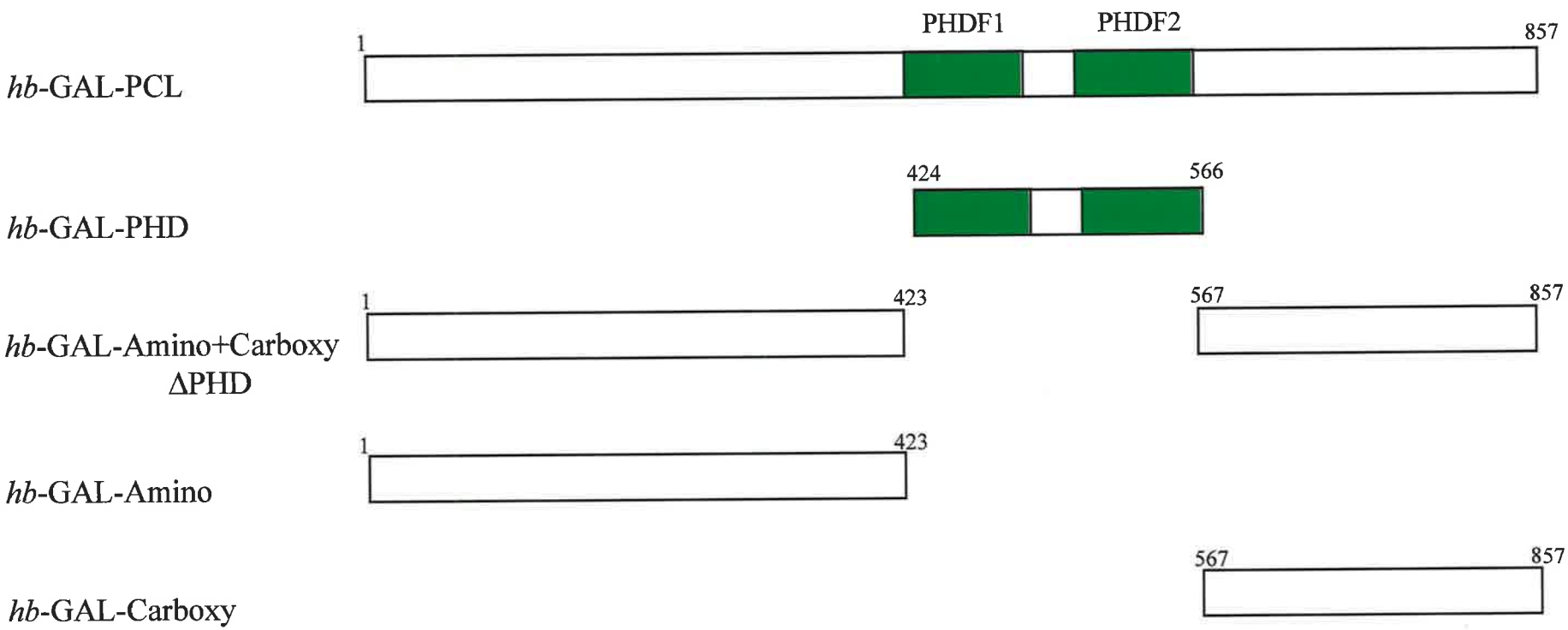

PHDF1
PHDF2
1
857
hb-GAL-PCL
424
566
hb-GAL-PHD
1
423
567
857
hb-GAL-Amino+Carboxy
ΔPHD
1
423
hb-GAL-Amino
567
857
hb-GAL-Carboxy

Figure 4.4: The PHD finger region of PCL is not responsible for the heritable repression conferred by GAL-PCL.
The PHD fingers of PCL are the only identified protein motif in PCL. They are known to mediate the interaction between PCL and E(Z) and may be important in initiating the heritable repression conferred by the GAL-PCL fusion protein. To determine whether this region of PCL was important in initiating heritable repression, two effector constructs were generated; a *hb-GAL-PHD* and a *hb-GAL-Amino+CarboxyΔPHD* construct. Flies homozygous for either *hb-GAL-PHD* or *hb-GAL-Amino+CarboxyΔPHD* were mated to flies carrying the *BGUZ* reporter construct and the *lacZ* expression pattern of their progeny determined at three time points; stage 10 ((A) and (D)), stage 12-14 ((B) and (E)) and stage 16 ((C) and (F)). *hb-GAL-PHD* was unable to repress reporter gene expression at all three time points ((A)-(C)) indicating that the PHD finger region of PCL is not responsible for mediating the repression of the GAL-PCL fragment. *hb-GAL-Amino+CarboxyΔPHD* was also unable to repress the reporter gene at all three time points ((D)-(F)). The anterior boundary of parasegment 6 is shown by an arrowhead.

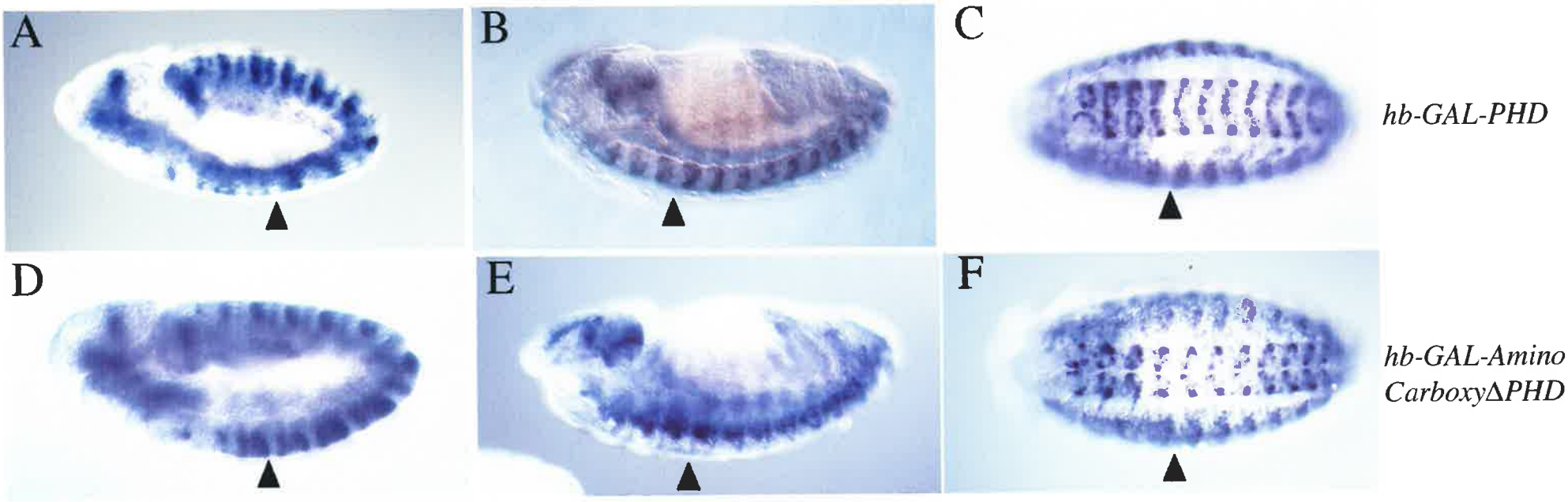
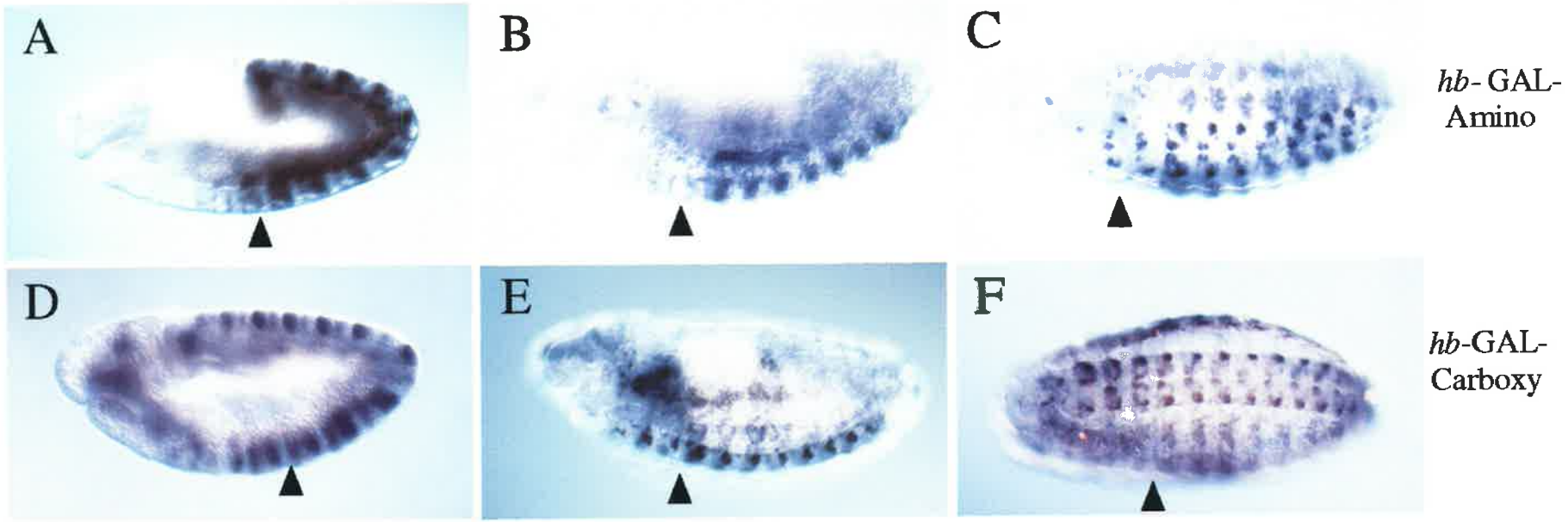

A
B
C
D
E
F
hb-GAL-PHD
hb-GAL-Amino
CarboxyΔPHD

Figure 4.5: The amino terminus of PCL is responsible for mediating heritable repression of the *lacZ* reporter gene.
Neither the *hb-GAL-PHD* nor the *hb-GAL-Amino+CarboxyΔPHD* construct were able to initiate heritable repression of the *lacZ* reporter gene, suggesting that the PHD fingers cannot repress the *lacZ* reporter gene in this assay. To determine which region of PCL is important, the amino terminus of PCL and the carboxy terminus of PCL were cloned into the *hb-GAL* vector to generate *hb-GAL-Amino* and *hb-GAL-Carboxy*. Transgenic lines homozygous for the *hb-GAL-Amino* or *hb-GAL-Carboxy* constructs were mated to flies containing the *BGUZ* reporter construct and the progeny examined for *lacZ* expression using an anti-β-galactosidase antibody at three different stages of embryogenesis; 10 ((A) and (D)), 12-14 ((B) and (E)) and 17 ((C) and (F)). (A)-(C) are embryos which contain the *hb-GAL-Amino* and *BGUZ* constructs, stained for β-galactosidase protein. Repression of the anterior expression of *lacZ* is observed at all stages of embryogenesis indicating that the amino terminus of PCL is able to establish heritable repression of the *lacZ* reporter gene. However the repression consistently appeared weaker than that observed for the *hb-GAL-PCL* construct as some β-galactosidase staining was present in parasegment 5. (D)-(F) are embryos containing the *hb-GAL-Carboxy* and *BGUZ* constructs which have been stained for *lacZ* expression. No repression of the *lacZ* expression is seen in the anterior half of the embryos at any stage of development examined. The arrowhead indicates the anterior boundary of parasegment 6.


A
B
C
hb- GAL-
Amino
D
E
F
hb-GAL-
Carboxy

GAL-Amino construct was able silence the *lacZ* reporter gene and this silencing was maintained in later embryos, indicating that the amino terminus of PCL is able to initiate heritable silencing *in vivo*. The silencing conferred by the GAL-Amino construct consistently appeared weaker in late embryos than the silencing conferred by GAL-PCL (compare Figure 4.2 (F) with Figure 4.5 (C)). The GAL-Amino fusion protein only maintained repression in parasegments 1-4, compared to 1-5 for GAL-PCL. The *hb-GAL-Carboxy* construct was unable to silence reporter gene expression (Figure 4.5, (D)-(F)), indicating that the carboxy terminus of PCL is not important in the function of PCL in this assay. Alternatively the GAL-Carboxy fusion protein is unable to fold correctly or is not being expressed.

4-4 The heritable silencing conferred by GAL-PCL and GAL-Amino is PcG dependant

The ability of the GAL-PCL and GAL-Amino fusion proteins to initiate repression that was maintained suggests that these proteins are able to recruit the endogenous PcG complex through protein-protein interactions. This recruitment would establish heritable repression by a mechanism analogous to that of gap protein recruitment of the PcG to target genes early in embryogenesis. Jurg Müller demonstrated that the heritable silencing initiated by a GAL-PC fusion protein was PcG dependent (Müller, 1995). If the GAL-PCL and GAL-Amino proteins are interacting with endogenous PcG members to initiate heritable repression, then heritable repression should also be dependent on endogenous PcG proteins.

To determine whether the silencing conferred by the GAL-PCL and GAL-Amino fusion proteins was PcG dependant, transgenic lines containing the *hb-GAL-PCL, hb-GAL-Amino* and BGUZ constructs were crossed into a Pcl^{E90} and $E(z)^{61}$ mutant background (Figure 4.6). Pcl^{E90} is a null allele of *Pcl* (Jurgens, 1985) and $E(z)^{61}$ is a temperature sensitive allele of *E(z)* (Jones and Gelbart, 1990). The embryo collects for $E(z)^{61}$ were performed at 29°C where both maternal and zygotic E(Z) is non-functional. The progeny of flies heterozygous for mutations in *Pcl* or homozygous for *E(z)* were collected and analysed for their *lacZ* expression using an anti-β-galactosidase antibody (Figure 4.6). All embryos stained contained a copy of the effector construct and therefore any derepression of the *lacZ* reporter gene expression will be due to a requirement for endogenous PCL and E(Z). 50% of the stained embryos were homozygous mutant for Pcl^{E90} and 100% of the stained embryos were mutant for $E(z)^{61}$. Embryos were examined at stage 10 and stage 16. No derepression was observed in either the *hb-GAL-PCL* (Figure 4.7 compare (A), (G) and (I)) or the *hb-GAL-Amino* (Figure 4.7 compare (A), (C) and (E)) embryos mutant for Pcl^{E90} at stage 10. *Pcl* is a maternally deposited transcript (Lonie *et al.*, 1994) and it is therefore unable to be determined if endogenous PCL is required at this stage of the tethering assay. At stage 16, derepression of the *lacZ* reporter gene was observed in Pcl^{E90} mutant embryos for both the *hb-GAL-PCL*

(Figure 4.7 compare (B), (H) and (J)) and *hb-GAL-Amino* (Figure 4.7 compare (B), (D) and (F)) constructs at the expected frequency of 50%. Endogenous PCL is therefore required for the maintenance of heritable repression conferred by GAL-PCL and GAL-Amino fusion proteins. Stage 10 embryos mutant for *E(z)* showed derepression of the anterior border of *lacZ* staining in both the *hb-GAL-PCL* (Figure 4.8 compare (A), (G) and (I)) and the *hb-GAL-Amino* embryos (Figure 4.8 compare (A), (C) and (E)), indicating that E(Z) is required for initiation of repression conferred by the GAL-PCL and GAL-Amino constructs. Stage 16 *E(z)* embryos carrying the *hb-GAL-PCL* (Figure 4.8 compare (B), (H) and (J)) or *hb-GAL-Amino* constructs (Figure 4.8 compare (B), (D) and (F)) were unable to repress reporter gene expression in the anterior region of the embryo. It is not clear whether maintenance of repression is affected in *E(z)* mutants, because repression was not initiated.

4-5 The amino terminus of PCL interacts with full length PCL

Given that repression of the *lacZ* reporter gene by the GAL-PCL and GAL-Amino fusion proteins was PcG dependent, it was reasonable to assume that PCL, specifically the amino terminus of PCL, was interacting with PcG protein(s) to mediate initiation of heritable repression. To determine which PcG proteins are interacting with the amino terminus of PCL two approaches were used; far western analysis and a yeast two hybrid assay.

Far western analysis was used to determine whether the amino terminus of PCL was capable of interacting with the full length PCL. The amino terminus of PCL (aa1-423) was cloned into the pGex3X vector to generate a bacterially expressed GST-PCL-Amino fusion protein. GST-PCL had previously been generated by S. Robert (1997). Far western analysis was then performed to determine whether full length PCL is able to interact with the GST-PCL and GST-PCL-Amino fusion proteins. Equal amounts of GST-fusion protein and GST alone were loaded onto an 11% SDS-polyacrylamide gel, electrophoresed, western blotted and probed with radiolabelled PCL (^{35}S PCL) after denaturation and renaturation of the blot (see 2-3.11). As shown in Figure 4.9 ^{35}S PCL interacted with both GST-PCL (Lane 1) and GST-PCL-Amino (Lane 2). ^{35}S PCL failed to interact with GST alone (Figure 4.9 Lane 4), indicating that the interaction with the GST-PCL and GST-PCL-Amino fusion proteins was occurring through the PCL portion of the fusion protein.

To refine the region required for the interaction between the amino terminus of PCL and full length PCL, the first 200aa of PCL was cloned into pGex3X to generate a GST-PCL$_{1\text{-}200}$ fusion protein and far western analysis was performed to determine whether this region was responsible for mediating the interaction with ^{35}S PCL. As shown in Figure 4.9 (Lane 3), GST-PCL$_{1\text{-}200}$ is able to bind with ^{35}S PCL, indicating that the first 200aa of PCL is sufficient for the homotypic interactions of PCL. GAL-PCL-Amino may therefore be able to repress *lacZ* reporter gene expression by recruiting full length PCL.

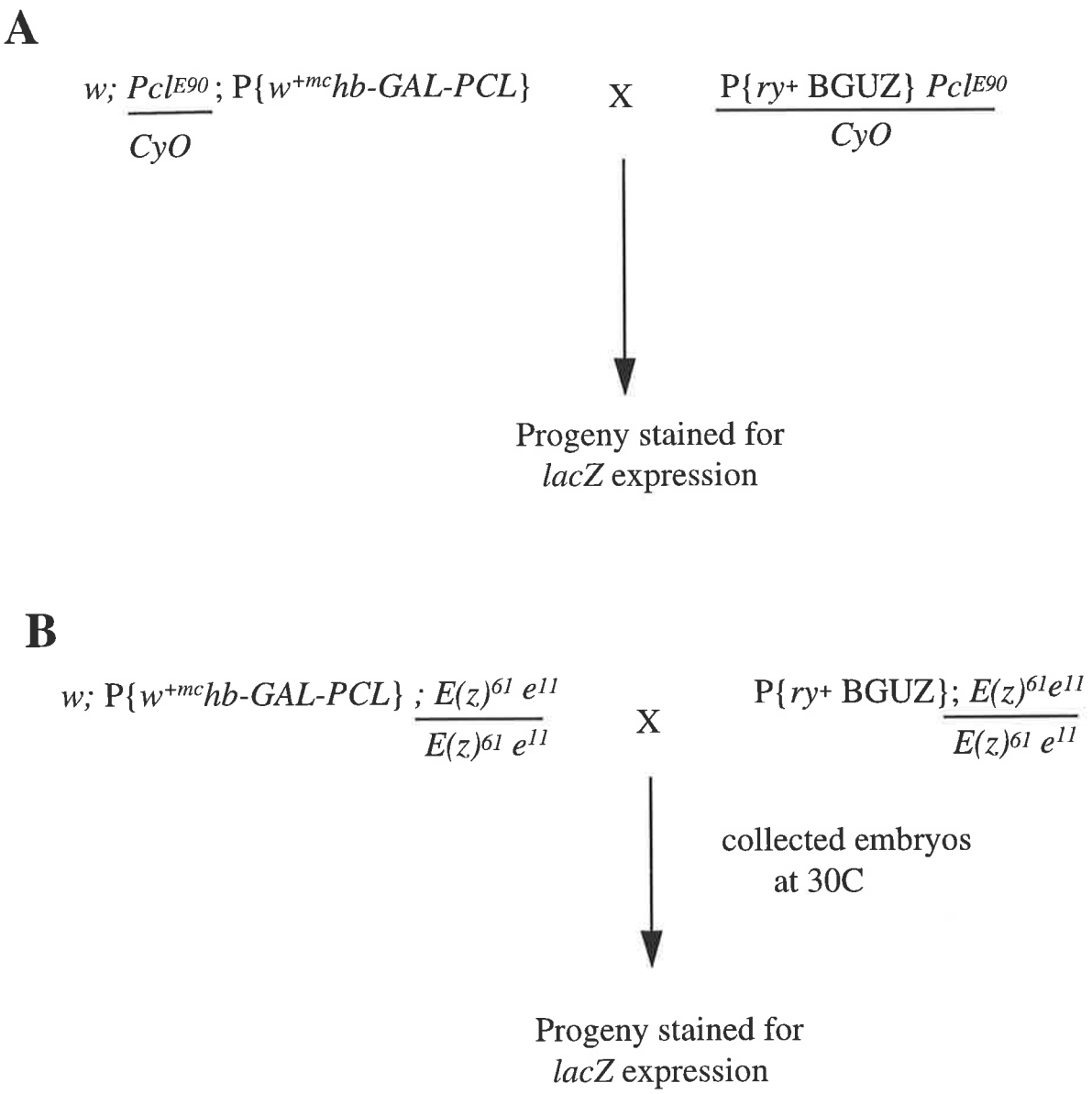


Figure 4.6: A summary of the fly crosses used to test PcG dependency of silencing.
A schematic of the fly strains used and the fly crosses performed to generate embryos carrying either the *hb-GAL-PCL* or the *hb-GAL-Amino* effector construct with the BGUZ reporter construct, in a Pcl^{E90} (A) or $E(z)^{61}$ (B) mutant background. Staged embryo collects were performed at 25°C for Pcl^{E90} and at 30°C for $E(z)^{61}$. 50% of the stained embryos will be mutant for Pcl^{E90} and 100% of the embryos will be mutant for $E(z)^{61}$.

Figure 4.7: Endogenous PCL is required for the heritable repression conferred by the GAL-PCL and GAL-Amino fusion constructs.

To determine whether the heritable repression of the *lacZ* reporter gene initiated by the GAL-Amino and GAL-PCL fusion constructs was PcG dependent, the tethering assay was performed in a *Pcl* mutant background. The genotypes of the embryos are as follows:
(A) Stage 10 *BGUZ*
(B) Stage 16 *BGUZ*
(C) Stage 10 *BGUZ; hb-GAL-Amino*
(D) Stage 16 *BGUZ; hb-GAL-Amino*
(E) Stage 10 *BGUZ; hb-GAL-Amino;* not known whether this embryo is homozygous or heterozygous for the Pcl^{E90} mutant allele, but the embryo is representative of 100% of the embryos observed.
(F) Stage 16 *BGUZ; hb-GAL-Amino;* Pcl^{E90}/ Pcl^{E90}; this embryo is representative of 50% of the embryos observed and is thought to be a homozygous Pcl^{E90}mutant.
(G) Stage 10 *BGUZ; hb-GAL-PCL*
(H) Stage 16 *BGUZ; hb-GAL-PCL*
(I) Stage 10 *BGUZ; hb-GAL-PCL*; not known whether this embryos is homozygous or heterozygous for the Pcl^{E90} mutant allele, but the embryos is representative of 100% of the embryos observed.
(J) Stage 16 *BGUZ; hb-GAL-PCL;* Pcl^{E90}/ Pcl^{E90}; this embryo is representative of 50% of the embryos observed and is thought to be a homozygous Pcl^{E90}mutant.

When (E) is compared with (A) and (C) it is apparent that no derepression of the *lacZ* reporter gene can occur in Pcl^{E90}/ Pcl^{E90} embryos at this stage of development. (F) is representative of 50% of the embryos and is presumably a homozygous Pcl^{E90} mutant. When compared to (D), (F) shows derepression of the reporter gene into the anterior half of the embryo, demonstrating that maintenance of the repression conferred by the GAL-Amino fusion protein requires endogenous PCL. (I) is representative of 100% of the embryos observed showing no depression of *lacZ* expression in the anterior region of the embryo when compared with (A) and (G). (J) is representative of 50% of the embryos and is presumably a homozygous Pcl^{E90} mutant. When compared to (H), (J) shows depression of reporter gene expression in the anterior half of the embryo demonstrating that the heritable repression conferred by the GAL-PCL fusion protein is dependent on endogenous PCL.

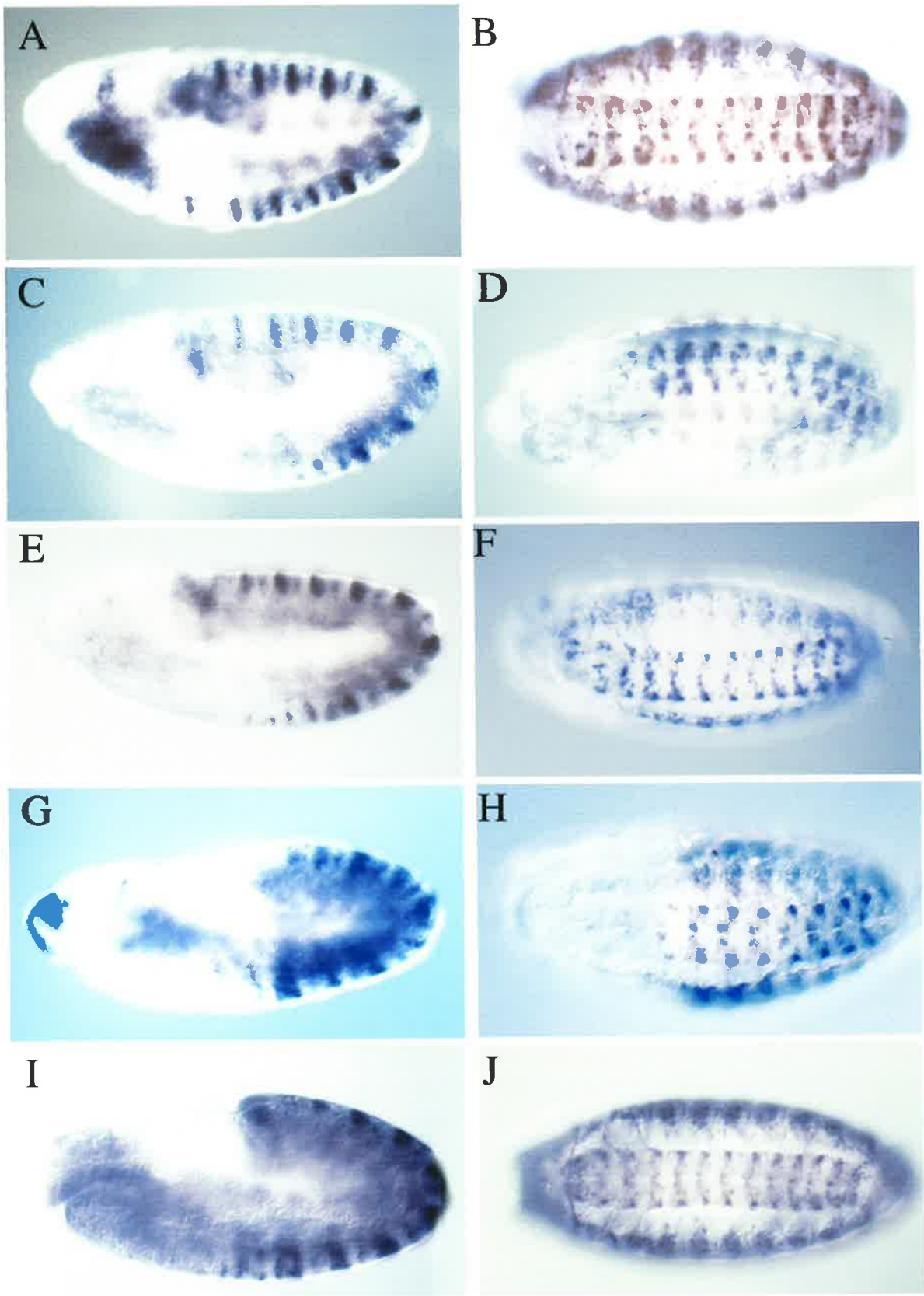

A
B
C
D
E
F
G
H
I
J

Figure 4.8: Endogenous E(Z) is required for the heritable repression conferred by the GAL-PCL and GAL-Amino fusion proteins.

To determine whether the heritable repression of the *lacZ* reporter gene, initiated by the GAL-Amino and GAL-PCL fusion proteins is PcG dependent, the tethering assay was performed in an *E(z)* mutant background. The genotypes of the embryos are as follows:

(A) Stage 10 *BGUZ*

(B) Stage 16 *BGUZ*

(C) Stage 10 *BGUZ / hb-GAL-Amino*

(D) Stage 16 *BGUZ / hb-GAL-Amino*

(E) Stage 10 *BGUZ / hb-GAL-Amino; $E(z)^{61}$/ $E(z)^{61}$*

(F) Stage 16 *BGUZ / hb-GAL-Amino; $E(z)^{61}$/ $E(z)^{61}$*

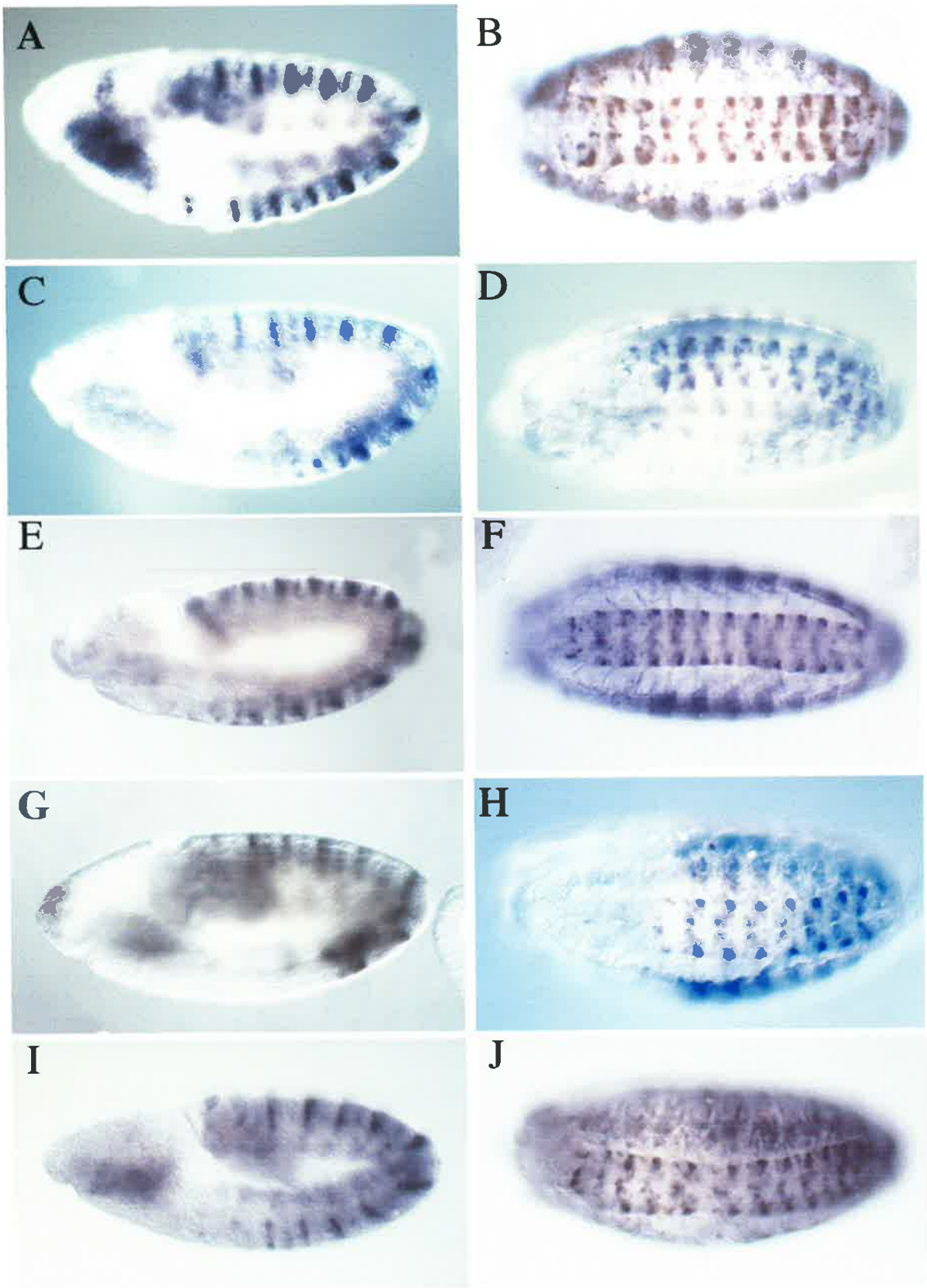
(G) Stage 10 *BGUZ / hb-GAL-PCL*

(H) Stage 16 *BGUZ / hb-GAL-PCL*

(I) Stage 10 *BGUZ / hb-GAL-PCL; $E(z)^{61}$/ $E(z)^{61}$*

(J) Stage 16 *BGUZ / hb-GAL-PCL; $E(z)^{61}$/ $E(z)^{61}$*

When (E) is compared with (A) and (C) it is apparent that derepression of the *lacZ* reporter gene has occurred. This indicates that repression by the GAL-Amino fusion protein is dependent on endogenous E(Z) at stage 10. Derepression of the reporter gene has also occured in (F) (compare to (D)) at stage 16, indicating that maintenance of the repression initiated by GAL-Amino is dependent on endogenous E(Z). Repression conferred by the GAL-PCL fusion protein is also dependent on endogenous E(Z) at both stage 10 (compare (A), (G) and (I)) and stage 16 (compare (B), (H) and (J)).


A
B
C
D
E
F
G
H
I
J

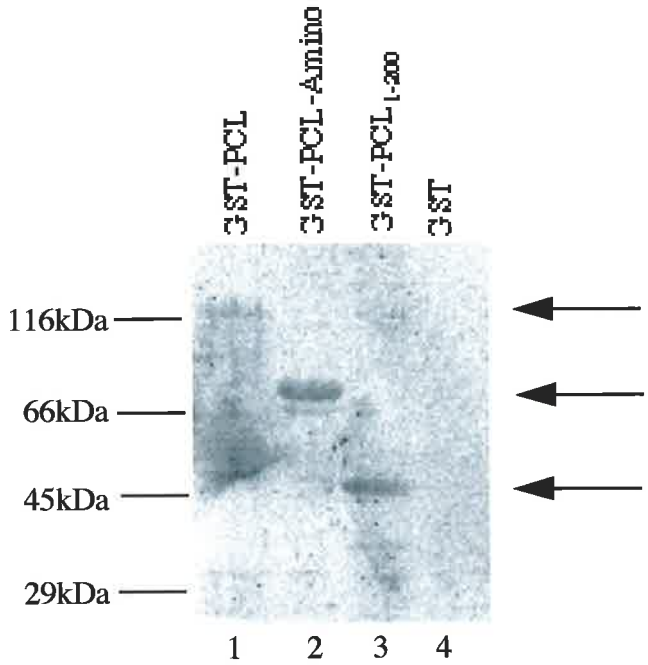


Figure 4.9: The amino terminus of PCL interacts with full length PCL.

To determine whether the amino terminus of PCL mediates homotypic interactions, far western analysis was performed. The amino terminus of PCL (aa1-424) and the PCL_{1-200} fragment (aa1-200) were cloned into pGEX3 and the proteins induced generating GST-PCL-Amino and GST-PCL_{1-200} respectively. A western blot containing GST-PCL (generated by S. Robert), GST-PCL-Amino, GST-PCL_{1-200} and GST alone was probed with ^{35}S-labelled PCL. GST-PCL (Lane 1), GST-PCL-Amino (Lane 2) and GST-PCL_{1-200} (Lane 3) were able to interact with ^{35}S PCL . ^{35}S PCL interacted specifically with the PCL portion of the GST-fusion proteins as no interaction was observed between ^{35}S PCL and GST alone (Lane 4).

The predicted size of each of the proteins is

GST-PCL- 120kDa

GST-PCL-Amino- 72kDa

GST-PCL_{1-200} - 48kDa

GST- 29kDa

4-6 Polycomb interacts with the amino terminus of PCL in the two hybrid assay

The self-association of the amino terminus of PCL may be sufficient to initiate the repression observed by GAL-PCL and GAL-Amino. However, given that no other region of PCL was sufficient to initiate heritable repression it was thought that the amino terminus of PCL would interact with other PcG members to enable heritable repression. In order to identify any potential interactions between the amino terminus of PCL and PcG proteins, the amino terminus (aa1-423) of PCL was cloned into pEG202 to generate LexA-PCL-Amino. Although this construct was generated using a high fidelity DNA polymerase (*Pfu*), subsequent sequence analysis of the LexA-PCL-Amino construct identified several mutations, many of which produced truncated proteins approximately 130aa into the amino terminus (data not shown). This construct was regenerated three times using independent PCR amplification reactions. Each time, several errors had been introduced into the PCL ORF. It is therefore likely that the amino terminus of PCL was being expressed in *E. coli* and was lethal to the bacteria in the unmutated form, so that only bacteria containing mutated constructs were able to grow.

The LexA-PCL-Amino fusion is under the control of a *S. cerevisiae* constitutive pADH promoter which must not be completely inactive in *E. coli,* leading to leaky expression of the LexA-PCL-Amino fusion protein. Similar effects have been observed by others (D. Kortschak pers. comm.). It was known however that GST-PCL-Amino was not lethal to bacteria. The expression of these proteins is under the control of an IPTG inducible *tac* promoter. Given the ability of GST-PCL-Amino to be expressed under an inducible promoter, it was thought that cloning the amino terminus into the galactose inducible LexA yeast two hybrid vector, pGilda could overcome the lethality. pGilda-PCL-Amino was generated and, when sequenced, was in frame and contained no mutations (data not shown). LexA-PCL-Amino was then transformed into EGY48 and western analysis performed using the LexA antiserum to ensure that the fusion protein was being expressed (data not shown). LexA-PCL-Amino was tested for the ability to self-activate a *LEU2* reporter gene and was unable to do so (data not shown).

LexA-PCL-Amino was tested for an interaction with the following set of constructs, which were generously provided by Michael Kyba and Hugh Brock (UBC); AD-AsxA, AD-AsxQ, AD-AsxZn, AD-E(Pc)Ybox, AD-E(z), AD-Pc, AD-PcΔ3', AD-PcΔchromo, AD-ph, AD-ph-D, AD-phΔN, AD-phΔS AD-Psc, AD-PscΔB, AD-SCM, AD-Su(Z)2 and AD-Su(Z)2ΔB. To complete the set of AD-PcG proteins, AD-esc and AD-PHO were generated using high fidelity PCR and cloning into pJG4-5 to generate the AD-fusion proteins (see 2-2.12 and 2-2.13). The vector pJG4-5 contains a HA epitope tag which enables confirmation of the expression of constructs. Both AD-PHO and AD-esc were transformed into EGY48 and western analysis was performed using the anti-HA antisera, confirming that the fusion proteins were expressed (data not shown).

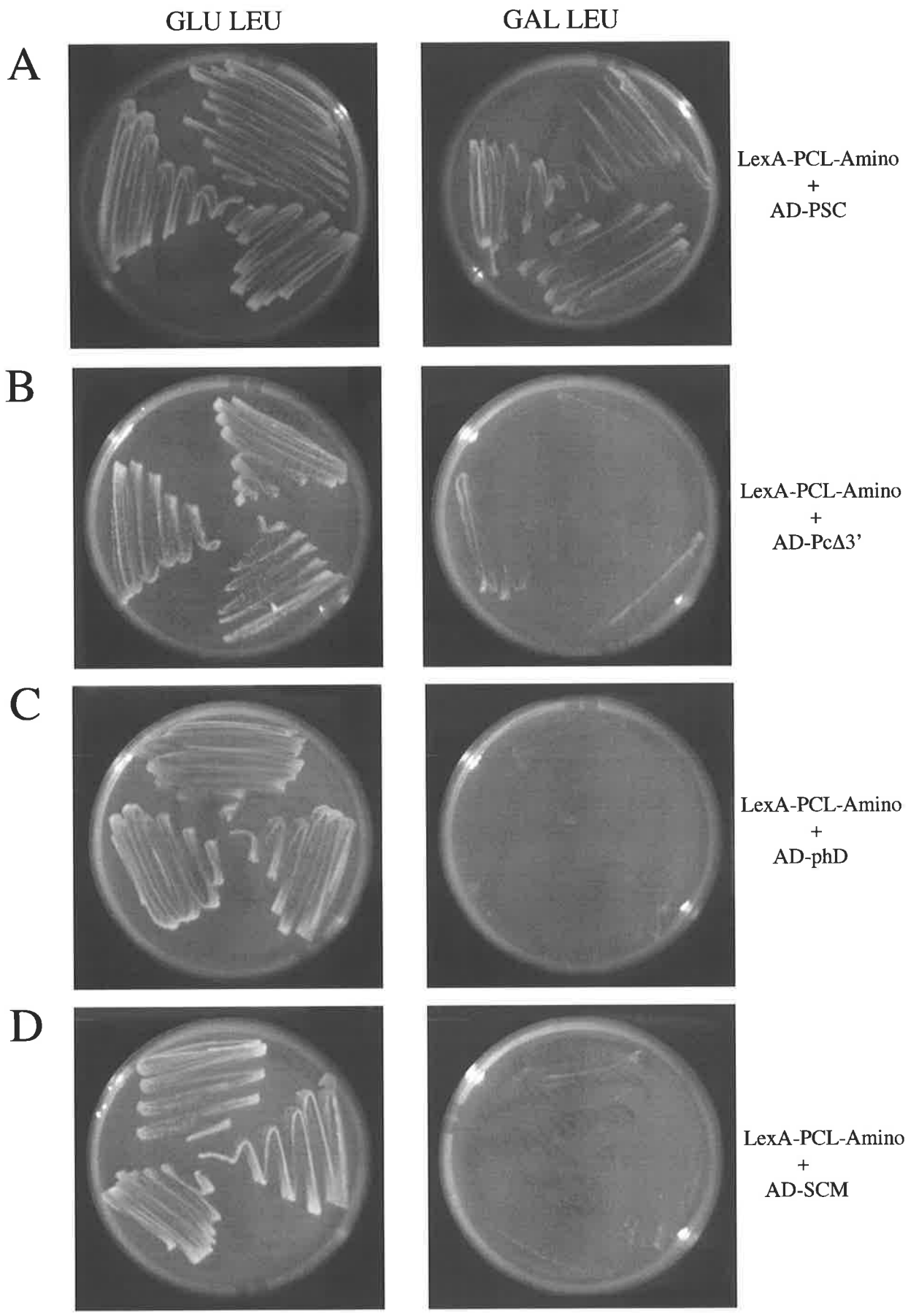
When LexA-PCL-Amino was tested for an interaction with the AD-PcG constructs only AD-ph-D, AD-PcΔ3' and AD-Scm displayed an interaction. AD-ph-D encodes aa1297-1589 which is inclusive of the SPM domain. AD-PcΔ3' encodes a truncated form of PC, lacking the shadow domain at the C-terminus. AD-Scm encodes the entire SCM open reading frame.

The interactions observed between LexA-PCL-Amino and the three AD constructs were not straight forward. Instead of the usual growth on GAL plates versus no growth on GLU plates, lethality was observed on GAL LEU plates and no growth was observed on GAL or GLU plates. Normal growth was observed on GLU LEU. Figure 4.10 is a representation of the lethality observed for LexA-PCL-Amino and AD-Scm, AD-ph-D and AD-PcΔ3' compared to the normal growth of the AD-Psc combination. The lethality observed on GAL LEU plates must result from the induction of both the LexA-PCL-Amino construct and the AD-PcG construct, as no growth impairment was observed on GLU LEU plates. One explanation for the lethality is that the expression of the amino terminus of PCL with either ph-D, Scm or PcΔ3' produces a highly repressive complex in yeast which represses transcription of endogenous yeast genes required for viability.

To determine whether the interaction between the amino terminus and ph-D, Scm and PcΔ3' was the result of the formation of a repressive complex *in vivo* or the result of the lethality of the individual proteins, several control experiments were performed. Firstly, each construct was tested individually for its ability to grow on glucose containing media versus galactose containing media. AD-ph-D was lethal in combination with LexA alone (data not shown) so that an interaction between LexA-PCL-Amino and AD-ph-D could not be tested. AD-Scm was not lethal in combination with LexA alone but was lethal in combination with several control LexA-constructs, including LexA-DBL, a fusion of LexA with a fragment of the Pebble protein involved in cytokinesis, LexA-cyclinE, a fusion of LexA with a G1/S-phase regulator and LexA-DRI-N, LexA-DRI-ARID and LexA-DRI-C, all fusions with LexA and fragments of the Dead ringer protein involved in early embryonic patterning (data not shown). This random lethality in combination with proteins that SCM would not be expected to interact with, indicates that the lethal interaction observed between LexA-PCL-Amino and AD-Scm is likely to be due to non-specific toxicity. An interaction between LexA-PCL-Amino and AD-Scm, therefore, cannot be tested by this approach. AD-PcΔ3' was not lethal in combination with LexA alone nor any of the other LexA constructs that were tested (data not shown). This suggests that the lethality observed between LexA-PCL-Amino and AD-PcΔ3' is the result of a direct association between these two proteins. See Table 4.1 for a summary of the interactions.

Figure 4.10: The combined expression of the LexA-PCL-Amino fusion protein together with the AD-PcΔ3', AD-ph-D or AD-SCM fusion proteins cause yeast lethality.

To determine whether the amino terminus of PCL is able to interact with PcG proteins, a LexA-PCL-Amino construct was generated and tested for an interaction with an array of AD-PcG constructs. No GAL dependent activation of the LEU2 reporter gene was observed for any AD-PcG construct. For three AD-PcG constructs, AD-PcΔ3' (B), AD-phD (C) and AD-SCM (D), a synthetic lethality was observed on GAL LEU medium when compared to all other constructs. No lethality was observed on GLU LEU medium as LexA-PCL-Amino, AD-PcΔ3', AD-phD or AD-SCM are under galactose inducible control. LexA-PCL-Amino and AD-PSC construct combination (A) serves as a growth comparison as it is representative of the growth observed for other constructs on GLU LEU and GAL LEU.


GLU LEU
GAL LEU
A
LexA-PCL-Amino
+
AD-PSC
B
LexA-PCL-Amino
+
AD-PcΔ3'
C
LexA-PCL-Amino
+
AD-phD
D
LexA-PCL-Amino
+
AD-SCM

	AD-PcΔ3'	AD-ph-D	AD-Scm	other AD constructs	AD alone
LexA alone	not lethal	lethal	not lethal	not lethal	not lethal
LexA-PCL-Amino	lethal	lethal	lethal	not lethal	not lethal
LexA-cycE	not lethal	lethal	lethal	not tested	not lethal
LexA-DBL	not lethal	lethal	lethal	not tested	not lethal
LexA-DRI-N	not lethal	lethal	lethal	not lethal	not lethal
LexA-DRI-ARID	not lethal	lethal	lethal	not lethal	not lethal
LexA-DRI-C	not lethal	lethal	lethal	not lethal	not lethal

Table 4.1: A summary of the lethal/non-lethal interaction between AD-PcΔ3', AD-ph-D, AD-Scm and control LexA constructs when plated on GAL LEU media.

To overcome this potential problem with repression, the amino terminus of PCL was broken up into two overlapping fragments to generate a LexA-PCL_{1-200} (aa1-200) and a LexA-$PCL_{132-423}$ (aa132-423) construct (see Figure 4.11 for a schematic diagram). LexA-PCL_{1-200} was generated in pEG202, $PCL_{132-423}$ however, was unable to be cloned into pEG202. The problem with cloning $PCL_{132-423}$ into pEG202 is similar to the inability to clone the amino terminus into pEG202 and is therefore probably due to leaky expression of the fusion protein in *E. coli* resulting in lethality. $PCL_{132-423}$ was therefore cloned into pGilda, the galactose inducible yeast two hybrid vector. Western analysis was performed on both proteins, using anti-LexA antisera to confirm their expression (Figure 4.11). When tested for an interaction with AD-PcΔ3' no interaction was observed between AD-PcΔ3' and LexA-PCL_{1-200} or LexA-$PCL_{132-423}$ and no lethality was observed on galactose-containing media (Figure 4.11). This indicates that either the whole amino terminus is required for the interaction with PC or the lethal interaction between PC and the amino terminus of PCL is artefactual.

4-7 Discussion

In order to identify the functional regions of PCL, an *in vivo* tethering assay was performed. Full length PCL had previously been shown to initiate heritable repression of a reporter gene when tethered to DNA via a GAL4 DNA binding domain (T. McGrath pers. comm.). This assay was used to identify the functional domains of PCL.

4-7.1 The amino terminus of PCL is sufficient to establish heritable repression

The amino terminus of PCL was the only region sufficient to initiate heritable repression. The repression conferred by GAL-Amino was significantly weaker than that of GAL-PCL.

It appeared to only extend from parasegment 1-4. Perhaps the high concentration of GAL-Amino at the anterior end of the embryo is sufficient to initiate heritable repression through the recruitment of PcG complexes, but as the concentration of GAL-Amino decreases towards the posterior end of the embryo, the efficiency of recruiting the endogenous PcG and initiating repression reduces.

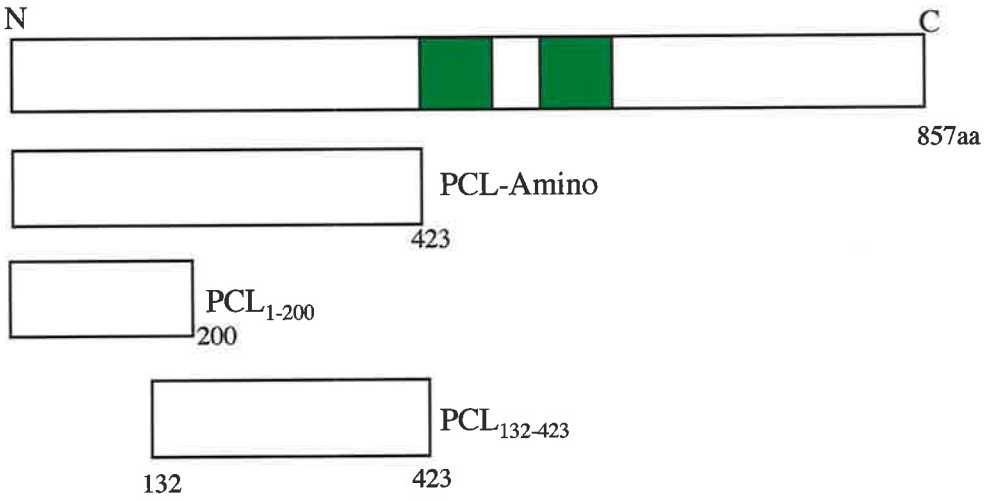
Heritable repression by GAL-Amino and GAL-PCL was dependent on endogenous PCL, suggesting that the amino terminus of PCL was directly interacting with PcG members to initiate heritable repression. Far western analysis demonstrated that the amino terminus of PCL, and specifically the first 200aa, is able to interact with full length PCL. It is not yet known whether regions outside of the amino terminus of PCL are involved in this homotypic interaction, however they are unlikely to be, given that no other region of PCL was able to initiate heritable repression. It is also not known whether this interaction is significant *in vivo*. *In vivo* co-immunoprecipitation experiments could be performed using PCL and GAL4 antisera to determine whether an association between the amino terminus of PCL and endogenous PCL occurs *in vivo*.

Although an interaction between the amino terminus of PCL and full length PCL was identified, it was not known whether this interaction would be sufficient to recruit the PcG complex to enable heritable repression, given that no other region of PCL was sufficient to initiate repression. It was therefore thought that the amino terminus of PCL would interact with other PcG members. This was tested using yeast two hybrid analysis. A synthetic lethality was observed between the amino terminus of PCL and a 3' deletion of PC. The lethality was specific to yeast containing both LexA-PCL-Amino and AD-PcΔ3' as no lethality was observed between LexA-PCL-Amino and AD-alone, or LexA-alone and AD-PcΔ3'. Although this is suggestive of an interaction between these two proteins, further work needs to be performed to confirm this. GST-pulldown assays and co-immunoprecipitations from embryonic extracts should be performed to confirm the interaction. The reason for the lethality in yeast is not apparent. Perhaps the interaction of the amino terminus of PCL with PcΔ3' is causing repression of genes essential to yeast viability. This would presumably require endogenous yeast proteins. Curiously no lethality was observed between AD-PC and LexA-PCL-Amino. Perhaps the amino terminus interacts more strongly with PcΔ3' than it does with PC, leading to lethality. A synthetic lethality was also generated by AD-ph-D and AD-Scm. An interaction between the amino terminus of PCL and these two constructs could therefore not be tested.

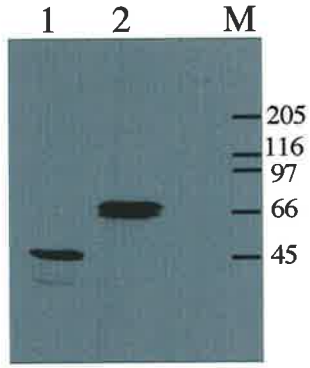
Given that the amino terminus of PCL was sufficient to establish heritable repression it was surprising to observe that GAL-Amino+CarboxyΔPHD was not (Figure 4.4). There are three potential reasons for this. The first is that the GAL-Amino+CarboxyΔPHD is not being stably expressed and nuclear localised. Western analysis and immunohistochemistry need to be

Figure 4.11: Neither the PCL_{1-200} nor the $PCL_{132-423}$ is able to interact with PcΔ3' in the two hybrid assay.
In order to confirm the interaction between the amino terminus of PCL and PcΔ3', the amino terminus of PCL was divided up into two overlapping fragments, a PCL_{1-200} fragment and a $PCL_{132-423}$ fragment. These fragments were generated using PCR and the PCL_{1-200} fragment (aa1-200) was cloned into pEG202 to generate LexA-PCL_{1-200}. The $PCL_{132-423}$ fragment (aa132-423) was cloned into pGilda to generate LexA-$PCL_{132-423}$. (A) is a schematic of the PCL protein showing the position of the PHD fingers in green. The position of the PCL-Amino, PCL_{1-200}, and $PCL_{132-423}$ fragments are represented. The numbers below represent the amino acid positions.
(B) shows that neither LexA-PCL_{1-200} nor LexA-$PCL_{132-423}$ interact with AD-PcΔ3' in the yeast two hybrid assay. Yeast were streaked sequentially onto GLU, followed by GAL, followed by GLU LEU plates
(C) is a western blot probed with the LexA-antisera showing the expression of LexA-PCL_{1-200} (Lane 1) and LexA-$PCL_{132-423}$ (Lane 2). Expected sizes for LexA-PCL_{1-200} and LexA-$PCL_{132-423}$ are 48kDa and 58kDa respectively. LexA-$PCL_{132-423}$ is larger than the predicted size which could indicate post-translational modification of the protein.

A



C



Lane 1:LexA-PCL_{1-200}
Lane 2:LexA-$PCL_{132-423}$
M: Molecular weight markers in kilodaltons

B

performed to test this. The second is that removal of the PHD fingers leads to improper folding of the amino and carboxy termini. This possibility could be examined using far western analysis. Given that the amino terminus of PCL interacts with full length PCL, this could be used to assay for the proper folding of the Amino+CarboxyΔPHD construct. The third possibility is that one role of the carboxy terminus of PCL could be to inhibit the function of the amino terminus and therefore prevent the association of the amino terminus with PCL and possibly PC. This possibility seems unlikely given that GAL-PCL is able to silence and it contains both the amino and carboxy terminus.

4-7.2 The PHD fingers are unable to initiate heritable repression

Given the interaction between the PHD fingers of PCL and E(Z) (see Chapter 3), it was thought that the PHD fingers may be responsible for initiating the heritable repression conferred by the GAL-PCL fusion protein through the recruitment of E(Z). However the GAL-PHD fusion protein was unable to initiate heritable repression. There are several possible reasons for this. The simplest explanation is that the GAL-PHD fusion protein is not being stably expressed or nuclear localised. Improper folding of the PHD fingers could also be inhibiting their ability to initiate heritable repression. However this possibility is unlikely given that the PHD fingers are stably produced and folded in yeast. The *in vivo* interaction between GAL-PHD and E(Z) could be tested with co-immunoprecipitation assays using GAL4 and E(Z) antibodies.

An alternative explanation is that the GAL-PHD fusion protein is being produced, but the interaction between the PHD fingers and E(Z) is not stable or strong enough to initiate heritable silencing. It is likely that establishment of heritable repression requires both recruitment of the E(Z)/ESC and PH/PC/SCM/PSC complexes. Recruitment of only one of these complexes to a target locus could be insufficient.

To date the only interactors of E(Z) are ESC and PCL (Tie *et al.*, 1998; Jones *et al.*, 1998; Chapter 3). The GAL-PHD construct could recruit the E(Z)/ESC complex, but would be unable to recruit the PH/PC/SCM/PSC complex as no interactions have been described between the ESC/E(Z) and PH/PC/SCM/PSC complexes. It would therefore appear as though the recruitment of E(Z) and therefore ESC is not sufficient for the establishment of heritable repression and that interactions between PCL and other PcG members are required, perhaps for complex stability (see Figure 4.12 for a schematic diagram). These other interactions are not able to be formed because no other region of PCL is present in the GAL-PHD construct.

The GAL-Amino fusion protein may therefore be able to establish heritable silencing because it is able to recruit both the E(Z)/ESC complex and the PH/PC/SCM/PSC complex. The E(Z)/ESC complex is recruited through the association of the GAL-Amino fusion with

endogenous PCL, which contains the PHD finger region. The GAL-Amino construct could also recruit the PH/PC/SCM/PSC complex through the association of the amino terminus of PCL with PC (see Figure 4.12 for a schematic representation). However this interaction needs to be confirmed. The recruitment of both the E(Z)/ESC and PH/PC/SCM/PSC complexes would lead to initiation of stable repression.

An alternative explanation is that the interaction between the PHD fingers of PCL and E(Z) is not required for repression in this assay and that a requirement for this interaction is bypassed by tethering PCL to DNA. This could be tested directly by mutating the PHD fingers in GAL-PCL to determine whether this interaction is required. There is some speculation that E(Z) and ESC form a PcG initiating (PcGi) complex which is responsible for initiating repression and recruiting the PcG maintenance (PcGm) complex, which then establishes heritable silencing (van Lohuizen, 1999). Evidence to support this comes from experiments in both *Drosophila* and mammals which show that ESC/E(Z) form a separate complex which does not appear to be associated with the PC/PH/PSC/SCM (PcGm) complex (reviewed in van Lohuizen, 1999; Tie *et al.,* 1998; Jones *et al.,* 1998, Shao *et al.,* 1999). No direct link between the PcGi and PcGm exists. Perhaps the PcGm is recruited to target loci by PcGi through the interaction of E(Z) with PCL and PCL with PC. Arguing against a bypassing of the requirement for the ESC/E(Z) complex, is the requirement for endogenous E(Z) in the early silencing conferred by both GAL-PCL and GAL-Amino (Figure 4.8). This suggests that E(Z) is an integral component of the repression initiated by both GAL-PCL and GAL-Amino.

Although E(Z) has not been detected in the Polycomb Repressive Complex 1 (PRC1), genetic evidence suggests that it has a role in maintaining PcG-mediated repression (Jones and Gelbart, 1990). Experiments performed using a temperature sensitive allele of *E(z)* also demonstrated a continual requirement for E(Z) throughout development (Jones and Gelbart, 1990). Furthermore, a new E(Z) binding site is generated at the site of a PRE insertion (Carrington and Jones, 1996). These are all consistent with E(Z) being a member of the PcGm.

An alternative explanation is that E(Z) is not an integral member of the PcGm but instead is required for chromosome integrity and stabilisation of PcGm on chromatin. Consistent with this is the abnormal morphology of mutant *E(z)* salivary glands polytene chromosomes and the dependence on E(Z) function for binding of PSC and SU(Z)2 to chromosomes (Rastelli *et al.,* 1993). Although *E(z)* mutants have an obvious effect on chromosome morphology, and E(Z) function is required for the binding of all PcG members tested, the function of other PcG members is unable to be assayed in this way as no other temperature sensitive mutants are available. It is therefore not known if mutations in all PcG members would lead to abnormal chromosome morphology or if this phenotype is specific for E(Z). Clearly, characterisation of the role of E(Z) as a member of either the PcGi or PcGm or both, is required. Characterisation of the role of E(Z) and other PcG proteins in chromosome morphology is also required.

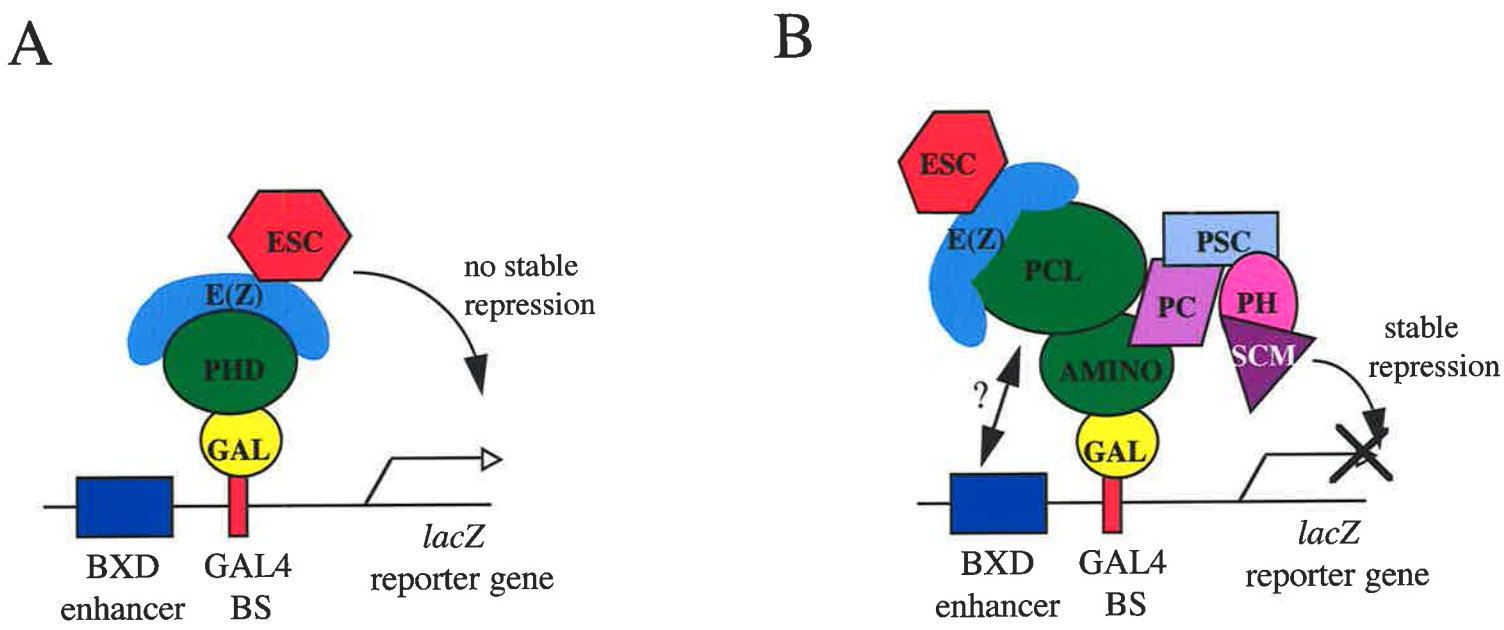


Figure 4.12: A schematic diagram of a model for why the GAL-PHD fusion protein did not initiate heritable repression.
(A) The GAL-PHD fusion protein may recruit the E(Z)/ESC complex but this complex is unable to recruit any other PcG members and therefore repression is not established. (B) The GAL-Amino fusion protein is able to initiate heritable silencing as it brings in endogenous PCL and perhaps PC which associates with PSC, PH and SCM. The PHD fingers of endogenous PCL will also recruit the E(Z)/ESC complex through their association with E(Z). The recruitment of the endogenous PcGi and PcGm leads to stable heritable repression. The ? represents the unknown link between the BXD enhancer and the PcG.

The results described in this chapter demonstrate a PcG dependent role for the amino terminus of PCL in establishment of a heritable repression complex. The role of the amino terminus in heritably repressing transcription is likely to depend on homotypic interactions with PCL and also possibly heterotypic interactions with PC.

Chapter 5: Identification and characterisation of interactors of the carboxy terminus of PCL.

5-1 Introduction

The aims of Chapter 3 and 4 of this thesis were to characterise a previously identified interaction between PCL and E(Z) and to identify functionally important regions of the PCL protein. The results identified the PHD fingers and the amino terminus of PCL as being important in its function. No functional activity has been attributed to the C-terminus of PCL. This region was not sufficient to initiate heritable repression of the *lacZ* reporter gene in the tethering assay, either because the C-terminus of PCL has no repressive function or because it has a repressive function that is not identified in this assay. It was therefore the aim of the work described in this chapter to identify and characterise any PcG interactors of the C-terminus using the yeast two hybrid assay.

5-2 Identification of PCL carboxy terminus interactors

In order to identify interactions between the carboxy terminus of PCL (aa567-857) and cloned members of the PcG, two LexA-fusion constructs were generated in pEG202; a LexA-PCL-Carboxy (LexA-PCL-C) (aa567-857) and a LexA-PCL-PHD+Carboxy (LexA-PCL-PHD+C) (aa424-857) construct. The overlapping LexA-PCL-PHD+C construct was generated to ensure that no protein/protein interaction domain(s) was disrupted by removal of the PHD fingers in the LexA-PCL-C construct. With the exception of AD-PHO, which had not been cloned at the time, the PHD fingers had already been tested for an interaction with the AD-PcG constructs (Robert, 1997). Both constructs were generated using high fidelity PCR and the nucleotide sequence of the LexA-PCL fusion constructs were sequenced to ensure the correct reading frame was maintained from LexA into PCL. Both LexA-PCL-PHD+C and LexA-PCL-C were transformed into EGY48 and western analysis was performed using anti-LexA antisera to confirm that the constructs were being expressed (data not shown). Both constructs were also tested for self-activation of the *LEU2* reporter gene in EGY48. LexA-PCL-C did not auto-activate the *LEU2* reporter gene (Figure 5.1 compare GAL plate with GLU plate). Note that the expression of the AD-construct only occurs on GAL media and therefore any growth on GLU media is due to the ability of the LexA-fusion construct to activate. LexA-PCL-PHD+C was able to weakly activate the reporter gene (data not shown) and this construct was therefore transformed into EGY191, a yeast strain containing only two LexA binding sites upstream of the endogenous *LEU2* gene. The reduced number of binding sites (2 in EGY191 compared to 6 in EGY48) leads to less activation of the reporter gene and therefore can reduce the ability of a construct to self-

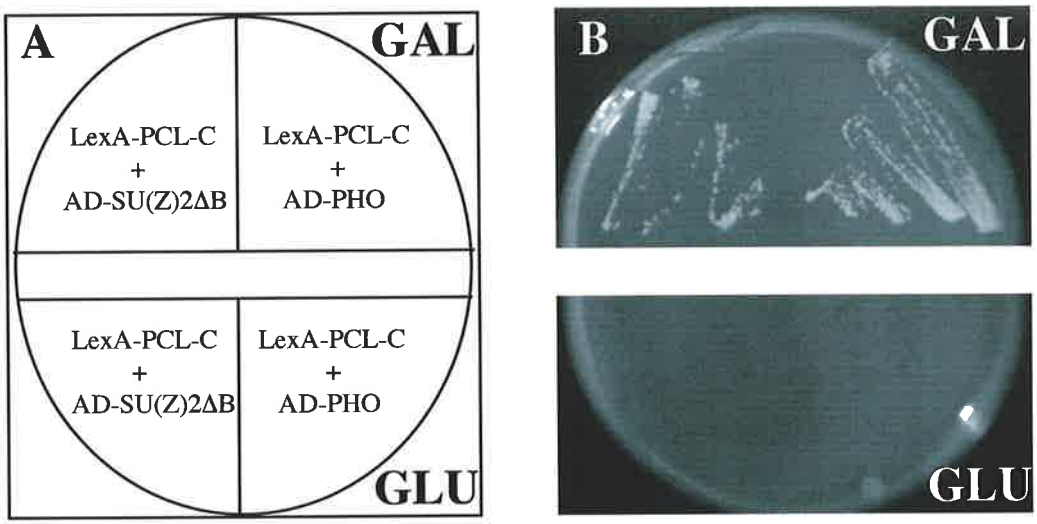


Figure 5.1: The carboxy terminus of PCL interacts with PHO and SU(Z)2ΔB in the yeast two hybrid assay. The carboxy terminus of PCL (aa 567-857) was cloned into the yeast two hybrid vector pEG202 to generate LexA-PCL-Carboxy. This was then tested for an interaction against an array of AD-PcG members (see Chapter 4 for list of AD-PcG members). AD-PHO and AD-SU(Z)2ΔB displayed an interaction with LexA-PCL-C. (A) is a schematic diagram of the arrangement in (B). The interaction between LexA-PCL-C and AD-PHO and AD-SU(Z)2ΔB is shown. These interactions are dependent on the induction of the AD-fusion protein, which is under galactose inducible control, as growth was only observed on GAL and not on GLU media.

activate (Ausubel *et al.*, 1995). The LexA-PCL-PHD+C construct was unable to self-activate the reporter gene in this yeast strain (Figure 5.2 compare (B) and (C)) and therefore EGY191 was used to test for interactions between LexA-PCL-PHD+C and the AD-PcG constructs. The AD-PcG constructs (see section 4.6 for a list) were tested for an interaction with LexA-PCL-C and LexA-PCL-PHD+C by assaying for growth in the absence of exogenously supplied leucine and by comparing growth on GAL with growth on GLU plates.

LexA-PCL-C interacted with two PcG members; AD-Pleiohomeotic (AD-PHO) and AD-Suppressor of zeste two ΔB (AD-SU(Z)2ΔB) (Figure 5.1). LexA-PCL-PHD+C interacted with three AD-PcG constructs; AD-E(Z), AD-PHO and AD-SU(Z)2ΔB (Figure 5.2). The interactions between LexA-PCL-C and LexA-PCL-PHD+C and the AD-PcG constructs were only observed on GAL media and not GLU media and therefore required induction of the AD-fusion protein (Figure 5.1 and Figure 5.2). The interaction between AD-E(Z) and LexA-PCL-PHD+C and not with the LexA-PCL-C fusion is consistent with the results discussed in Chapter 3 of this thesis, which demonstrated an interaction between E(Z) and the PHD fingers of PCL. The growth observed on GAL plates containing LexA-PCL-C and AD-Su(Z)2ΔB was consistently greater than the growth observed on GAL plates containing LexA-PHD+C and AD-Su(Z)2ΔB. Whether this reflects a difference in the strength of the interaction between these two fragments of PCL and Su(Z)2ΔB was not investigated. However, a correlation is generally observed between the strength of a two hybrid interaction and the strength of an interaction *in vitro* (Estojak *et al.*, 1995).

AD-Su(Z)2ΔB is a C-terminal deletion of SU(Z)2, encoding amino acids 1-478 (Hugh Brock, pers. comm.). This region of SU(Z)2 can be divided into two domains; a homology region (HR) which displays sequence similarity to two mammalian homologues Bmi-1 and Mel-18 and a region C-terminal to the HR, called the similar amino acid content region (SAACR)(see Figure 5.3 for a schematic)(Brunk *et al.*, 1991). Full length SU(Z)2 (AD-SU(Z)2) failed to interact with either LexA-PCL-C or LexA-PCL-PHD+C in this assay (data not shown), the implications of which are considered in the discussion. Further characterisation of the regions required for interaction between PCL/PHO and PCL/Su(Z)2 was undertaken. The results of these studies are discussed separately in the following sections.

5-3 Characterisation of the interaction between PCL and PHO

5-3.1 Identification of the PCL domain required for the interaction with PHO.

Pleiohomeotic is a recently cloned member of the PcG (Brown *et al.*, 1998). It encodes the only identified DNA binding member of the Polycomb group. Due to its recent cloning, it was not tested in the initial two hybrid screen that identified E(Z) as an interactor of the cDOM of PCL or full length PCL (LexA-PCL) (Robert, 1997). To complete the analysis, LexA-PCL and LexA-cDOM were tested against AD-PHO in EGY48 and no interaction was observed between either LexA-PCL or LexA-cDOM and AD-PHO (Figure 5.4) (see discussion). The failure of LexA-cDOM to interact with AD-PHO, confirms that the interaction between PCL and PHO is mediated through the C-terminus of PCL.

It is currently unknown whether the PcG form a single type of multimeric protein complex at each target locus or whether the PcG members form several different types of complexes, each differing in their components. Antibody stainings performed on salivary polytene chromosomes show only a partial overlap in binding patterns for some members of the PcG and a complete overlap in the binding pattern for others (Carrington and Jones, 1996; Franke *et al.*, 1993; Lonie *et al.*, 1994; Martin *et al.*, 1993; Rastelli *et al.*, 1993), thus favouring the model that several different multimeric PcG complexes exist within a nucleus. SU(Z)2 shares 53% of its polytene binding sites with PCL (Rastelli *et al.* 1993) and the polytene binding sites of PHO are yet to be published. It was therefore of interest to determine whether the region in PCL responsible for the interaction with PHO and SU(Z)2 was the same or different. If the two proteins interacted with the same region of PCL, this could provide an explanation for the formation of different PcG complexes and also imply a different function for SU(Z)2 versus PHO-containing complexes.

To determine whether PHO and SU(Z)2 interact with the same region of PCL, three overlapping LexA-constructs were generated by PCR; LexA-$PCL_{567\text{-}700}$ (aa567-700), LexA-$PCL_{634\text{-}800}$ (aa634-800) and LexA-$PCL_{700\text{-}857}$ (aa700-857) (see Figure 5.5 for a schematic). Overlapping constructs were generated to ensure that no unidentified domain was disrupted. All three constructs were transformed into EGY48 and tested for self-activation. All three were able to self-activate the *LEU2* reporter gene, they were therefore transformed into EGY191, containing fewer LexA binding sites, and when tested for self-activation were unable to do so. LexA-$PCL_{567\text{-}700}$, LexA-$PCL_{634\text{-}800}$ and LexA-$PCL_{700\text{-}857}$ in EGY191, were therefore tested for an interaction with AD-PHO. LexA-$PCL_{567\text{-}700}$ and LexA-$PCL_{634\text{-}800}$ were unable to interact with AD-PHO whereas, LexA-$PCL_{700\text{-}857}$ was able to interact with AD-PHO (Figure 5.6). The inability of LexA-$PCL_{634\text{-}800}$ and the ability of LexA-$PCL_{700\text{-}857}$ to interact with AD-PHO narrowed down the region of interaction to aa800-857.

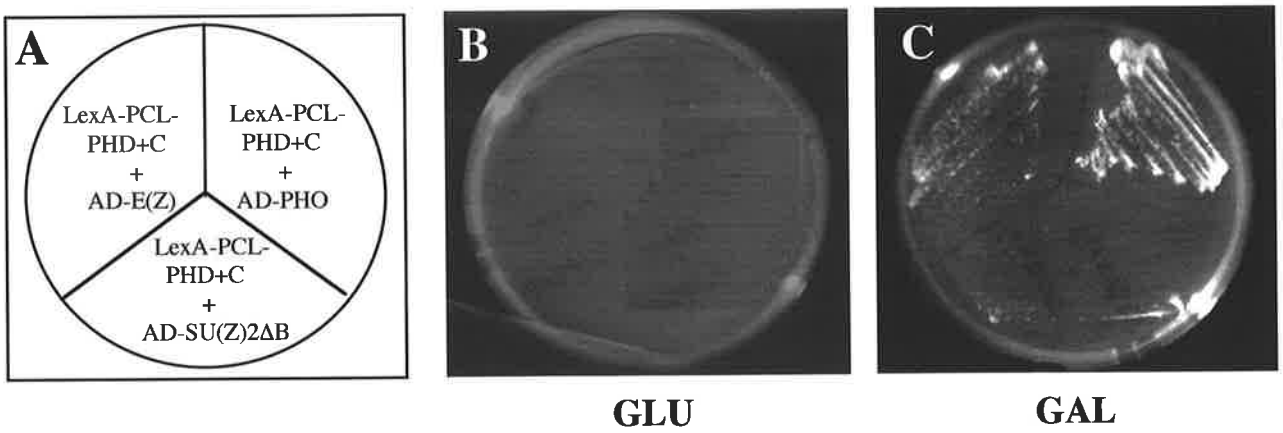


Figure 5.2: The PHD+Carboxy terminus of PCL interacts with E(Z), PHO and SU(Z)2ΔB in the yeast two hybrid assay. The PHD fingers and the carboxy terminus of PCL (aa 424-857) were cloned into the yeast two hybrid vector, pEG202 to generate LexA-PCL-PHD+C. This was then tested for an interaction against an array of AD-PcG members. AD-E(Z), AD-PHO and AD-SU(Z)2ΔB displayed an interaction with LexA-PCL-PHD+C. (A) is a schematic diagram of the arrangement in (B) and (C). (B) is a glucose control showing that LexA-PCL-PHD+C is unable to activate expression of the *LEU2* reporter gene and that the interaction between LexA-PCL-PHD+C and AD-PHO, AD-E(Z) and AD-SU(Z)2ΔB (shown in (C)) is dependent on induction of the AD-fusion proteins, which are under galactose inducible control. (C) shows the interaction between LexA-PCL-PHD+C and AD-PHO, AD-E(Z) and AD-SU(Z)2ΔB. The interaction between LexA-PCL-PHD+C and AD-SU(Z)2ΔB was consistently weaker t han the interaction between LexA-PCL-C and AD-SU(Z)2ΔB (see Figure 5.1 for a comparison).

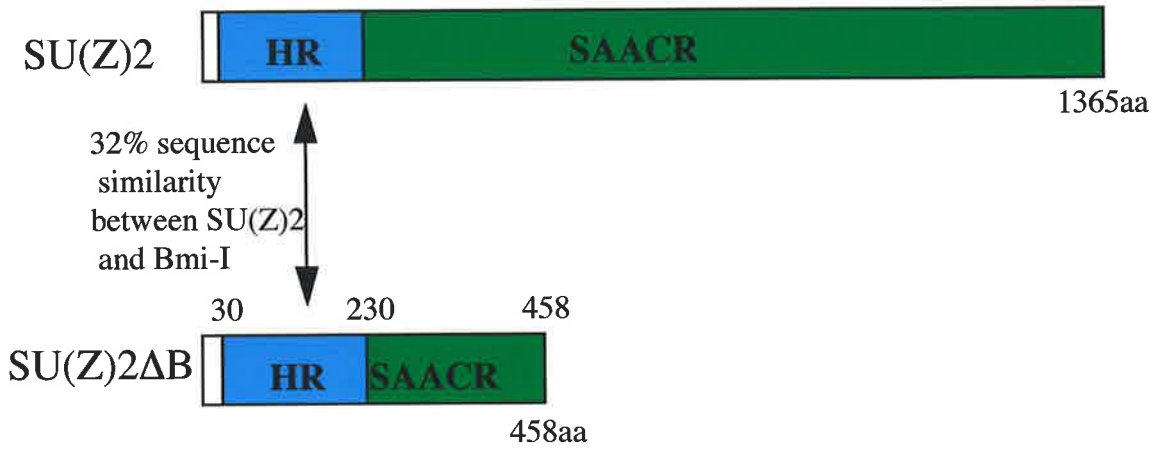


Figure 5.3: A schematic representation of the SU(Z)2ΔB fragment that was shown to interact with LexA-PCL-C and LexA-PCL-PHD+C.

Su(Z)2 is a 1365aa protein that shows sequence similarity to two mammalian homologues (Bmi-1 and Mel-18) and *Drosophila* Posterior Sex Combs (PSC) over a region termed the homology region (HR). The rest of the protein shows a similar amino acid composition to PSC and is termed the Similar Amino Acid Content Region (SAACR). The SU(Z)2ΔB fragment encodes the HR and 228aa of the SAACR. The SAACR in the mammalian homologues is much shorter than that of the Drosophila proteins.

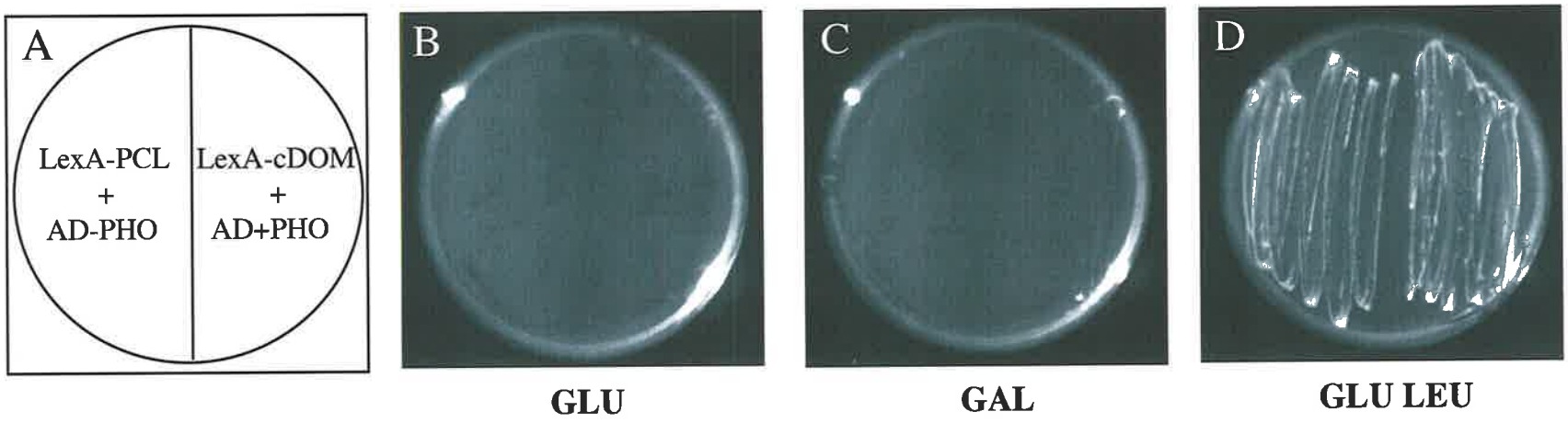


Figure 5.4: PHO does not interact with either full length PCL or the cDOM of PCL.

Pleiohomeotic is a recently cloned member of the PcG (Brown *et al.*, 1998) and was not tested in the two hybrid matrix performed by S. Robert using PCL and the cDOM of PCL. To complete the analysis LexA-PCL and LexA-cDOM were tested against AD-PHO in the yeast two hybrid assay. (A) is a schematic of the arrangement in (B), (C) and (D). (B) shows that LexA-PCL and LexA-cDOM are unable to self-activate the reporter gene on GLU medium. LexA-PCL and LexA-cDOM are unable to interact with AD-PHO (C). D is a positive control GLU LEU plate which was streaked last to ensure that yeast had been streaked onto (B) and (C).

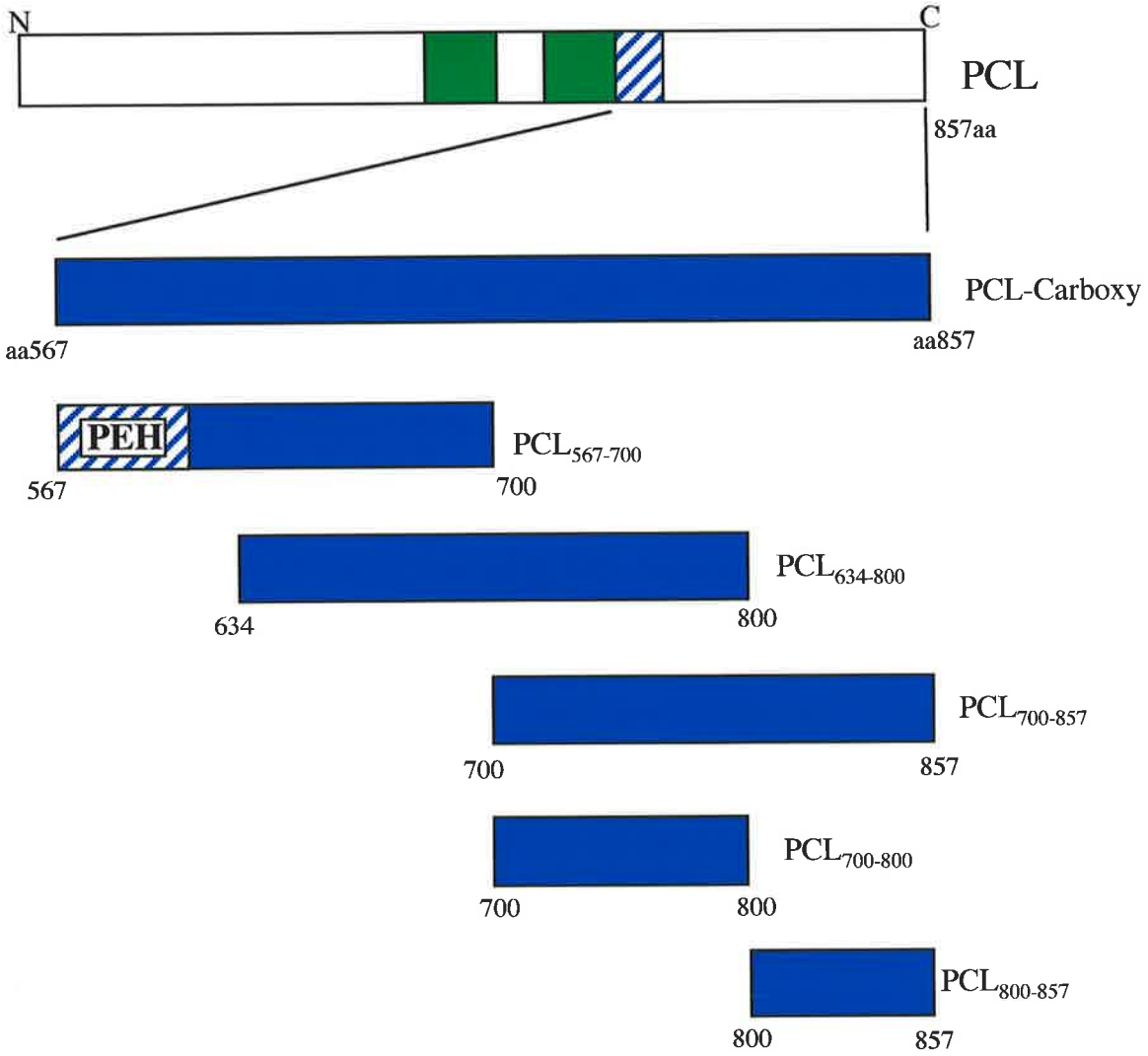


Figure5.5: A schematic representation of the carboxy terminus constructs generated to refine the PHO and SU(Z)2 interaction domains within PCL.

The C-terminus of PCL was divided up into three overlapping fragments; $PCL_{567-700}$ (aa567-700), $PCL_{634-800}$ (aa634-800) and $PCL_{700-857}$ (aa700-857). Both PHO and SU(Z)2 interacted with the $PCL_{700-857}$ fragment. This fragment was therefore further divided upinto a $PCL_{700-800}$ (aa700-800) and a $PCL_{800-857}$ (aa800-857) fragment to enable further refinement of the PCL interaction domains. The position of the two PHD finger motifs in full length PCL is shown in green, as is the position of the Polycomblike Extended Homology (PEH) domain (in blue)(see Chapter 3), which lies within $PCL_{567-700}$ fragment. The amino acid position of each fragment is shown with respect to the full length amino acid sequence of PCL.

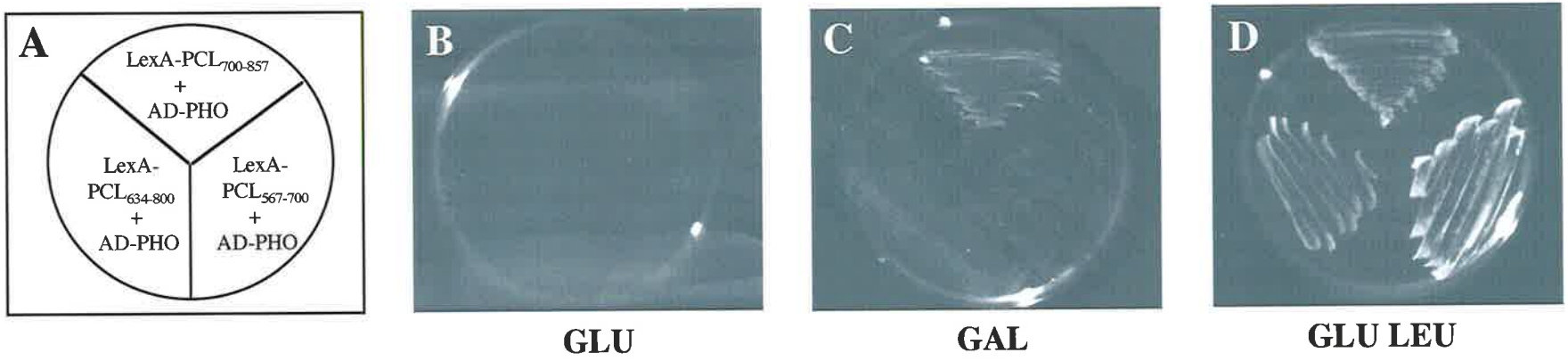


Figure 5.6: PHO interacts with the C-terminal 57aa of PCL.

The carboxy terminus of PCL was divided into three overlapping fragments; $PCL_{567-700}$, $PCL_{634-800}$ and $PCL_{700-857}$. Each fragment was cloned into pEG202 to generate LexA-$PCL_{567-700}$, LexA-$PCL_{634-800}$, and LexA- $PCL_{700-857}$ and tested for its ability to interact with AD-PHO. Three independant transformant colonies were streaked in each quadrant. GLU plates (B) were streaked first, followed by GAL (C), then GLU LEU (D) to ensure that yeast were being streaked onto each plate. (A) is a schematic diagram of the arrangement on (B), (C) and (D).

(B) shows that LexA-$PCL_{567-700}$, LexA-$PCL_{634-800}$, and LexA-$PCL_{700-857}$ were unable to self-activate expression of the LEU reporter gene. (C) shows that LexA-$PCL_{700-857}$ was able to interact with AD-PHO. Both LexA-$PCL_{567-700}$ and LexA-$PCL_{634-800}$ were unable to interact with AD-PHO indicating that the last 57 amino acids of PCL is responsible for mediating the interaction with AD-PHO. (D) is a positive control plate which was streaked last demonstrating that yeast was streaked onto the GLU and GAL plates.

To further refine this interaction, the $PCL_{700-857}$ fragment was divided into a LexA-$PCL_{700-800}$ (aa700-800) and a LexA-$PCL_{800-857}$ (aa800-857) construct (see Figure 5.5 for a schematic). Both constructs were transformed into EGY48 and tested for auto-activation. LexA-$PCL_{700-800}$ was unable to activate the reporter gene (Figure 5.7 compare (B) with (C)). LexA-$PCL_{800-857}$ however was able to strongly activate the *LEU2* reporter gene in both EGY48 and EGY191 and was therefore unable to be tested for an interaction with AD-PHO (Figure 5.7 B). LexA-$PCL_{700-800}$ did not interact with AD-PHO (Figure 5.7) inferring that aa800-857 of PCL is important for mediating the interaction between PCL and PHO. This sequence was used to search for any related sequences using the BLAST program (Atschul *et al.,* 1997)(http://www.ncbi.nlm.nih.gov). No proteins were identified that had any significant similarity to this region of PCL (data not shown). Analysis using the COILS program (Lucas *et al.,* 1991) (http://ulrec3.unil.ch/software/COILS_form.html.) failed to reveal a coiled-coil domain, while the program MOTIF (http://www.genome.ad.jp/) failed to identify any common, previously described motif in the PHO-interacting sequence.

5-3.2 The PHO interaction region

PHO is the *Drosophila* homologue of the mammalian transcription factor YY1 (Brown *et al.,* 1998). YY1 has been shown to act as both a transcriptional activator and a repressor depending on the context (reviewed in Shi *et al.,* 1997). The region of YY1 responsible for mediating repression of target genes has been mapped and corresponds to the Zn finger region (Bushmeyer, *et al.,* 1995; Austen *et al.,* 1997). This region is highly conserved in PHO (Brown et. al. 1998). The region mediating transcriptional activation of target genes has also been mapped in YY1 although these regions are not conserved in PHO (Bushmeyer *et al.,* 1995; Austen *et al.,* 1997; Brown *et al.,* 1998). To determine which region of PHO is responsible for mediating the interaction with PCL and whether this corresponds to the putative region of repression in PHO, two overlapping AD-PHO constructs were generated in pJG4-5 to produce AD-PHO-Amino (aa 1-384) and AD-PHO-Carboxy (aa 278-521). AD-PHO-Carboxy contains the region of PHO which encodes the four zinc fingers which are known to mediate repression in YY1 (Bushmeyer *et al.,* 1995 and Austen *et al.,* 1997). The vector pJG4-5 contains a HA epitope tag which enables confirmation of the expression of constructs. After transformation into EGY48, expression of AD-PHO-Amino and AD-PHO-Carboxy was confirmed by western analysis using the anti-HA antibody (Figure 5.8). Both AD-PHO-Amino and AD-PHO-Carboxy were then tested for an interaction with LexA-PCL-C on leucine deficient media. Neither AD-PHO-Amino nor AD-PHO-Carboxy were able to interact with LexA-PCL-C in this assay (Figure 5.8). This suggests that PHO contains a bipartite PCL interaction domain that was separated in the AD-PHO constructs. Alternatively the AD-PHO-Amino and AD-PHO-Carboxy are unable to fold properly.

5-4 Characterisation of the interaction between PCL and SU(Z)2

5-4.1 The PCL interaction domain

In order to identify the region of PCL responsible for the interaction with AD-Su(Z)2ΔB, and to determine whether this is the same as that of the PHO interaction region, AD-Su(Z)2ΔB was tested for an interaction with LexA-$PCL_{567-700}$, LexA-$PCL_{634-800}$ and LexA-$PCL_{700-857}$ (previously described in 5-3.1). LexA-$PCL_{567-700}$ was unable to interact with AD-Su(Z)2ΔB (Figure 5.9), LexA-$PCL_{634-800}$ and LexA-$PCL_{700-857}$ however were able to interact with AD-Su(Z)2ΔB (Figure 5.9). This interaction was dependent on the induction of the AD-Su(Z)2ΔB fusion protein (Figure 5.9 GAL vs GLU plates). The interaction between LexA-$PCL_{634-800}$ and AD-Su(Z)2ΔB consistently appeared weaker than the interaction between LexA-$PCL_{700-857}$ and AD-Su(Z)2ΔB. Given the inability of the AD-Su(Z)2ΔB fusion to interact with LexA-$PCL_{567-700}$ and the ability to interact with both LexA-$PCL_{634-800}$ and LexA-$PCL_{700-857}$, this narrowed down the SU(Z)2 interaction region of PCL to the region of overlap between these two constructs (aa700 -800). To test whether these amino acids were responsible for the interaction, LexA-$PCL_{700-800}$ (see section 5.5-2 for a description of this construct) was tested for an interaction with AD-Su(Z)2ΔB. LexA-$PCL_{700-800}$ was unable to interact (Figure 5.10). The inability of the LexA-$PCL_{700-800}$ construct to interact could potentially be due the interruption of a protein/protein interaction domain, or the presence of a bipartite interaction domain. Alternatively it could be due to improper folding of the LexA-$PCL_{700-800}$ fusion protein.

To determine whether this region of PCL encodes any domains that had not been identified in previous database searches, aa700-857 were compared with database sequences using the BLAST program (http://www.ncbi.nlm.nih.gov) (Altshul *et al.*, 1997). No proteins were identified that had any significant similarity to this region of PCL. This region of PCL was also analysed using the COILS (http://ulrec3.unil.ch/software/ COILS_form.html.) program to determine whether it was likely to encode a coiled-coil motif. A 25 amino acids region was identified as having a high probability of forming a coiled-coil (Figure 5.11). This region corresponded to the region of overlap between LexA-$PCL_{634-800}$ and LexA-$PCL_{700-857}$ (Figure 5.11), both of which were able to mediate an interaction with AD-Su(Z)2ΔB. Whether this coiled-coil mediates the interaction between PCL and SU(Z)2 is not yet known, however it is known that coiled-coils mediate protein/protein interactions with other coiled-coil motifs (Ho *et al.*, 1994). If the coiled-coil in PCL is responsible for mediating the interaction between PCL and SU(Z)2, SU(Z)2 must also contain a coiled-coil motif. Su(Z)2ΔB was therefore analysed using the COILS program. As shown in Figure 5.12, Su(Z)2ΔB has a strong probability of containing two coiled-coil motifs and a weaker probability of containing a third. All three putative coiled-coil domains map to the SAACR of SU(Z)2.

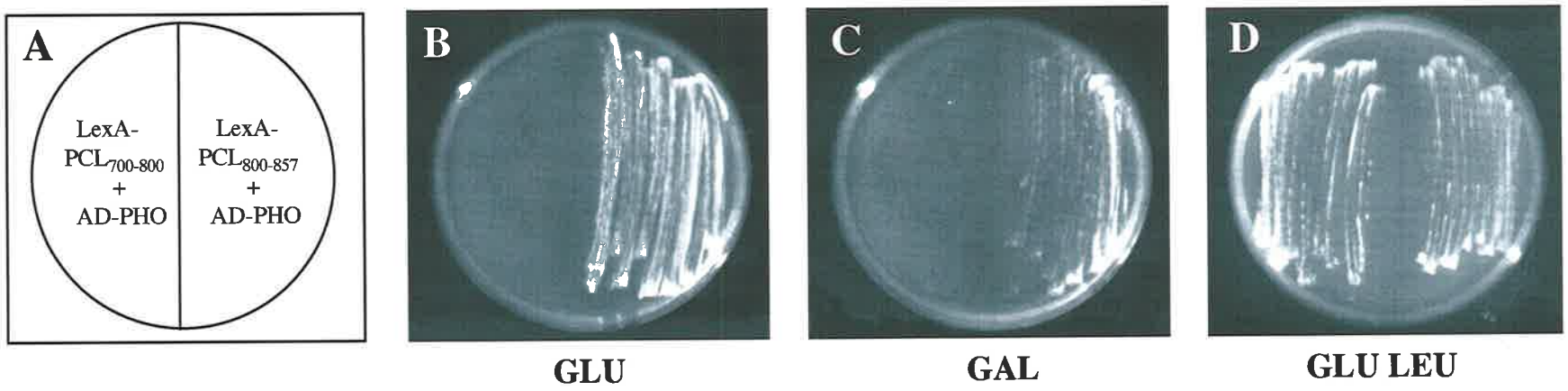
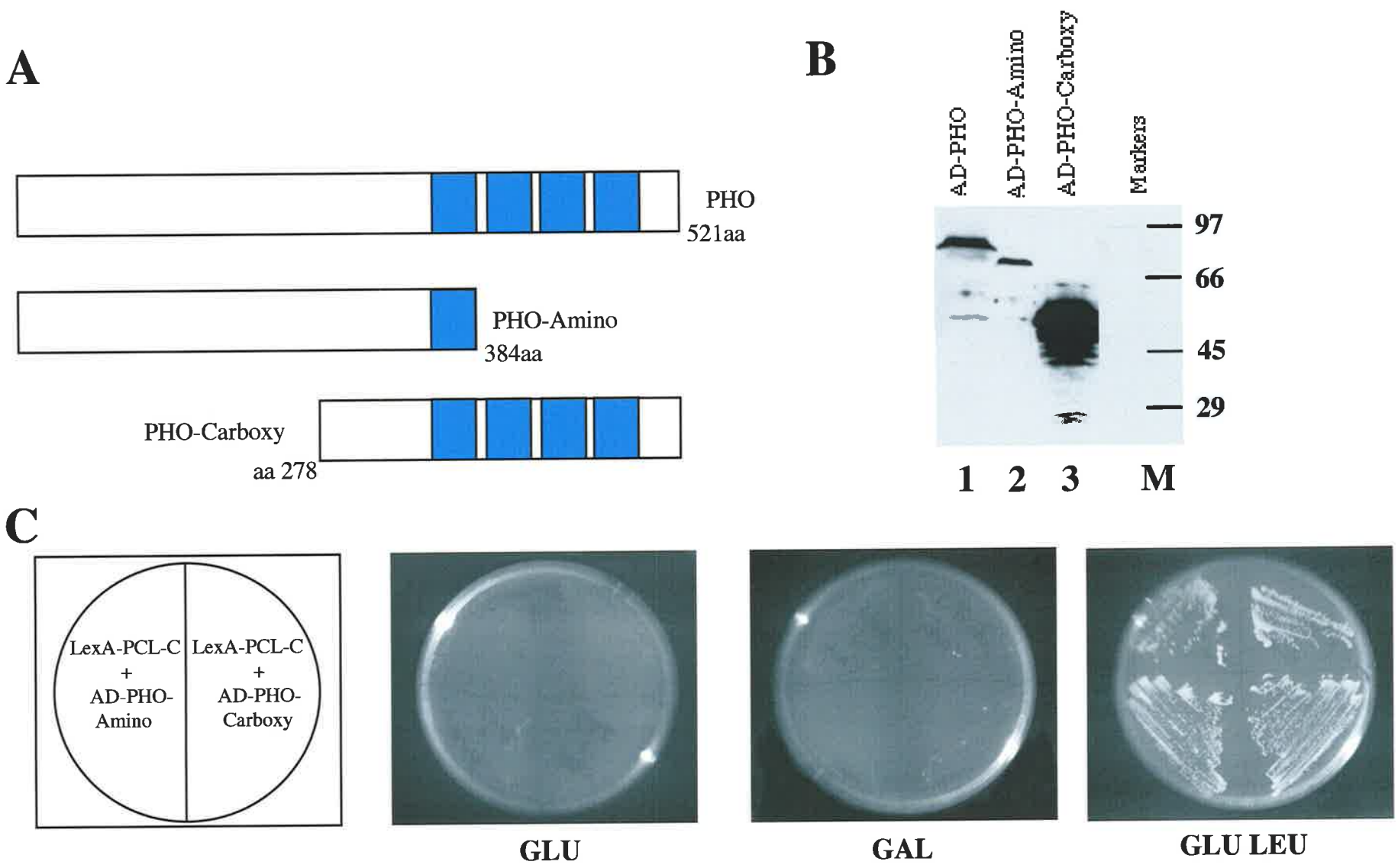


Figure 5.7: Interaction of AD-PHO with the carboxy terminus of PCL

The $PCL_{700-857}$ fragment of PCL was divided into two fragments and cloned into pEG202 to generate LexA-$PCL_{700-800}$ and LexA-$PCL_{800-857}$. These constructs were then tested for their ability to interact with AD-PHO. GLU plates (B) were streaked first, followed by GAL (C), then GLU LEU (D) to ensure that yeast were being streaked onto each plate. (A) is a schematic diagram of the arrangement on (B), (C) and (D). (B) shows that LexA-$PCL_{800-857}$ is able to activate expression of the *LEU2* reporter gene in the absence of the AD-PHO protein, which is not expressed on GLU plates. An interaction between AD-PHO and LexA-$PCL_{800-857}$ was therefore unable to be tested. LexA-$PCL_{700-800}$ is unable to activate expression of the reporter gene. (C) shows the expected growth of the LexA-$PCL_{800-857}$ + AD-PHO yeast on GAL media given its growth on GLU, and it also shows that LexA-$PCL_{700-800}$ is unable to mediate an interaction with AD-PHO. (D) shows that yeast were streaked onto both the GLU and GAL plates.

Figure 5.8: Neither AD-PHO-Amino nor AD-PHO-Carboxy are able to interact with LexA-PCL-Carboxy.
To determine which region of PHO is responsible for mediating the interaction with LexA-PCL-C, PHO was divided into two overlapping fragments and cloned into the yeast two hybrid vector pJG4-5 to generate AD-PHO-Amino and AD-PHO-Carboxy. (A) is a schematic of the PHO protein and the AD-PHO-Amino and AD-PHO-Carboxy constructs. The four Zn finger motifs are shown in blue. The region spanning the Zn fingers is the region responsible for mediating repression by YY1, the mammalian homologue of PHO. This entire region is present in the AD-PHO-Carboxy fragment (A). (B) is a western blot probed with the anti-HA antisera showing the expression of AD-PHO (lane 1), AD-PHO-Amino (lane 2) and AD-PHO-Carboxy in EGY48. Expected sizes for HA-PHO-Amino and HA-PHO-Carboxy are 54kDa and 39kDa respectively. Both HA-PHO-Amino and HA-PHO-Carboxy are larger than the predicted size which could indicate post-translational modification of the proteins. (C) shows that neither AD-PHO-Amino nor AD-PHO-Carboxy were able to interact with LexA-PCL-Carboxy in the yeast two hybrid assay. Yeast were sequentially streaked onto GLU, followed by GAL, followed by GLU LEU plates.


A
PHO
521aa
PHO-Amino
384aa
PHO-Carboxy
aa 278
B
AD-PHO
AD-PHO-Amino
AD-PHO-Carboxy
Markers
97
66
45
29
1
2
3
M
C
LexA-PCL-C + AD-PHO-Amino
LexA-PCL-C + AD-PHO-Carboxy
GLU
GAL
GLU LEU

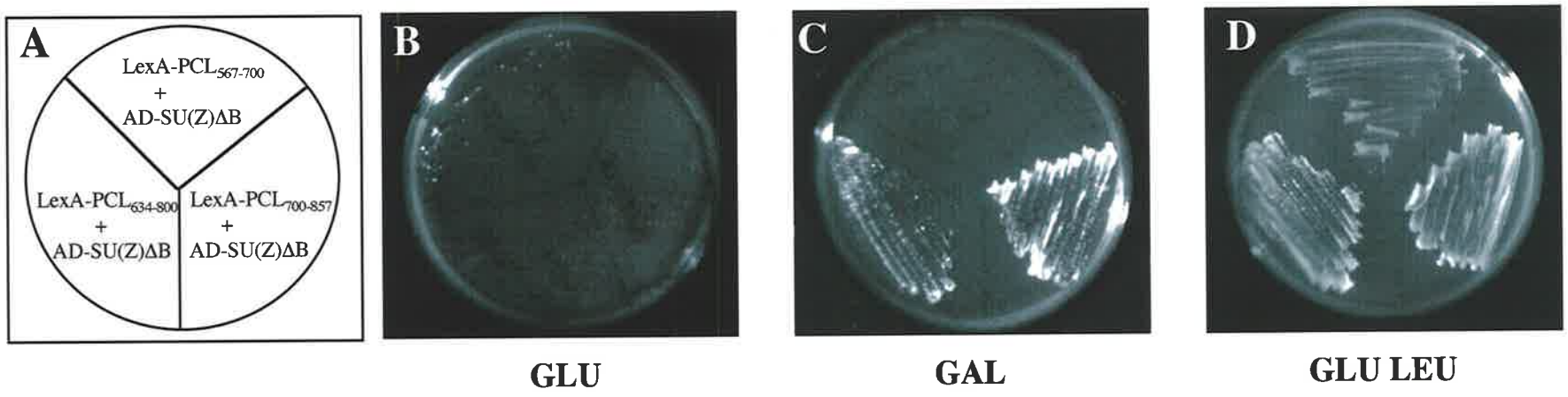


Figure 5.9: AD-SU(Z)ΔB interacts with the $PCL_{634-800}$ and $PCL_{700-857}$ fragments of PCL.

To determine which region of the C-terminus of PCL interacts with AD-SU(Z)ΔB, LexA-$PCL_{567-700}$, LexA-$PCL_{634-800}$ and LexA-$PCL_{700-857}$ were tested for an interaction with AD-SU(Z)ΔB in the yeast two hybrid assay. (A) is a schematic representation of the arrangement on (B), (C) and (D). The GLU plate (B) was streaked first, followed by GAL (C), then GLU LEU (D), to ensure that yeast were streaked onto each plate. (B) shows that LexA-$PCL_{567-700}$, LexA-$PCL_{634-800}$ and LexA-$PCL_{700-857}$ are unable to self-activate *LEU2* reporter gene expression. (C) shows the interaction of LexA-$PCL_{634-800}$ and LexA-$PCL_{700-857}$ with AD-SU(Z)ΔB. The interaction between LexA-$PCL_{634-800}$ and AD-SU(Z)ΔB was consistently weaker than the interaction between LexA-$PCL_{700-857}$ and AD-SU(Z)ΔB. (D) is a growth control plate to show that yeast were streaked onto the GLU and GAL plates.

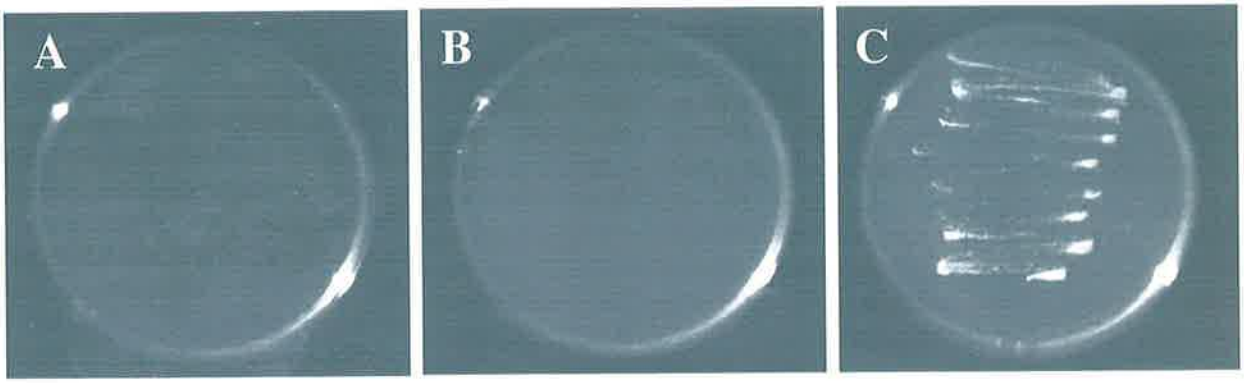
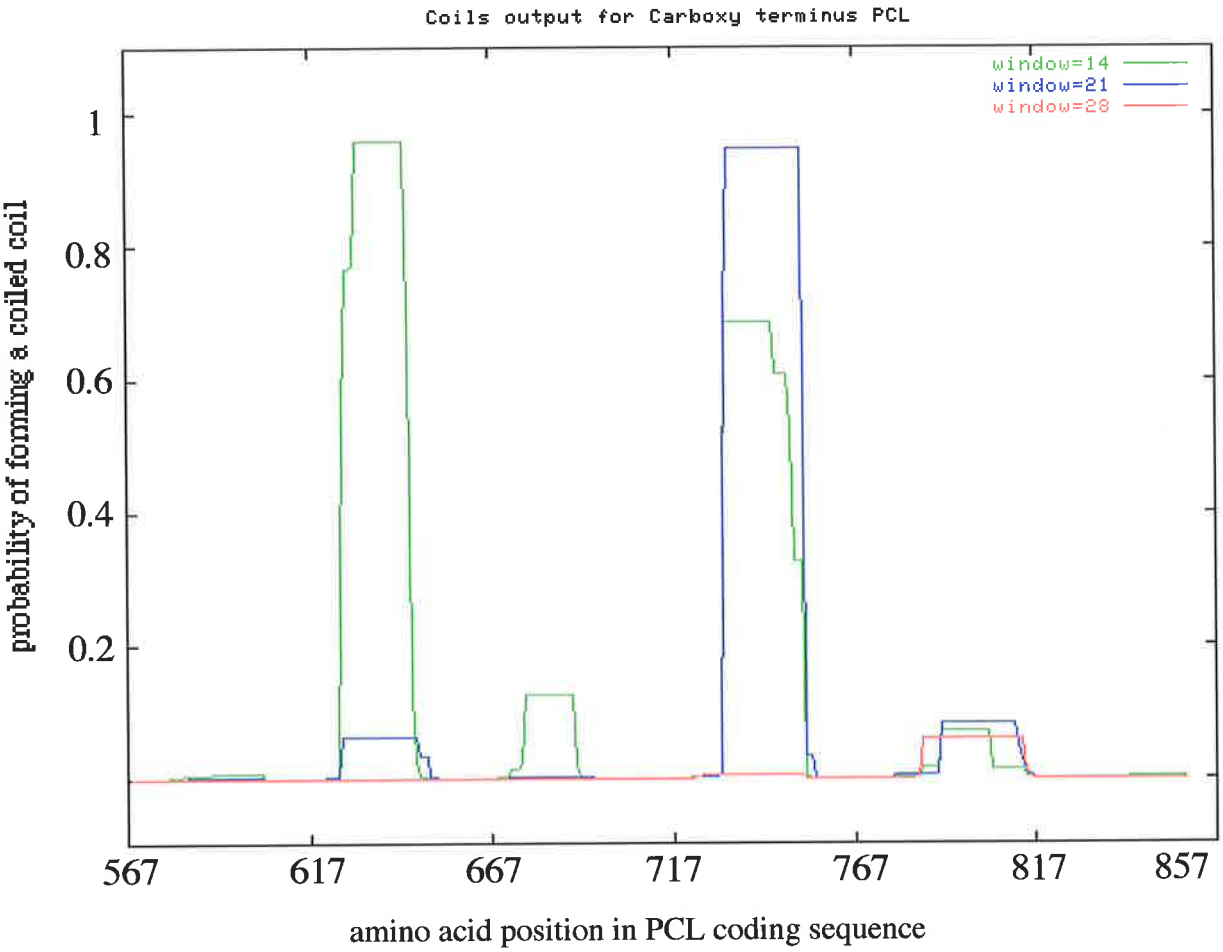


Figure 5.10: AD-SU(Z)2ΔB does not interact with the region of overlap between the LexA-$PCL_{634-800}$ and LexA-$PCL_{700-857}$ fragment.

To determine whether the region of overlap between the $PCL_{634-800}$ and $PCL_{700-857}$ fragment of PCL was responsible for the interaction between PCL and SU(Z)2, a LexA-$PCL_{700-800}$ (aa700-800) construct was generated. The GLU plate (A) was streaked first, followed by the GAL (B) and GLU LEU (C) plate to ensure that yeast was streaked onto each plate. (A) is a self-activation control and shows that LexA-$PCL_{700-800}$ is unable to self-activate reporter gene expression. (B) shows that LexA-$PCL_{700-800}$ is unable to interact with AD-SU(Z)2ΔB in this assay. (C) shows that yeast was streaked onto each plate.

Figure 5.11: The $PCL_{700-857}$ and $PCL_{634-800}$ fragments of the PCL carboxy terminus encode a putative coiled-coil motif.

The carboxy terminus of PCL was analysed using the COILS program to determine whether it encoded a coiled-coil motif. The COILS program produces a graphical output with the probability of forming a coiled-coil on the vertical axis and the amino acid position of the input protein on the horizontal axis (Lucas *et al.*, 1991). The prediction for the carboxy terminus of PCL is that there are two regions which have a high probability of forming a coiled-coil motif. Firstly a region in the $PCL_{567-700}$ fragment and secondly a region in the overlap between the $PCL_{634-800}$ and $PCL_{700-857}$ fragment. Below the graphical output of the COILS program is a schematic of the position of the $PCL_{567-700}$, $PCL_{634-800}$ and $PCL_{700-857}$ fragments with respect to the peaks on the graph. The amino acid positions of the fragments is also shown above the blue bars.


Coils output for Carboxy terminus PCL
window=14
window=21
window=28
probability of forming a coiled coil
1
0.8
0.6
0.4
0.2
567
617
667
717
767
817
857
amino acid position in PCL coding sequence

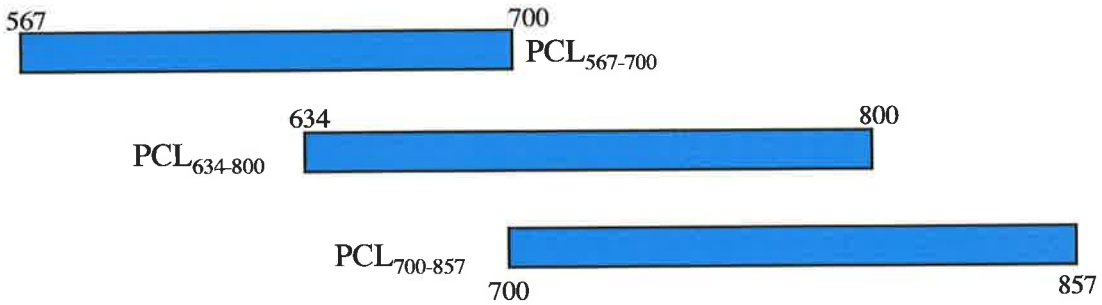
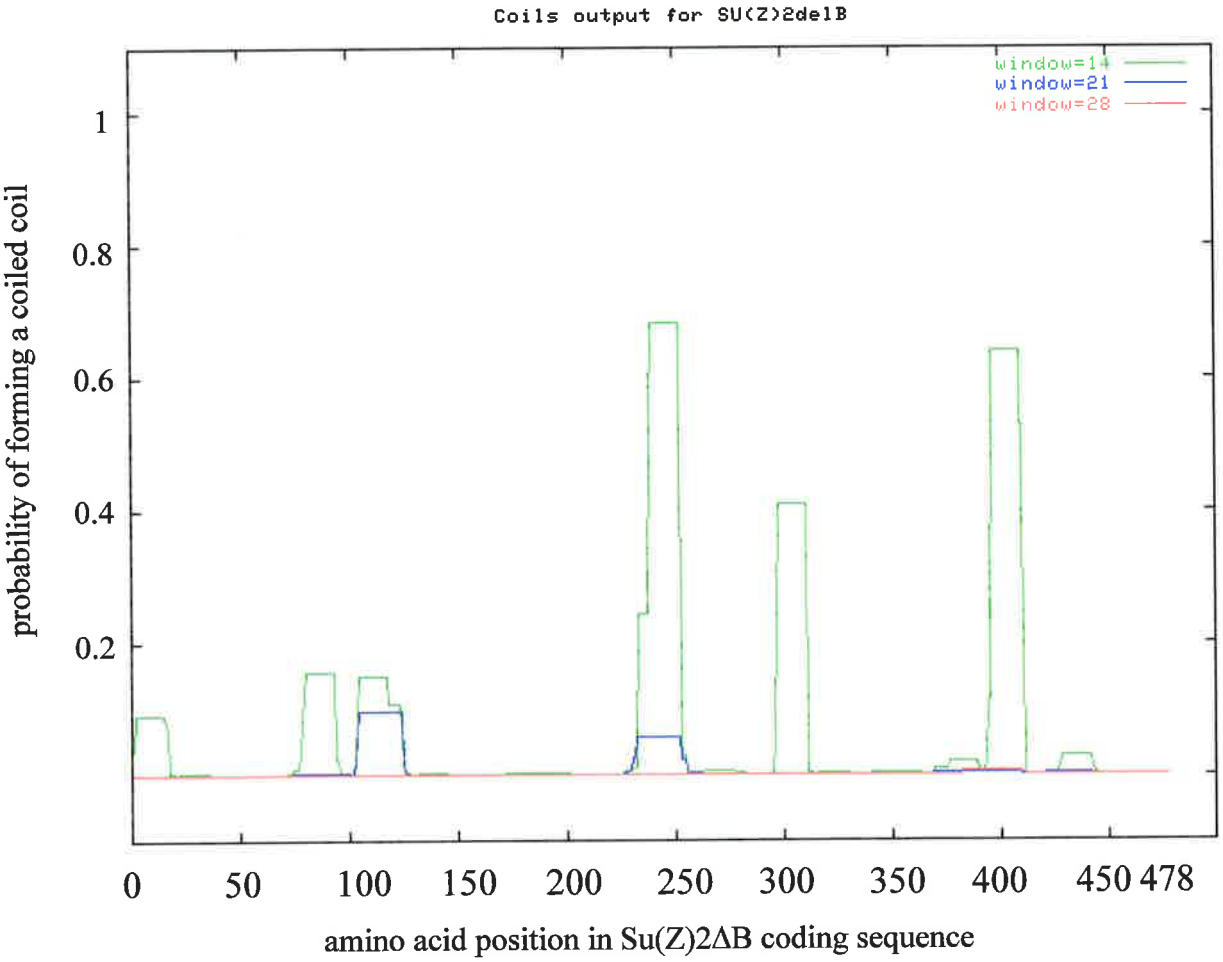
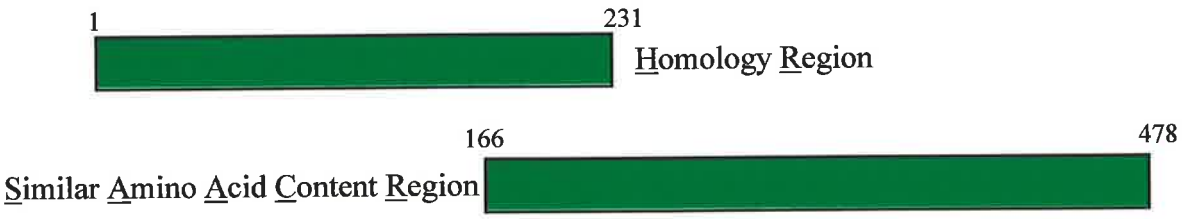

567
700
PCL567-700
634
800
PCL634-800
PCL700-857
700
857

Figure 5.12: The Similar Amino Acid Content Region of SU(Z)2ΔB encodes two potential coiled-coil motifs.

The SU(Z)2ΔB was analysed using the COILS program to determine whether it encoded for putative coiled-coil domain (Lucas *et al.*, 1991). The graphical output of the program shows that SU(Z)2ΔB has two regions which could encode for coiled-coils and a third region that has a reduced probability. All three putative coiled-coils are encoded by the SAACR of SU(Z)2ΔB. The position of the HR and SAACR are shown schematically below, together with the amino acid position of the fragments.


Coils output for SU(Z)2delB
window=14
window=21
window=28
probability of forming a coiled coil
1
0.8
0.6
0.4
0.2
0
50
100
150
200
250
300
350
400
450
478
amino acid position in Su(Z)2ΔB coding sequence


1
231
Homology Region
166
478
Similar Amino Acid Content Region

5-4.2 The SU(Z)2 interaction region

To determine which region of SU(Z)2 is responsible for the interaction with PCL, two overlapping AD-Su(Z)2ΔB constructs were generated in pJG4-5. These constructs were generated to include either the HR domain (AD-SU(Z)2-HR) (aa1-236) or the SAACR (AD-SU(Z)2-SAACR) (aa167-478). The AD-SU(Z)2-SAACR construct encoded the three putative coiled-coil domains predicted to be present in the SU(Z)2ΔB sequence. AD-SU(Z)2-HR and AD-SU(Z)2-SAACR were transformed into EGY48 and their expression confirmed using the anti-HA antisera (Figure 5.13). AD-SU(Z)2-HR and AD-SU(Z)2-SACCR were then tested for an interaction with LexA-PCL-C on leucine deficient media. Neither AD-SU(Z)2-HR or AD-SU(Z)2-SACCR were able to interact with LexA-PCL-C (Figure 5.13) in this assay suggesting that neither region of SU(Z)2ΔB alone is sufficient for the interaction with PCL.

Hugh Brock and colleagues have noticed that AD-SU(Z)2ΔB interacts with a wide range of PcG proteins in the yeast two hybrid assay (H. Brock, pers. comm.). To rule out general stickiness of the AD-SU(Z)2ΔB construct, five LexA fusion constructs; LexA-DBL, LexA-cyclinE, LexA-DRI-N, LexA-DRI-ARID and LexA-DRI-C were tested for an interaction with AD-SU(Z)2ΔB. LexA-DBL (a domain of the Pebble (Pbl) protein, which is involved in cytokinesis) and LexA-cyclinE (involved in the G1-S phase transition of the cell cycle) were found not to interact with AD-SU(Z)2ΔB when tested. Fragments of Dead ringer (Dri), a protein involved in early embryonic patterning, were able to weakly interact with AD-SU(Z)2ΔB (data not shown). Dri, unlike Pbl and CyclinE may be expected to interact with members of the PcG. In some developmental contexts Dri acts as a repressor and could potentially recruit the PcG protein to target genes (T. Shandala, pers. comm.), however it would be unlikely that the three non-overlapping regions of DRI would each be able to mediate an interaction with SU(Z)2ΔB. The interaction between PCL and SU(Z)2 is therefore highly questionable and further *in vitro* and *in vivo* analysis clearly needs to be performed to confirm the significance of the interaction between PCL and SU(Z)2.

5-5 Discussion

The previous two chapters provided evidence for the role of the cDOM and amino terminus in the function of PCL. In order to identify a function for the carboxy terminus of PCL, yeast two hybrid assays were performed using LexA-PCL-C and LexA-PCL-PHD+C together with all available PcG members. This analysis identified PHO and SU(Z)2ΔB as potential interactors of PCL.

5-5.1 The PCL/PHO interaction

PHO is the only member of the PcG which has been shown to encode a DNA binding domain (Brown *et al.,* 1998). The role of PHO in the recruitment of the PcG to target genes and the mechanism of repression of these target genes is unclear, but PHO presumably targets at least some PcG complexes to target genes via its interaction with DNA. PHO has not been shown to interact with any member of the PcG and no polytene antibody stainings have been reported to date. The identification of the interaction between PHO and PCL raises the possibility that PHO recruits the PcG via PCL. The region of PCL responsible for mediating the interaction with PHO was narrowed down to aa 800-857 but was unable to be further refined. This region of PCL shows no sequence similarity to any other protein currently present in the GENBANK database and no characterised motif has been identified in this region.

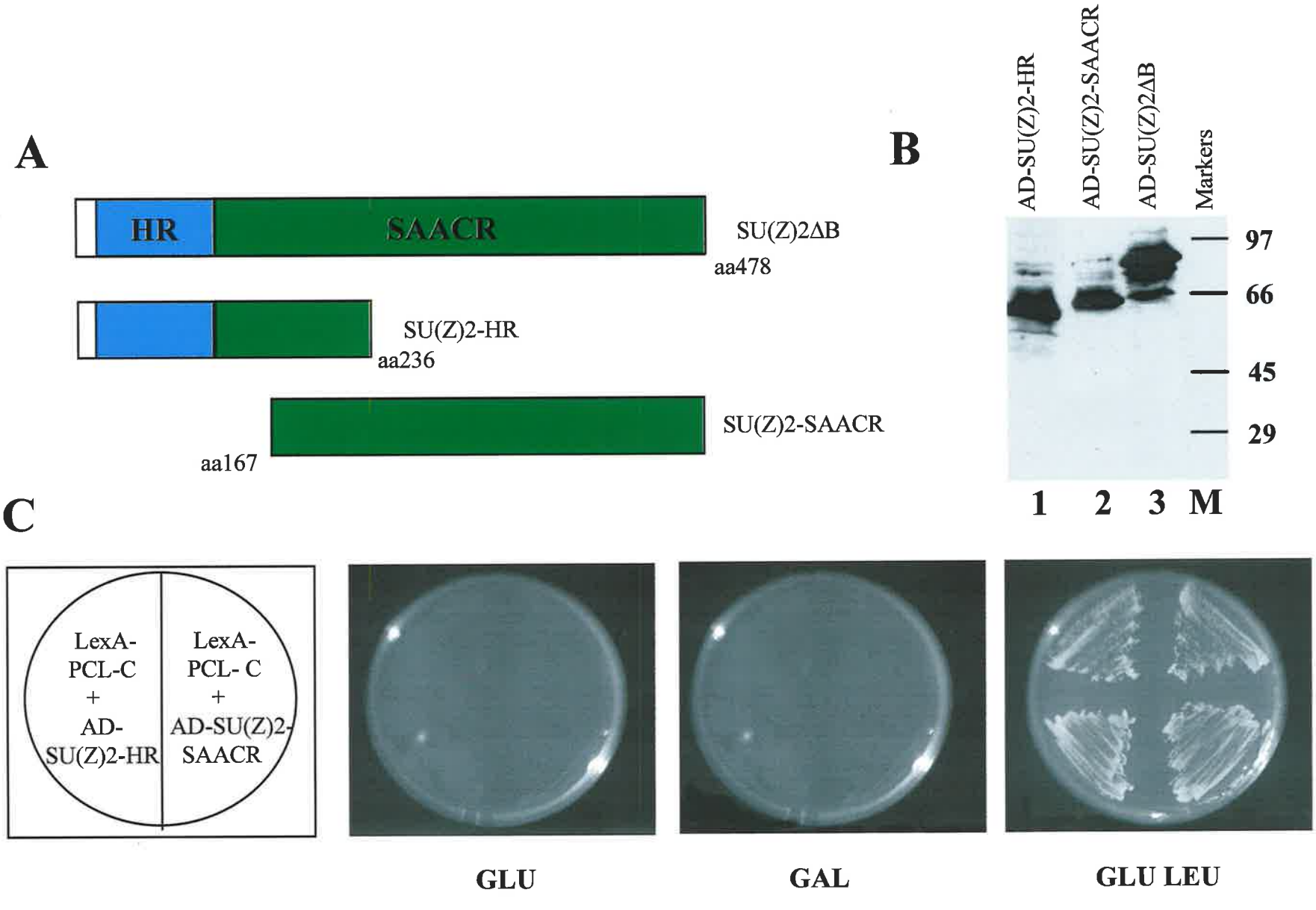
An attempt was made to narrow down the PCL interaction domain in PHO by generating overlapping amino and carboxy fragments of PHO. The region of the protein responsible for mediating repression in YY1, the mammalian homologue of PHO, is the Zn finger region. The high level of sequence conservation between PHO and YY1 across this region suggests the possibility that the Zn fingers in PHO could also be responsible for mediating repression of target genes. The repression mediated by the Zn fingers could be due to the interaction between PHO and PcG proteins such as PCL. The AD-PHO-Carboxy fragment encoded the entire Zn finger region. However, neither AD-PHO-Amino nor AD-PHO-Carboxy were able to interact with PCL, although both fusion proteins were being expressed. This suggests that PHO has a bipartite interaction domain or the region of overlap between the two fragments was not sufficient to prevent the disruption of the PCL interaction domain. Further work needs to be performed to confirm the two hybrid interaction between PCL and PHO. *In vitro* co-immunoprecipitations were attempted, but PCL interacted with the protein A/G beads and this problem was not overcome by incubating the beads in high concentrations of BSA (data not shown). Further *in vitro* experiments such as far western analysis or GST-pulldowns should be performed.

5-5.2 The interaction with SU(Z)2

SU(Z)2 is an unusual member of the PcG. Unlike other members of the PcG, alleles of *Su(Z)2* do not have a haplo-insufficient homeotic extra sex combs phenotype or dominantly enhance the homeotic phenotype of mutations in other PcG genes (Adler *et al.,* 1989). *Su(Z)2* alleles do not derepress homeotic gene expression in the embryo. The role of SU(Z)2 in the action of the PcG is not known, it does however share 53% of its polytene binding sites with PC/PH/PCL (Rastelli *et al.,* 1993), suggesting that SU(Z)2 functions with the PcG complex at many chromosomal sites.

Figure 5.13: Neither the HR nor the SAACR of SU(Z)2ΔB interact with the carboxy terminus of PCL.

To determine whether the putative coiled-coil motifs present in SU(Z)2ΔB are responsible for mediating the interaction with LexA-PCL-C, the region encoding the putative coiled-coils was cloned into the yeast two hybrid vector pJG4-5 to generate AD-SU(Z)2-SAACR. An overlapping AD-SU(Z)2-HR construct was also generated to to rule out the amino terminus as having a role in mediating the interaction with LexA-PCL-C. (A) is a schematic of the SU(Z)2ΔB coding sequence and also shows the positioning of the SU(Z)2-HR and SU(Z)2-SAACR fragments that were cloned into the pJG4-5 vector. (B) is a western blot probed with anti-HA antisera showing the expression of AD-SU(Z)2ΔB (lane 3), AD-SU(Z)-HR (lane 1) and AD-SU(Z)2-SAACR (lane 2) in EGY48. The predicted sizes for HA-HR and HA-SAACR are 38kD and 49kDa respectively. Both HA-HR and HA-SAACR migrated more slowly than expected for the predicted size, which could indicate post-translational modification. (C) shows that neither AD-SU(Z)2-HR nor AD-SU(Z)2-SAACR were able to interact with LexA-PCL-C in the two hybrid assay. Yeast were streaked sequentially onto GLU, GAL and GLU LEU plates.


A
HR
SAACR
SU(Z)2ΔB
aa478
SU(Z)2-HR
aa236
SU(Z)2-SAACR
aa167
B
AD-SU(Z)2-HR
AD-SU(Z)2-SAACR
AD-SU(Z)2ΔB
Markers
97
66
45
29
1
2
3
M
C
LexA-PCL-C + AD-SU(Z)2-HR
LexA-PCL- C + AD-SU(Z)2-SAACR
GLU
GAL
GLU LEU

This chapter raises the possibility that SU(Z)2 interacts with PCL, a 'true' member of the PcG. Two fragments of PCL were cloned into the LexA yeast two hybrid vector, pEG202 to generate LexA-PCL-PHD+C and LexA-PCL-C. Both constructs interacted with AD-SU(Z)2ΔB to allow growth of the yeast in the absence of exogenously supplied leucine, indicating activation of the *LEU2* reporter gene. When comparing the growth of yeast on GAL plates versus GLU plates the growth of yeast containing LexA-PCL-PHD+C and AD-SU(Z)2ΔB containing yeast was considerably weaker than the growth of yeast containing LexA-PCL-C and AD-SU(Z)2ΔB. The cause of this difference has not been determined but could be due to partial repression of the *LEU2* reporter gene by the PHD+C/SU(Z)2ΔB protein complex. Perhaps a yeast protein is present that interacts with the PHD+C protein and enables the formation of a repressive complex at the reporter gene site.

There is evidence to suggest that PcG proteins may be able to repress reporter gene expression in yeast (Kyba and Brock, 1998a). Full length PC is unable to interact with full length PSC in the yeast two hybrid assay. However when the proteins are broken up into smaller fragments and used in the two hybrid system an interaction can be observed (Kyba and Brock, 1998a). GST-pulldown assays with the full length proteins show that the full length proteins are capable of interacting (Kyba and Brock, 1998a). This ability to repress yeast transcription could also explain why full length SU(Z)2 was unable to interact with either LexA-PHD+C, LexA-PCL-C (this study) or LexA-PCL (Robert, 1997).

The region of PCL responsible for mediating the interaction with SU(Z)2 was narrowed down using the overlapping constructs LexA-$PCL_{567-700}$, LexA-$PCL_{634-800}$ and LexA-$PCL_{700-857}$. Both LexA-$PCL_{634-800}$ and LexA-$PCL_{700-857}$ interacted with AD-SU(Z)2ΔB, however the interaction between AD-SU(Z)2ΔB and LexA-$PCL_{634-800}$ was consistently weaker than the interaction with LexA-$PCL_{700-857}$. Whether this reflects a quantitative difference in the ability of these fragments to interact with AD-SU(Z)2ΔB is not known, however a general correlation has been observed between the strength of the interaction observed in the two hybrid assay and *in vitro* interaction assays (Estojak *et al.*, 1995). This difference in the strength of the interaction could be due to the disruption of a domain in the LexA-$PCL_{634-800}$ construct. The region of PCL encoded by LexA-$PCL_{700-857}$ has a high probability of encoding a coiled-coil motif. This region of PCL corresponds to the region of overlap between LexA-$PCL_{700-857}$ and LexA-$PCL_{634-800}$ (see Figure 5.11). The weaker interaction observed between LexA-$PCL_{634-800}$ and AD-SU(Z)2ΔB compared with the interaction between LexA-$PCL_{700-857}$ and AD-SU(Z)2ΔB may therefore be due to a requirement for stabilising sequences present in the C-terminus of the protein.

Sequences flanking coiled-coil motifs have been shown to be important in stabilising the structure of the coiled-coil and therefore in maintaining an interaction between two proteins (Pelletier *et al.*, 1997). As coiled-coils have only been shown to interact with other coiled-

coil motifs, SU(Z)2 was also analysed using the COILS program to determine whether it contained a coiled-coil. The SAACR of SU(Z)2ΔB has a strong probability of forming two coiled-coil motifs. However this region of SU(Z)2 was unable to interact with LexA-$PCL_{700-857}$. This could be due to the requirement of stabilising sequences in the N-terminus of SU(Z)2. Alternatively, the coiled-coil motifs of SU(Z)2ΔB may not be responsible for mediating the interaction between PCL and SU(Z)2. To further test this a careful mutational analysis needs to be performed in which key residues critical to the formation of the coiled-coil motifs (generally hydrophobic residues) in both PCL and SU(Z)2, would be mutated to determine the importance of these regions in the interaction between these two proteins.

Before further analysis is carried out to determine the domains in both SU(Z)2 and PCL important in mediating the interaction between the two proteins, *in vivo* confirmation of the interaction is required. As mentioned previously, Hugh Brock and colleagues have noticed that AD-SU(Z)2ΔB is able to interact with a wide variety of PcG proteins in the yeast two hybrid assay. AD-SU(Z)2ΔB was tested against five other non-PcG LexA-fusion proteins and was unable to interact with two of these proteins (LexA-Cyclin E and LexA-DBL). However AD-SU(Z)2ΔB was able to interact with three other fragments tested (LexA-DRI-N, LexA-DRI-ARID and LexA-DRI-C). Whether these interactions are of *in vivo* significance is not known, however both *in vitro* and *in vivo* evidence should be gathered before further work is performed on the interaction between PCL and SU(Z)2. Co-immunoprecipitation experiments could be used to confirm that the two proteins are associated in the same complex *in vivo*. Further *in vitro* analysis, such as far western analysis or "GST-pulldown" analysis could be carried out to ensure that the interaction between the two proteins is direct and not mediated by a yeast protein, and the interaction is 'real' and not an artefact of the two hybrid assay.

The work presented in this chapter raises the possibility that PCL interacts with PHO and SU(Z)2. If this was to be confirmed by *in vivo* and *in vitro* experiments, it would further support the growing body of evidence that members of the PcG form many protein/protein interactions. The next step is to translate what we know about the protein compositions of the group into how it is able to heritably repress gene expression.

Chapter 6: Discussion

6-1 Introduction

The broad aim of this thesis has been to investigate the molecular role of *Pcl* as a member of the PcG of genes. *Pcl* is an important member of the PcG. Genetically, *Pcl* alleles interact strongly with other members of the PcG (Campbell *et al.*, 1994; Jurgens, 1985) and *Pcl* mutants show derepression of *Ubx* and *AbdB* outside of their normal domains of expression (Lonie *et al.*, 1994; Soto *et al.*, 1995). Immunolocalisation studies on polytene chromosomes demonstrate that the binding of PCL is identical to that of PC and therefore PH and must therefore be overlapping with other members of the PcG (Lonie *et al.*, 1994). Less was known about the molecular role of PCL in the PcG. Yeast two hybrid analysis had identified an interaction between a conserved region of PCL and E(Z) (Robert, 1997). No function had been attributed to any other region of PCL. The aim of this thesis was to confirm, extend and characterise interactions between PCL and other PcG proteins.

6-2 Summary of results

6-2.1 PCL/E(Z) interaction

The association between PCL and E(Z) *in vivo* was confirmed by co-immunoprecipitation from cellular extracts. *In vitro* mutagenesis was used to demonstrate that the interaction between PCL and E(Z) is mediated through the two conserved PHD finger motifs. Prior to this work, it was not known whether the PHD fingers mediated protein-protein or DNA-protein interactions. It now appears, at least in the case of PCL, that they are responsible for mediating a protein-protein interaction between PCL and E(Z). The role of E(Z) as a member of the PcG and trxG, raised the possibility that the PHD fingers of PCL and TRX mediated the role of E(Z) as a member of the PcG or trxG respectively. This theory was tested using yeast two hybrid analysis but no interaction was detected between the PHD fingers of TRX and E(Z). This does not, however, rule out the possibility that the PHD fingers of PCL are required for the PcG-function of E(Z).

The role of the PCL/E(Z) interaction was investigated in a tethering assay. It had previously been shown that PCL, when tethered to DNA via a GAL4 DNA binding domain, was able to initiate heritable silencing of a reporter gene when linked to sequences from the *Ubx* promoter (T. McGrath pers. comm.). To determine whether the PHD fingers were important in initiating this heritable repression, they were fused to the GAL4 DNA BD to generate a *hb-GAL-PHD* construct. This construct was assayed for its ability to initiate heritable silencing of the *lacZ* reporter gene in the *BGUZ* construct and was unable to do so. It therefore appears that the interaction between the PHD fingers and E(Z) is not sufficient

for heritable silencing of a reporter gene. The alternative possibility, that the PHD fingers are not being expressed or folded properly, also needs to be examined. The expression of the GAL-PHD fusion protein could be examined using western analysis. The requirement of the PHD fingers could be further examined by mutating the PHD fingers in the *hb-GAL-PCL* construct and determining the effect on both the initiation and maintenance of repression.

6-2.2 The function of the amino terminus

The tethering assay described in Chapter 4 demonstrated that the amino terminus of PCL was responsible for initiating heritable repression and this repression was dependent on endogenous *Pcl* and *E(z)*. The ability of PCL and the amino terminus to initiate heritable repression of a reporter gene suggests that PCL is capable of interacting with endogenous PcG members. A homotypic interaction was identified between full length PCL and the amino terminus. Two other PcG proteins, PH and SCM, have also been demonstrated to form homotypic interactions (Peterson *et al.,* 1997). The role of these homotypic interactions have not been explored. However it can be proposed that these interactions will aid in the formation and stabilisation of PcG complexes at target genes. PcG complexes are thought to spread several kilobases from the site of repression (Strutt *et al.,* 1997) and this spreading could be facilitated by homotypic interactions that could mediate interactions between otherwise distinct PcG complexes. Evidence for a heterotypic interaction between PCL and PC was also presented, but remains to be confirmed. This heterotypic interaction could be important in the recruitment of the endogenous PcG by the GAL-Amino and GAL-PCL fusion proteins.

6-2.3 Carboxy terminus interactors

The tethering assay did not demonstrate a role for the C-terminus of PCL in the recruitment of the repression complex. Chapter 5, however, describes the interaction of the C-terminus of PCL with PHO and SU(Z)2. These interactions have to be confirmed *in vitro* and *in vivo,* but if they do occur, why was the carboxy terminus unable to initiate heritable repression in the tethering assay if it interacts with PHO and SU(Z)2?

Very little is known about the role of SU(Z)2 in the PcG. Mutations in *Su(z)2* do not result in a homeotic derepression phenotype, but instead act to enhance the phenotype of mutations in other PcG members. PHO is a recently identified DNA binding member of the PcG and has been shown to bind to PREs *in vitro* and *in vivo*. Although it binds to PRE sequences, it is not known whether PHO is involved in the initiation of heritable repression.

PHO appears to be required for PcG-mediated repression throughout development, perhaps through the anchoring of PcG complexes to DNA (Fritsch *et al.,* 1999), however PHO has

not been detected in PRC1 (Shao *et al.*, 1999). The role of PHO as a DNA binding domain protein involved in anchoring the PcG to DNA raises an interesting explanation for the failure of the GAL-Carboxy fusion protein to initiate heritable repression. Perhaps PHO initiates heritable repression through recruitment of the PcG complex via its interaction with the carboxy domain of PCL. The role of the carboxy terminus would then become redundant when it is tethered to DNA via the GAL4 DNA binding domain. The requirement for PHO in the tethering assay can be tested by performing the tethering experiment in a *pho* mutant background. If GAL-Amino and GAL-PCL are able to initiate heritable repression in a *pho* mutant, then the requirement for PHO is bypassed by tethering full length PCL and the amino terminus to DNA. The requirement for PHO in maintenance of repression could also be tested.

6-3 PcGi vs PcGm

Recently there has been some speculation about the formation of two PcG complexes (reviewed in van Lohuizen, 1999), a PcGi complex which initiates PcG-mediated silencing and a PcGm complex which maintains PcG-mediated silencing. The precise roles of these complexes is yet to be established. The PcGi complex is thought to contain ESC and E(Z). ESC is a unique member of the PcG which is only required early in embryogenesis to initiate PcG-mediated repression and plays no role in the maintenance of repression (Struhl, 1981; Struhl and Brower, 1982). Its association with E(Z) has raised the possibility that it functions with E(Z) in initiating PcG-mediated repression. A 560-600MDa complex containing at least E(Z) and ESC has been identified from *Drosophila* (Peter Harte pers. comm.). This complex also contains several other as yet, unidentified proteins. Whether PCL is a member of this complex is not yet known. The exact role of the PcGi in establishing PcG-mediated repression is not known. A link between the PcGi and the gap protein repressors, such as HB, is yet to be established but dMi-2 could be a member of the PcGi, and its interaction with HB targets PcGi to the *Ubx* gene.

The PcGm complex is thought to contain at least PC, PH, PSC, and SCM. These proteins are associated *in vivo* in a 2MDa complex which has been isolated and shown to inhibit SWI/SNF mediated nucleosomal remodelling (Shao *et al.*, 1999). PCL has not been detected in this complex, which would suggest that it is not a member of the PcGm. However, this would be inconsistent with the genetic and immunohistochemical evidence which suggests that PCL is important in maintaining PcG-mediated repression. The requirement for *Pcl* in maintaining repression of the *lacZ* reporter gene conferred by GAL-PCL and GAL-Amino in the tethering assay suggests that PCL is an integral member of the PcGm complex.

6-4 Where does PCL fit in?

Evidence is accumulating for three key steps in the initiation and maintenance of heritable silencing at a target locus (see Figure 6.1 for a schematic diagram).

The first of these is the recruitment of the PcGi. Evidence suggests that the PcG interacts with the gap repressor proteins and intermediary proteins such as dMi-2, which interacts with HB and PC. PHO, as a DNA binding protein could also be involved in recruiting the PcGi complex to target loci. PHO could be a member of the PcGi or it could act independently. The recruitment function of PHO could be mediated through the interaction with the carboxy terminus of PCL.

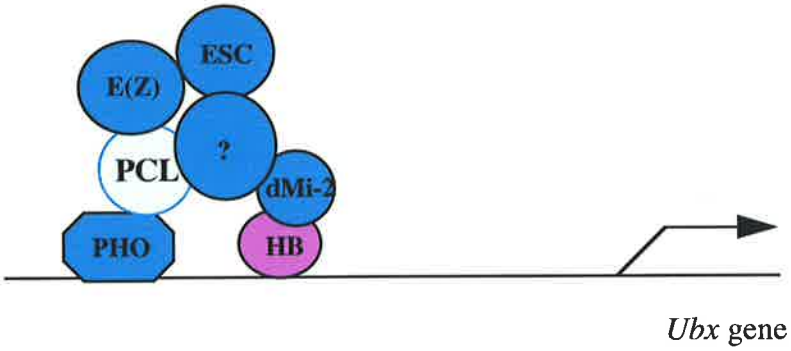
The second step is the recruitment of the PcGm. Although there appear to be two distinct PcG complexes in *Drosophila,* both of which are required for heritable repression, no direct link has been made between the PcGi and PcGm complexes. The interaction between PCL and E(Z), and the possible interaction between PCL and PC provides a potential link between the PcGi and the PcGm. Perhaps PCL is a member of both complexes. Its role in the PcGi complex would be to recruit the PcGm complex, which could be mediated through homotypic interactions with PCL in the PcGm and also heterotypic interactions with PC in the PcGm complex.

The third step is the maintenance of heritable silencing. Maintenance of repression involves two processes: actively preventing gene transcription during interphase and re-establishing repression after DNA replication. Neither process has been well characterised. All PcG members tested, with the exception of ESC, are required for maintenance of PcG-mediated repression and this repression is thought to be mediated through the formation of multimeric protein complexes such as PRC1. The role of PCL in this complex is not yet clear, although the interactions between PCL and other PcG proteins are likely to be crucial to the formation of a stable complex which is able to maintain PcG-mediated repression. PHO may be important in re-establishing the PcG on target loci after DNA replication and its association with PCL could play an important role in this process.

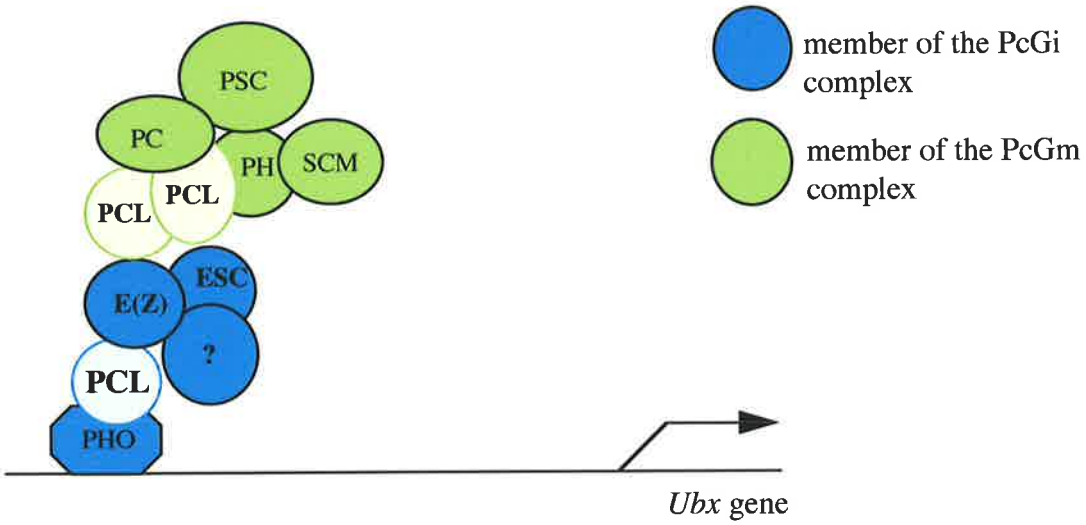
PCL is an essential member of the PcG. It makes contacts with E(Z) and may also make contacts with PHO, PC and SU(Z)2. Confirmation of the interactions with PHO, PC and SU(Z)2 would provide a direct link between the PcGi and PcGm complexes, a link that is likely to be important in the establishment and maintenance of PcG-mediated repression.

Figure 6.1: A schematic diagram of the possible role of PCL in the PcG complexes.
(i) The first step in initiating heritable silencing is the recruitment of the PcGi to target genes. HB, which initiates silencing of the *Ubx* gene has been shown to interact with dMi-2. *dMi-2* interacts genetically with PcG genes. PHO, as a DNA binding protein, may be involved in recruitment of the PcGi to target genes. PCL could be a member of the PcGi through its interaction with E(Z). Recruitment of the PcGi to target genes could therefore occur through the interaction between PCL and PHO.
(ii) The second step is the recruitment of the PcGm by the PcGi complex. PCL, as a member of the PcGi, could recruit the PcGm through its interaction with PC and also its homotypic interactions with PCL.
(iii) The third step in PcG-mediated silencing is the maintenance of silencing throughout cell division. As yet, it is unclear what the mechanism of maintenance is, although it is known that PCL, PC, PH, PSC, PHO, E(Z), SCM, MSX, SCE and CRM are required for maintenance of PcG-mediated repression.

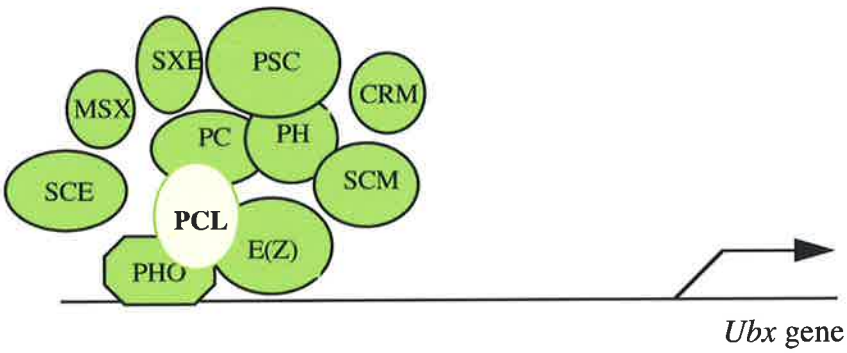
(i) targeting of the PcG



(ii) recruitment of the PcGm



(iii) maintenance of repression



References

Aasland, R., Gibson, T. and Stewart, A. (1995). The PHD finger: implications for chromatin-mediated transcriptional regulation. *Trends in Biochemical Sciences*, **20**; 56-59.

Abel, K.J., Brody, L.C., Valdes, J.M., Erdos, M.R., Mckinley, D.R., Castilla, L.H., Merajver, S.D., Couch, F.J., Friedman, L.S., Ostermeyer, E.A., Lynch, E.D., King, M-C., Welsch, P.L., Osborne-Lawrence, S., Spillman, M., Bowcock, A.M., Collins, F.S. and Weber, B.L. (1996). Characterisation of *EZH1*, a human homolog of *Drosophila Enhancer of zeste* near *BRCA1*. *Genomics*, **37**; 161-171.

Adler, P., Charlton, J., and Brunk, B. (1989). Genetic Interactions of the *Suppressor 2 of zeste* Region Genes. *Developmental Genetics*, **10**: 249-260.

Alkema, M., Jacobs, J. Voncken, J., Jenkins, N.A., Copeland, N.G., Satijn, D.P.E., et al. (1997). *Mpc2*, a new murine homologue of the *Drosophila* Polycomb protein is a member of the mouse Polycomb transcriptional repressor complex. *Journal of Molecular Biology*, **273**; 993-1003.

Alkema, M., Bronk, M., Verhoeven, E., Otte, A., van't Veer, L., Berns, A. and van Lohuizen, M. (1997b). Identification of Bmi-1-interacting proteins as constituents of a multimeric mammalian Polycomb complex. *Genes and Development*, **11**;226-240.

Altshul, F., Madden, T., Schaffer, A., Zhang, J, Zhang, Z., Miller, W. and Lipman, D. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acid Research*, **25**:3389-3402.

Austen, M., Luscher, B. and Luscher-Fuzlaff, J. (1997). Characterisation of the Transcriptional Regulator YY1. *Journal of Biological Chemistry*, **272**: 1709-1717.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1995). Current Protocols in Molecular Biology. Vol 1-3.

Bornemann, D., Miller, E. and Simon, J. (1996). The *Drosophila Polycomb* group gene *Sex combs on midleg (Scm)* encodes a zinc finger protein with similarity to polyhomeotic protein. *Development*, **122**; 1621-16-30.

Brand, A. and Perrimon, N. (1993). Targeted gene expression as a means of altering cells fates and generating dominant phenotypes. *Development*, **118**: 401-415.

Brannan, C. and Bartolomei, M. (1999). Mechanisms of genomic imprinting. *Current Opinions in Genetics and Development*, **9**;164-170.

Breen, T.R. and Duncan, I.M. (1986). Maternal expression of genes that regulate the bithorax complex of *Drosophila melanogaster*. *Developmental Biology*, **118**; 442-456.

Breen, T.R. and Harte, P.J. (1993). *trithorax* regulates multiple homeotic genes in the bithorax and Antennapedia complexes and exerts different tissue-specific, parasegment-specific and promoter specific effects on each. *Development*, **117**; 119-134.

Brunk, B., Martin, E. and Adler, P. (1991). Drosophila genes *Posterior sex combs* and *Supressor two of zeste* encode chromosomal proteins with homology to the murine *bmi-1* oncogene. *Nature*, **353**; 351-353.

Buchenau, P., Hodgson, J., Strutt, H. and Arndt-Jovin, D. (1998). The distribution of Polycomb-Group proteins during cell division and development in *Drosophila* embryos: Impact on models for silencing. *Journal of Cell Biology*, **141**; 469-481.

Buck, S. and Shore, D. (1995). Action of RAP1 carboxy-terminal silencing domain reveals an underlying competition between *HMR* andtelomeres in yeast. *Genes and Development*, **9**; 370-384.

Bushmeyer, S., Park, K. and Atchinson, M. (1995). Characterisation of Functional Domains within the Multifunctional Transcription Factor, YY1. *Journal of Biological Chemistry*, **270**: 30213-30220.

Campbell, R., Sinclair, D., Couling, M. and Brock, H.W. (1994). Genetic interactions and dosage effects of *Polycomb* group genes of *Drosophila*. *Molecular and General Genetics*, **246**; 291-300

Carrington, E. and Jones, R. (1996). The *Drosophila Enhancer of zeste* gene encodes a chromosomal protein: examination of wild-type and mutant protein distribution. *Development*, **122**:4073-4083.

Cavalli, G. and Paro, R. (1998). Chromo-domain proteins: linking chromatin structure to epigenetic regulation. *Current Opinion in Cell Biology*, **10**; 354-360.

Chen, H., Rossier, C. and Antonarakis, S. (1996). Cloning of a human homologue of the *Drosophila Enhancer of zeste* gene *(EZH2)* that maps to chromosome 21q22.2. *Genomics*, **38**; 30-37.

Cheng, N., Sinclair, D., Campbell, R. and Brock, H.W. (1994). Interactions of *polyhomeotic* with *Polycomb* Group genes of *Drosophila melanogaster*. *Genetics*, **138**; 1151-1162.

Cockell, M. and Gasser, S.M. (1999). Nuclear compartments and gene regulation. *Current Opinion in Genetics and Development*, **9**; 199-205.

Coulson M., Robert S., Eyre H.J., Saint R. (1998). The identification and localization of a human gene with sequence similarity to Polycomblike of Drosophila melanogaster. *Genomics* **48**;381-3

Csink, A. and Henikoff, S. (1996). Genetic modification of heterochromatic association and nuclear organisation in *Drosophila*. *Nature*, **381**; 529-531.

DeCamilis, M., Cheng, N., Pierre, D. and Brock, H.W. (1992). The *polyhomeotic* gene of *Drosophila* encodes a chromatin protein that shares polytene chromosomes binding sites with *Polycomb*. *Genes and Development*, **6**; 223-232.

Denisenko, O., Shnyreva, M., Suzuki, H. and Bomsztyk, K. (1998). Point mutations in the WD-40 domain of Eed blocks its interaction with Ezh2. *Molecular and Cellular Biology*,**18**; 5634-5642.

Dreesen, T., Henikoff, S. and Loughney, K. (1991). A pairing-sensitive element that mediates *trans*-inactivation is associated with the *Drosophila brown* gene. *Genes and Development*, **5**;331-340.

Duncan, I. (1982). *Polucomblike*: a gene that appears to be required for the normal expression of the bithorax and Antennepedia gene complexes of *Drosophila melanogaster*. *Genetics*, **102**; 49-70.

Dura, J-M., Brock, H.W., and Santamaria, P. (1985). *Polyhomeotic*; a gene in *Drosophila melanogaster* required for correct expression of segment identity. *Molecular and General Genetics*, **198**; 220-231.

Dura, J-M. and Ingham, P. (1988). Tissue- and stage-specific control of homeotic and segmentation gene expression in *Drosophila* embryos by *polyhomeotic*. *Development*, **103**; 733-741.

Estojak, J., Brent, R., and Golemis, E. (1995). Correlation of two hybrid affinity data with in vitro measurements. *Molecular and Cellular Biology*, **15**:5820-5829.

Fischer, J., Giniger, E., Maniatis, T. and Ptashne, M. (1988). GAL4 activates transcription in *Drosophila*. *Nature*, **332**; 853-856.

Franke, A., DeCamillis, A., Zink, D., Cheng, N., Brock, H. and Paro, R. (1992). *Polycomb* and *pleiohomeotic* are constituents of a multimeric protein complex in chromatin of *Drosophila melanogaster*. *EMBO J.*, **11**;2941-2950.

Fritsch, C., Brown, L., Kassis, J. and Muller, J. (1999). The DNA-binding Polycomb group protein Pleiohomeotic mediates silencing of a *Drosophila*homeotic gene. *Development*, **126**; 3905-3913.

Gatti, M. and Pimpinelli, S. (1992). Functional elements in *Drosophila melanogaster* heterochromatin. *Annual Review of Genetics*, **26**; 239-275.

Girton,J.R., and Jeon, S.H., (1994). Novel embryonic and adult homeotic phenotypes are produced by *pleiohomeotic* mutations in *Drosophila*. *Developmental Biology*, **161**; 393-407.

Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E. and Coupland, G. (1997). A Polycomb-group gene regulates homeotic gene expression in *Arabidopsis*. *Nature*, **386**; 44-51.

Gottschling, D., Aparacio, O., Billington, B. and Zakian, V. (1990). Position effect at *S.cerevisiae* telomeres: a reversible repression of polII transcription. *Cell*, **55**;27-39.

Gottschling, D. (1992). Telomere-proximal DNA in Saccharomyces cerevisiae is refractory to methyltransferase activity in vivo. *PNAS*, **89**; 4062-4065.

Grossniklaus, U., Vielle-Calzdada, J-P., Hoeppner, M. and Gagliano, W.B. (1998). Maternal control of embryogenesis by *MEDEA* a *Polycomb* group gene in *Arabidopsis*. *Science*, **280**; 446-450.

Gunster, M., Satijin, D., Hamer, K., den Blaauwen, J., de Bujin, D., Alkema, M., van Lohuizen, M., van Dreil, R. and Otte, A. (1997). Identification and characterisation of interactions between vertebrate Polycomb-Group protein BMI1 and human homologues of Polyhomeotic. *Molecular and Cellular Biology*, **17**; 2326-2335.

Gyuris, J., Golemis, E.A., Chertkov, H. and Brent, R. (1993). Cdi1, a human G1- and S-phase protein phosphatase associates with Cdk2. *Cell*, **75**; 791-803.

Harlow, E. and D. Lane (1988). *Antibodies: a laboratory manual.* New York, Cold Spring Harbour Laboratory.

Hartzof, G. and Winston, F. (1997). Nucleosomes and transcription: recent lessons from genetics. *Current Opinion in Genetics and Development*, **7**; 192-198.

Hecht A., Laroche T., Strahl-Bolsinger S., Gasser S.M. and Grunstein M. (1995). Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: a molecular model for the formation of heterochromatin in yeast. *Cell*, **80**; 583-92.

Henikoff S., Jackson J.M. and Talbert, P.B. (1995). Distance and pairing effects on the brownDominant heterochromaticelement in Drosophila. *Genetics* 1995, **140**; 1007-17

Ho, C-Y., Adamson, J.G., Hodges, R. and Smith, M. (1994). Heterodimerisation of the yeast MATa1 and MATα2 proteins is mediated by two leucine zipper-like coiled-coil motifs. *EMBO J.*, **13**; 1403-1413.

Hobert, O., Sures, I., Ciossek, T., Fuchs, M. and Ulrich, A. (1996). Isolation and developmental expression analysis of *Enx-1I,* a novel mouse Polycomb group gene. *Mechanism of Development*, **55**; 171-184.

Holdeman, R., Nerht, S. and Strome, S. (1998). MES-2, a maternal protein essential for viability of the germline in *Caenorhabditis elegans*, is homologous to a *Drosophila* Polycomb group protein. *Development*, **125**; 2457-2467.

Ingham, P. (1984). A gene that regulates the bithorax complex differently in larval and adult cells of *Drosophila* . *Cell*, **37**; 815-823.

Ingham, P.W. (1988). The molecular genetics of embryonic pattern formation in *Drosophila*.. *Nature*, **335**; 25-34.

Ishida, A., Asano, H., Hasegawa, M., Koseki, H., Ono, T., Yoshida, M., Taniguchi, M. and Kanno, M. (1993). Cloning and chromosome mapping of the human *Mel-18* gene which encodes a DNA-binding protein with a new 'RING-finger' motif. *Gene*, **129**; 249-255.

James, T. and Elgin, S. (1986). Identification of a nonhistone chromosomal protein associated with heterochromatin in *Drosophila* and its gene. *Molecular and Cellular Biology*, **6**; 3862-3872.

Jenuwein, T., Laible, G., Dorn, R. and Reuter, G. (1998). SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cellular and Molecular Life Sciences*, **54**; 80-93.

Jones, R. and Gelbart, W. (1990). Genetic analysis of the *Enhancer of zeste* locus and its role in gene regulation in *Drosophila melanogaster*. *Genetics*, **126**; 185-199.

Jones, R. and Gelbart, W. (1993). The *Drosophila* Polycomb-Group gene *Enhancer of zeste* shares a domain of sequence similarity with *trithorax*. *Molecular and Cellular Biology*, **13**; 6357-6366.

Jones, C., Ng, J., Peterson, A., Morgan, K., Simon, J. and Jones, R. (1998). The *Drsophila*esc and E(z) proteins are direct partners in Polycomb Group mediated repression. *Molecular and Cellullar Biology*, **18**; 2825-2834.

Jurgens, G. (1985). A group of genes controlling the spatial expression of the bithorax complex in *Drosophila*. *Nature*, **316**; 153-155.

Kehle, J., Beuchle, D., Treuheit, S., Christen, B., Kennison, J., Bienz, M. and Muller, J. dMi-2, a Hunchback-Interacting protein that functions in *Polycomb* repression. *Science*, **282**; 1897-1900.

Koonin EV, Zhou S, and Lucchesi JC (1995). The chromo superfamily: new members, duplication of the chromo domainand possible role in delivering transcription regulators to chromatin. *Nucleic Acids Research*, **23**; 4229-33

Korf, I., Fan, Y. and Strome, S. (1998). The Polycomb Group in *Canorhabditis elegans* and maternal control of germline development. *Development*, **125**; 2469-2478.

Kyba, M., and Brock, H. (1998a). The *Drosophila* Polycomb Group Proteins Psc Contacts ph and Pc through Specific Conserved Domains. *Molecular and Cellular Biology*, **18**; 2712-2720.

Kyba, M. and Brock, H. (1998b). The SAM domain of Polyhomeotic, RAE28, and Scm mediates specific interactions through conserved residues. *Developmental Genetics*, **22**; 74-78.

Labile, G., Wolf, A., Dorn, R., Reuter, G. , Noslow, C., Lebersorger, A., Popkin, D., Pillus, L. and Jenuwein, T. (1997). Mammalian homologues of the *Polycomb*-group gene *Enhancer of zeste* mediate silencing in *Drosophila* heterochromatin and at *S. cerevisiae* telomeres. *EMBO J.*, **16**; 3219-3232.

Latham, K. (1996). X-chromosome imprinting and inactivation in the early mammalian embryo. *Trends in Genetics*, **12**; 134-138.

Lennon, G.G., Auffray, C., Polymeropoulos, M. and Soares, M.B. (1995). The I.M.A.G.E. Consortium: An integrated Molecular Analysis of genomes and their Expression. *Genomics*, **33**, 151-152.

Lewis, E. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature*, **276**; 565-570.

Lonie, A., D'Andrea, R., Paro, R. and Saint, R. (1994). Molecular characterisation of the *Polycomblik* gene of *Drosophila melanogaster*, a trans-acting negative regulator of homeotic gene expression. *Development*, **120**; 2629-2636.

Lindsay, D.L. and Zimm, G.G. (1992). *In*, The genome of *Drosophila melanogaster*. Academic Press, San Diego.

Lucas, A., Van Dyke, M. and Stock, J. (1991). Predicting Coiled Coils from Protein Sequences. *Science*, **252**; 1162-1164.

Lustig, A. (1998). Mechanisms of silencing in *Saccharomyces cerevisiae*. *Current Opinion in Gentics and Development*, **8**; 233-239.

Martin, E. and Adler, P. (1993). The *Polycomb* group gene *Posterior Sex Combs* encodes a chromosomal protein. *Development*, **117**: 641-655.

McCall, K., O;Connor, M. and Bender, W. (1994). Enhancer traps in the bithorax complex mark parasegmental domains. *Genetics*, **138**; 387-399.

McCall, K. and Bender, W. (1996). Probes for chromatin accessibility in the *Drosophila* bithorax complex respond differently to *Polycomb*-mediated repression. *EMBO J.*, **15**; 569-580.

Messmer, S., Franke, A. and Paro, R. (1992). Analysis of the functional role of the *Polycomb* chromodomain in *Drosophila melanogaster*. *Genes and Development*, **6**; 1241-1254.

Moazad, D. and O'Farrell, P. (1992). Maintenance of the *engrailed* expression pattern by *Polycomb* group genes in *Drosophila*. *Development*, **116**; 805-810.

Montini, E., Buchner, G., Spalluto, C., Andolfi, G., Caruso, A., den Dunnen, J., Trump, D., Rocchi, M., Ballabio, A. and Franco, B. (1999). Identificatoin of SCML2, a second Human gene homologous to the *Drosophila Sex combs of midleg (Scm)*: a new gene cluster on Xp22. *Genomics*; **58**; 65-72.

Moore, G., Procunier, J., Cross, D. and Grigliatti, T. (1979). Histone gene deficiencies and position-effect variegation. *Nature*, **282**; 312-314.

Muller, J. and Bienz, M. (1991) Long range repression conferring boundaries of *Ultrabithorax*expression in the *Drsophila* embryo. *EMBO J.*, **10**; 3147-3155.

Muller, J. (1995). Transcriptional silencing by the Polycomb protein in *Drosophila* embryos. *EMBO J.*, **14**; 1209-1220.

Nibu, Y., Zhang, H. and Levine, M. (1998). Interaction of short-range repressors with *Drosophila* CtBP in the embryo. *Science*, **280**; 101-104.

Nomura, M., Takihara, Y. and Shimada, K. (1994). Isolation and characterisation of retinoic acid-inducible cDNA clones in F9 cells: one of the early inducible clones encodes a novel protein sharing several highly homologous regions with a *Drosophila polyhomeotic* proteins. *Differentiation*, **57**; 39-50.

Orlando, V. and Paro, R. (1993). Mapping Polycomb-repressed domains in the bithorax complex using *in vivo* formaldehyde cross-linked chromatin. *Cell*, **75**; 1187-1198.

Paro, R. and Hogness, D. (1991). The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. *PNAS*, **88**;263-267.

Pazin, and Kadonaga, J. (1997) What's up and down with histone deacetylation and transcription. *Cell*, **89**; 325-328.

Pearce, J., Singh, P. and Gaunt, S. (1992). The mouse has a *Polycomb*-like chromobox gene. *Development*, **114**; 921-929.

Pelletier, A., Do, F., Brisebois, J., Lagace, L. and Cordingley, M. (1997). Self-association of herpes simplex virus type 1 ICP35 is via coiled-coil interactions and promotes stable interactions with the major capsid protein. *Journal of Virology*, **71**; 5197-5208.

Petetson, A., Kyba, M., Bornemann, D., Morgan, K., Brock, H.W. and Simon, J. (1997). A domain shared by the *Polycomb* Group Proteins Scm and ph mediates Heterotypic and Homotypic interactions. *Molecular and Cellular Biology*,**17**; 6683-6692.

Pirrotta, V. (1995). Chromatin complexes regulating gene expression in *Drosophila*. *Current Opinion in Genetics and Development*, **5**;466-472.

Pirrotta, V. (1997). PcG complexes and chromatin silencing. *Current Opinion in Genetics and Development*, **7**; 249-258.

Platero, J., Hartnett, T. and Eissenberg, J.C. (1995). Functional analysis of the chromo domain of HP1. *EMBO J.*, **14**; 3977-3986.

Poortinga, G., Watanabe, M. and Parkhurst, S. (1998). *Drosophila* CtBP; a Hairy interacting protein required fr embryonic segmentation and hairy mediated transcriptional repression. *EMBO J.*, **17**; 2067-2078.

Qian, S., Capovilla, M. and Pirrotta, V. (1991). The *bx* region enhancer, a distant cis-control element of the *Drosophila Ubx* gene and its regulation by *hunchback*and other segmentation genes. *EMBO J.*, **10**; 1415-1425.

Qian, S., Capovilla, M. and Pirrotta, V. (1993). Molecular mechanisms of pattern formation by the BRE enhancer of the *Ubx* gene. *EMBO J.*, **12**; 3865-3877.

Rastelli, L., Chan, C.S., and Pirrotta, V., (1993). Related chromosome binding sites for *zeste*, *supressor of zeste* and *Polycomb* Group proteins in *Drosophila* and their dependence on *Enhancer of zeste* function. *EMBO J.*, **12**; 1513-1522.

Reijen, M., Hamer, K., den Blaauwen, J., Lambrechts, C., Schoveld, R., van Driel, R. and Otte, A. (1995). *Polycomb* and *bmi-1* homologs are expressed in overlapping patterns in *Xenopus* embryos and are able to interact with each other. *Mechanisms of Development*, **53**;35-46.

Renauld, H., Aparicio, O., Zierath, P., Billington, B., Chhablani, S. and Gottschling, D. (1993). Silent domains are assembled continuously from the telomere and are defined by promoter distance and strength and by *SIR3* dosage. *Genes and Development*, **7**; 1133-1145.

Roberts, S. (1997). Characterisation of Polycomblike and an interactor. Ph.D. thesis, University of Adelaide, Adelaide, Australia.

Rundlett SE, Carmen AA, Kobayashi R, Bavykin S, Turner BM, Grunstein M (1996). HDA1 and RPD3 are members of distinct histone deacetylase complexes that regulate silencing and transcription. *PNAS*, **93**;14503-8.

Santamaria, P. and Randsholt, N. (1995). Characterisation of a region of te X chromosome of *Drosophila* including *multi sex combs* (*mxc*), a *Polycomb* group gene which also functions as a tumour supressor. *Molecular and General Genetics*, **246**; 282-290.

Sato, T., Hayes, P.and Denell, R. (1984). Homeosis in Drosophila: Maternal effect of Enhancer of Polycomb locus and its interaction with Polycomb and related loci. *Developmental Genetics*, **4**; 185-198.

Satijn, D., Olson, D., van der Vlag, J., Hamer, K., Lambrechts, C., Masselink, H., Gunster, M., Sewalt, R., van Driel, R. and Otte, A. (1997). Intereference with the expression of a novel human Polycomb proteins, hPc2, results in cellular transformation and apoptosis. *Molecular and Cellular Biology*, **17**; 6076-6086.

Saurin, A., Shiels, C., Williamson, J., Satjin, D., Otte, A., Sheer, D. and Freemont, P. (1998). The human Polycomg Group complex associates with Pericentromeric Heterochromatin to form a novel nuclear domain. *Journal of Cell Biology*, **142**; 887-898.

Schloβherr, J., Eggert, H. Paro, R., Cremer, S. and Jack, R. (1994). Gene inactivation in *Drosophila* mediated by the *Poycomb* gene product or by position effect variegation does not involved major changes in the accessibility of the chromatin fibre. *Molecular and General Genetics*, **243**; 453-462.

Schumacher, A., Faust, C. and Magnuson, T. (1996). Positional cloning of a global regulator of anterior-posterior patterning in mice. *Nature*, **383**; 250-253.

Sedkov, Y., Tillib, S., Mizrokhi, L. and Mazo, A. (1994). The bithorax complex is regulated earlier by *trithorax* during *Drosophila* embryogenesis than is the Antennapedia complex, correlating with a bithorax-like expression pattern of distinct early *trithorax* transcripts. *Development*, **120**; 1907-1917.

Seeling, H., Moosbruger, I, Ehrfeld, H., Fink, T., Renz, M. and Genth, E. (1995). The major dermatomysis-specific autoantigen is a presumed helicase involved in transcriptional activation. *Arthritis andRheumotology*, **38**; 1389-1399.

Shi, Y., Lee, J-S. and Galvin, K. (1997). Everything you ever wanted to know about Yin Yang 1. *Biochimica et Biophysica Acta* 1332 F49-F66.

Sidow, A. (1996). Gen(om)e duplications in the evolution of early vertebrates. *Current Opinion in Genetics and Development*, **6**; 715-722.

Simon, J., Bornemann, D., Lunde, K.and Schwartz, C. (1995). The *extra sex combs* product contains WD40 repeats and its time of action implies a role distinct from other *Polycomb* group products. *Mechanisms of Development*, **53**; 197-208.

Simon, J. (1995). Locking in stable states of gene expression: transcriptional control during *Drosophila* development. *Current Biology*, **7**; 376-385.

Sinclair, D., Milne, T., Hodgson, J., Shellard, J., Salinas, C., Kyba, M., Randazzo, F. and Brock, H.W. (1998). The *Additional sex combs* gene of *Drosophila* encodes a chromatin protein that binds to shared and unique Polycomb group sites on polytene chromosomes. *Development*, **125**; 1207-1216.

Stankunas, K., Berger, J., Ruse, C., Sinclair, D., Randazzo, F. and Brock, H.W. (1998). The *Enhancer of Polycomb* gene of *Drosophila* encodes a chromatin protein conserved in yeast and mammals. *Development*, **125**;4055-4066.

Strouboulis, J., Damjanovski, S., Vermaak, D., Meric, F. and Wolffe, A. (1999). Transcriptional repression by Xpc1, a new Polycomb homologue in *Xenopus laevis* embryos, is independent of histone deacetylase. *Molecular and Cellular Biology*, **19**; 3958-3968.

Struhl, G. (1981). A gene product required for correct initiation of segmental determination in *Drosophila*. *Nature*, **293**;36-41.

Struhl, G. and Brower, D. (1982). Early role of the *esc+* gene product in the determination of segments in *Drsophila*. *Cell*, **31**; 285-292.

Strutt, H. and Paro, R. (1997). The Polycomb group protein complex of *Drosophila melanogaster* has different compositions at different target genes. *Molecular and Cellular Biology*, **17**; 6773-6783.

Tie, F., Furuyama, T. and Hatre, P.J. (1998). The *Drosophila* Polycomb Group proteins ESC and E(Z) bind directly to each other and co-localize at multiple chromosomal sites. *Development*, **125**; 3483-3496.

van de Vosse, E., Walpole, S., Nicolaou, A., van der Brent, P., Cahn, A., Vaudin, M., Ross, M., Durham, J., Pavitt, R., Wilkinson, J., Grafham, D., Bergen, A., van Ommen, G-J., Yates, J., den Dunnem, J. and Trump, D. (1998). Characterisation of *SCML1*, a new gene in Xp22, with homology to developmental Polycomb genes. *Genomics*, **49**; 96-102.

van Lohuizen, M., Frasch, M., Wientijens, E. and Berns, A. (1991). Sequence similarit between the mammalian *bmi-1* proto-oncogene and the *Drosophila* regulatory genes *Psc* and *Su(z)2*. *Nature*, **353**; 353-355.

van Lohuizen, M. (1998). Functional analysis of mouse Polycomb group genes. *Cell and Molecular Life Sciences*, **54**; 71-79.

van Lohuizen M., Tijms M., Voncken J.W., Schumacher A., Magnuson T., Wientjens E .(1998). Interaction of mouse polycomb-group (Pc-G) proteins Enx1 and Enx2 with Eed: indication for separate Pc-G complexes. *Mol. Cell. Biol.*, **18**; 3572-9

van Lohuizen, M. (1999). The trithorax-group and Polycomb-Group chromatin modifiers: implications for disease. *Current Opinion in Genetics and Development*, **9**: 355-361.

Wakimoto, B. (1998). Beyond the Nucleosome: Epigenetic aspects of Position-Effect Variegation in *Drosophila*. *Cell*, **93**; 321-324.

Weigel, D. (1997). Flower development: Repressing reproduction. *Current Biology*, **7**; R373-R375.

Yamamoto, Y., Girard, F., Bello, B., Affolter, M. and Gehring, W. (1997). The *cramped* gene of *Drosophila* is a member of the *Polycomb* group, and interacts with *mus209*, the gene encoding Proliferating Cell Nuclear Antigen. *Development*, **124**; 3385-3394.

Zhang, C-C., Muller, J., Hoch, M., Jackle, H. and Bienz, M. (1991). Target sequences for *hunchback* in a control region conferring *Ultrabithorax* expression boundaries. *Development*, **113**; 1171-1179.

Zhang, C-C. and Bienz, M. (1992). Segmental determintion in *Drosophila* conferred by hunchback (*hb*), a repressor of the homeotic gene Ultrabithorax (*Ubx*). *PNAS*, **89**; 7511-7515.

Zhang, Y., LeRoy, G., Seeling, H-P., Lane, W. and Reinberg, D. (1998). The dermatomysis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodelling activity. *Cell*, **95**; 279-289.

Zink, B., Engstrom, Y., Gehring, W. and Paro, R. (1991). Direct interactions of the *Polycomb* protein with *Antennapedia*regulatory sequences in poytene chromosomes of *Drosophila melanogaster*. *EMBO J.*, **10**; 153-162.

Zink, B. and Paro, R. (1995). *Drosophila Polycomb*-group regulated chromatin inhibits the accessibility of a *trans*-activator to its target DNA. *EMBO J.*, **14**; 5660-5671.