

Ref. Only

UNIVERSITY OF LONDON THESIS

Degree PhD Year 1997 Name of Author DAVIES, Joanna Marie

A.

This copy has been deposited in the Library of Institute of Cancer Research
(not BpMF)

This copy has been deposited in the University of London Library, Senate House, Malet Street, London WC1E 7HU.

LIBRARY
INSTITUTE OF CANCER RESEARCH
FULHAM RD.

The Identification of Novel POZ Domain Zinc Finger Proteins: Characterisation of a Heterodimeric Partner of BCL-6

Joanna Marie Davies

Submitted for the degree of PhD to the University of London

May 1997

Leukaemia Research Fund Centre
Institute of Cancer Research
London

Abstract

POK proteins are characterised by the presence of a highly conserved N-terminal POZ domain, which can mediate protein interactions and transcriptional repression, as well as a C-terminal *Krüppel*-related C₂-H₂ zinc finger region. The POK protein PLZF was first identified as a result of the molecular characterisation of a rare t(11;17) translocation associated with acute promyelocytic leukaemia. An additional POK family member of interest is BCL-6 which is frequently translocated and mutated in diffuse large cell lymphoma. To contribute towards an understanding of the molecular function of PLZF, the aim of this study was to identify additional, PLZF-related, POK proteins. Of the four POZ domain encoding genes cloned, the most closely related to PLZF, termed LRF, localises to human chromosome 19p13.3 and encodes a novel 60 Kdalton POK protein. Characterisation of the chicken, mouse and human homologues of LRF has demonstrated that, in analogy to other POK proteins, both its POZ domain and zinc fingers have been highly conserved through evolution. Northern blot and insitu hybridisation analyses have revealed a wide spread pattern of LRF expression. As detected for PLZF, at 10.0 days post coitus, LRF was strongly expressed in the developing limb bud with its levels of expression increasing along the proximo-distal axis. The yeast two-hybrid assay was used to demonstrate an interaction of LRF with itself and surprisingly, with the PLZF interacting BCL-6 protein but not with PLZF itself. This in vivo interaction between LRF and BCL-6 was confirmed by co-immunoprecipitation. Deletion analysis has revealed that in contrast to homodimerisation, which requires only the POZ domain, the heterodimeric interaction between BCL-6 and LRF is mediated through both the POZ domain and the *Krüppel*-like zinc fingers. In addition, a unique and potentially regulatory, intramolecular interaction between the POZ domain and zinc fingers of LRF has been detected. Whilst LRF is co-expressed with BCL-6 in the lymphoid compartment and co-localises with BCL-6 in the nucleus, preliminary co-transfection experiments have revealed no effect of LRF on BCL-6 mediated transcriptional repression. However, a general, potentially POZ domain dependant, repressive effect of LRF on transcription has been demonstrated. These studies suggest that POK proteins may function as homo- and heterodimeric complexes, their formation and resultant effect being determined by the levels and domains of expression of the partner proteins.

Acknowledgements

This PhD was funded by the National Institute of Health (USA). Support for purchase of materials was received from the Leukaemia Research Fund (Great Britain).

I am grateful to my supervisor Dr Arthur Zelent and all my colleagues and friends within the Institute for their guidance and support.

<u>Table of contents</u>	<u>Page</u>
<u>Abstract</u>	ii
<u>Acknowledgments</u>	iii
<u>Table of Contents</u>	iv
<u>List of Tables</u>	vi
<u>List of Figures</u>	vi
<u>Abbreviations</u>	viii
<u>Chapter 1: Introduction</u>	<u>1-28</u>
1.1 Regulation of gene transcription	1-17
1.1.1 General transcription factors	2
1.1.2 Regulatory transcription factors	3
1.1.3 Developmental transcriptional regulation	6
1.1.4 POZ domain transcriptional regulators	8
1.2 Leukaemias and lymphomas	17-27
1.2.1 Type I translocations of the RAR α locus in APL	19
1.2.2 Type II translocations and mutations of the BCL6 gene in NHL	24
1.3 The identification of regulatory heterodimeric POK protein complexes	27-28
<u>Chapter 2: Materials and Methods</u>	<u>29-62</u>
2.1 The identification of PLZF related POZ domain proteins	29-43
2.1.1 Zoo blot procedure	29
2.1.2 cDNA phage library screen	32
2.1.3 Subcloning phage DNA into Bluescript	34
2.1.4 POZ protein isolation by degenerate RT-PCR	37
2.1.5 Whole mount in-situ hybridisation	40
2.2 A novel POK protein: LRF	43-50
2.2.1 High stringency screen of mouse λ ZapII library	43
2.2.2 Isolation of genomic LRF clones	44
2.2.3 mLRF Zooblot	46

2.2.4 Northern blot analysis	46
2.2.5 Chromosomal localisation of LRF	49
2.3 The characterisation of a heterodimeric partner of LRF	50-62
2.3.1 Yeast two-hybrid assay	50
2.3.2 Co-immunoprecipitation	53
2.3.3 Co-immunofluorescence	55
2.3.4 Luciferase reporter assay	56
2.3.5 Construct design	58
<u>Chapter 3: Results</u>	<u>63-96</u>
3.1 The identification of PLZF related POZ domain proteins	63-73
3.1.1 The cloning of four POZ domain proteins	63
3.1.2 Embryonic expression of clones 1-4 at 10 dpc.	70
3.1.3 Clone 1 is the closest homologue of PLZF	72
3.2 A novel POK protein: LRF	74-86
3.2.1 LRF structure	74
3.2.2 Evolutionary conservation	78
3.2.3 Northern blot analysis	82
3.2.4 Chromosomal localisation	85
3.3 The characterisation of a heterodimeric partner of LRF	87-96
3.3.1 Identification of a heterodimeric partner for LRF	87
3.3.1.1 Yeast two-hybrid	87
3.3.1.2 Co-immunoprecipitation	90
3.3.1.3 Co-immunofluorescence	92
3.3.2 The functional consequences of a BCL-6/LRF interaction	92
3.4 Summary of results	95
<u>Chapter 4: Discussion</u>	<u>97-106</u>
4.1 PLZF related POZ domain proteins	97
4.2 Chromosomal localisation of LRF	98
4.3 A biologically significant POK protein heterodimerisation	99
4.4 Functional significance of a BCL-6/LRF interaction	101
4.5 Transcriptional regulation by POK protein dimerisation: a model	104
4.6 In conclusion	106
<u>References</u>	<u>107-119</u>

List of Tables **page**

Table 1	POZ domain protein, class I.	11
Table 2	Yeast two-hybrid controls.	89

List of Figures **page**

Figure 1	Alignment of the entire POZ domain of mPLZF and 19 other POZ domain protein regions selected from the data base.	8-9
Figure 2	A schematic representation of wildtype PLZF and RAR α and the APL associated PLZF-RAR α and RAR α -PLZF chimeras.	22
Figure 3	Zooblot analysis of mPLZF.	64
Figure 4	Deduced amino acid sequence of the mPLZF B protein and its conservation.	65
Figure 5	Chicken clone 1 (LRF) nucleotide and predicted amino acid sequences.	67-68
Figure 6	Nucleotide and predicted amino acid sequences of mouse clones 2-4.	69
Figure 7	In situ hybridisation of mouse clones 1-4 at 10dpc.	71
Figure 8	Alignment of the N terminal half of the POZ domain of mPLZF with mouse clones 1-4.	73
Figure 9	Nucleotide and predicted amino acid sequence of full length LRF.	75-76
Figure 10	Alignment of 5' LRF isoforms and genomic mLRF.	77
Figure 11	Zooblot analysis of mLRF	79
Figure 12	Evolutionary sequence conservation of the LRF POZ domain and zinc finger regions.	80-81
Figure 13	Northern blot expression of mouse LRF.	83

Figure 14	Northern blot expression of human LRF.	84
Figure 15	The chromosomal localisation of human and mouse LRF.	86
Figure 16	Mutational analysis of the BCL-6/LRF interaction by the yeast two-hybrid assay.	88
Figure 17	Co-immunoprecipitation of the BCL-6 and LRF proteins.	91
Figure 18	Nuclear co-localisation of the BCL-6 and LRF proteins	93
Figure 19	The effect of mLRF on BCL-6 mediated transcriptional repression.	94
Figure 20	A model of POK protein mediated transcriptional regulation.	105

Abbreviations

AGM	Aorta, gonad, mesonephros
ALL	Acute lymphoblastic leukaemia
AML	Acute myeloid leukaemia
Ap	Alkaline phosphatase
APL	Acute promyelocytic leukaemia
ATRA	All-trans retinoic acid
Bab	Bric a brac
BCIP	5 bromo-4 chloro 3 indolyl phosphate
BCL-6	B cell lymphoma-6
BMP	Bone morphogenic protein
BsK	Bluescript plasmid
c	Chicken
CHO	Chinese hamster ovarian
CML	Chronic myelogenous leukaemia
CTD	C-terminal domain
DEPC	Diethyl pyrocarbonate
DLCL	Diffuse large cell lymphoma
EKLF	Erythroid <i>Krüppel</i> -like factor
EtOH	Ethanol
FGF	Fibroblast growth factor
FITC	Flourescein isothiocyanate
FL	Follicular lymphoma
GTF's	General transcription factors
h	Human
HPLC	High purification liquid chromatography
hr	Hour(s)
Kr	<i>Krüppel</i>
KRAB	<i>Krüppel</i> associated box
LRF	Leukaemia/lymphoma related factor
m	Mouse
min	Minute(s)
NBT	Nitroblue tetrazolium salt
NCoR	Nuclear co-repressor
NEB	New England Biolabs
NHL	Non-Hodgin's lymphoma
NPM	Nucleophosmin
NuMa	Nuclear mitotic apparatus
o/n	Over night
ONPG	O-nitrophenyl- β -D-galactopyranoside
PCR	Polymerase chain reaction
PEV	Position effect variegation
PIC	Pre-initiation complex
PLZF	Promyelocytic leukaemia zinc finger
PODs	PML oncogenic domains
POK	POZ and <i>Krüppel</i>

POZ	POX virus and zinc fingers
preB	Precursor B cells
proB	Progenitor B cells
RA	Retinoids
RAR	Retinoic acid receptor
RE	Response element
rm temp	Room temperature
rpm	Revolutions per minute
RT	Reverse transcription
RXR	Retinoic X receptors
sec	Second(s)
SHH	Sonic hedgehog
ssDNA	Salmon sperm DNA
TAFs	TBP associated factors
TBP	TATA binding polypeptide
TRITC	Tetra methyl rodamine
U	Units
UV	Ultra violet
Zn	Zinc
(-/-)	Knock out

Chapter 1: Introduction.

The POZ (**P**ox virus and **Z**inc Finger) domain, is a newly recognised protein interaction motif present in a number of potential transcriptional regulators. Two POZ domain proteins, PLZF (**P**romyelocytic Leukaemia **Z**inc Finger) and BCL-6 (**B** Cell Lymphoma-6), have been implicated in oncogenesis. In order to gain a better understanding of the molecular and biological functions of the POZ domain proteins, either in oncogenesis or in the control of growth and differentiation, the aim of this study was to identify additional members of this gene family. Section 1.1, of this introduction, reviews the potential function of POZ domain proteins in light of the current understanding of transcriptional regulation whilst in Section 1.2, the perceived role of both PLZF and BCL-6 in the development of their associated haematopoietic malignancies is introduced.

1.1 Regulation of gene transcription

The control of the rate of transcriptional initiation by both cell type specific and ubiquitous transcription factors represents the fundamental mechanism of regulating gene expression. Even before the discovery of DNA structure it was recognised that the differential expression of genes is central to the processes through which a given cell acquires specific characteristics (Avery, MacLeod et al. 1944; Hershey and Chase 1952). The classical work of Jacob and Monod on the *Escherichia coli* lac operon established a general model that gene expression is controlled by DNA binding regulator proteins (Jacob and Monod 1961). Today it is well established that gene expression is regulated through a whole combination of transcription factors which assemble around the control sequences of a given gene by virtue of DNA binding and/or protein dimerisation; the complexity of this process ensures the accuracy of regulation (De La Brousse and Mcknight 1993; Tjian and Maniatis 1994). Some of these regulatory factors are essential for mediating basal transcription whereas other cell specific factors activate or repress this basal activity.

1.1.1 General transcription factors

The promoter of a gene represents the site at which transcription factors exert their effects and is comprised of both the core promoter and the upstream (or downstream) enhancer sequences (Tjian and Maniatis 1994). The core promoter of all eukaryotic genes represents the minimal DNA elements that are necessary and sufficient for basal transcription initiation by RNA polymerase II. For accurate transcription, RNA polymerase II requires six additional interacting general transcription factors (GTFs) which comprise the core promoter bound basal transcription or pre initiation complex (PIC). These six GTFs are highly conserved between yeast and man and include TFIIA, TFIIB, TFIIE, TFIIIF, TFIIH and the TBP (TATA binding polypeptide) subunit of TFIID (Burley and Roeder 1996).

As discussed below, there are a number of core promoter permutations and both the PIC constituents and their assembly pathway can vary from promoter to promoter. With respect to the TATA core promoter, it is recognised that the formation of the PIC is initiated by the binding of the TBP to the TATA box DNA element. This leads to DNA conformational changes which bring sequences upstream and downstream of the TATA box into closer proximity. This complex is stabilised by TFIIA binding to both the TBP and upstream DNA sequences although, for the TATA core promoter, the presence of TFIIA is not essential for basal transcription initiation. TFIIB binds to the TBP as well as upstream and downstream DNA sequences. This then recruits the pre-formed TFIIIF-RNA polymerase II complex to the core promoter, through its interaction with TFIIB, which consequently mediates the start site selection of the RNA polymerase II. The phosphorylation of the C-terminal domain (CTD) of RNA polymerase II is necessary for the transition from transcriptional initiation to elongation and subsequent termination. TFIIH has kinase activity and binds to RNA polymerase II bound TFIIE and, in the presence of ATP, mediates efficient initiation, elongation and promoter clearance to allow room for subsequent rounds of initiation. TFIIH also contains a helicase which is thought to be required to separate the DNA strands around the transcriptional start site prior to elongation. Despite this relatively well defined step-by step assembly of the PIC, there is currently much debate as to whether this reflects the in vivo

setting owing to the reports of variable combinations of pre-assembled RNA polymerase II-GTF complexes (Roeder 1996).

The PIC is formed at the core promoter of all transcribed genes however the DNA sequence of the core promoter can vary. The TATA box and initiator element (Inr) are the two recognised key elements of the core promoter. Some core promoters contain both elements, others contain either one and null promoters contain neither (Novina and Roy 1996). Specific element combinations are evident in the regulatory genes of different developmental processes, for example the TATA⁻ Inr⁺ is common among the haematopoietic lineage specific genes (Eichbaum 1994). The existence of different 'general' core promoters has led to the suggestion that they might mediate distinct transcriptional signals; this is supported by the observation that mutations in the activation domain of a given gene specific regulator results in differing transcriptional responses on different core promoters (Das, Hinkley et al. 1995).

1.1.2 Regulatory transcription factors

Given that there are a limited number of core promoter permutations, the most effective generation of a variety of transcriptional signals is mediated through the interaction of a multitude of gene specific regulatory transcription factors (regulators) with their specific response elements within the upstream or downstream enhancer regions of the promoter. These regulators activate or repress the GTF mediated basal transcription and the subject of much debate is how their presence at enhancer sites can be converted into a transcriptional response.

The binding to specific response elements is mediated through a number of different motifs such as zinc fingers and helix turn helix domains but to exert an effect, transcriptional regulators must also possess activator or repressor domains (Ptashne and Gann 1990; Hanna-Rose and Hansen 1996). The multiple steps during the formation of the PIC have been shown to be regulated by individual activators and repressors through direct protein-protein interactions (Tjian and Maniatis 1994; Burley and Roeder 1996). Likewise it is anticipated, although not proven, that activators and repressors can regulate transcriptional steps immediately following the initiation event,

such as elongation and promoter clearance (Goodrich and Tjian 1994). Whilst enhancer elements can be in close proximity to the basal transcription complex thus facilitating a direct interaction, they may also be at some distance, upstream or downstream from the core promoter. Therefore to mediate a direct interaction, the binding of a regulator to the enhancer region must cause a conformational change in the DNA to bring the general transcriptional apparatus into close proximity (Ptashne 1988).

The failure of *in vitro* transcription systems reconstituted with purified general transcription factors to respond to regulators was the first indication that communication with the PIC may also be mediated indirectly through co-factors (Meisterernst, Roy et al. 1991). TAFs (TBP associated factors) are activatory co-factors which bind directly to the TBP to form the TFIID complex. Numerous activatory regulators have been demonstrated to interact with distinct TAF subunits. These TAFs, therefore, provide a bridge between the regulators and the TBP. This is believed to enhance the formation of the initiation complex by actively recruiting the TBP to the core promoter (Verrijzer and Tjian 1996). The mediator is a recently recognised multi-protein co-factor complex which binds to the CTD of RNA polymerase II to form the holoenzyme. Different components of the mediator are thought to enhance or inhibit the phosphorylation dependant transition between initiation and elongation. How the gene specific regulators influence this mediator activity has yet to be established (Bjorklund and Kim 1996).

There is an increasing number of recognised general co-factors that can be inhibitory or activatory and are thought to be specific for different families of transcriptional regulators. The activity of these co-factors is best illustrated by recent advances in characterising the retinoic acid (RA) mediated transcriptional regulation. The retinoic acid receptors ($RAR\alpha$, $RAR\beta$, $RAR\gamma$) and retinoic X receptors ($RXR\alpha$, $RXR\beta$, $RXR\gamma$) are members of the nuclear receptor superfamily and are important mediators of the biological activities of retinoids (RA). They act as RA inducible transcription factors which bind as homodimers and/or heterodimers to the promoters or enhancers of RA responsive genes (Chambon 1994). In the absence of RA, RXRs can form DNA binding heterodimeric complexes with RARs which mediate the repression of basal transcription through their interaction with the repressive co-factor NCoR (nuclear receptor co-repressor). However, upon binding of the RA ligand, the NCoR dissociates from the complex (Horlein, Naar et al.

1995). It has been proposed that a co-activator, such as SRC-1, then binds to the ligand bound heterodimeric complex and mediates or enhances its activation of basal transcription (Onate, Tsai et al. 1995). This is, however, an over simplification of events as both RARs and RXRs have been demonstrated to interact directly with components of the basal transcriptional apparatus (Bianco, Wang et al. 1995; Schulman, Chakravarti et al. 1995).

Co-factors such as the GAGA protein can facilitate transcriptional activation transiently by increasing the accessibility of regulators to promoter DNA by disrupting the naturally repressive nucleosomes, see section 1.1.4. In this respect, transcriptional regulation in eukaryotes, in contrast to prokaryotes, has conventionally been associated with activation against a background of histone mediated repression. Whilst activators can enhance the rate of basal transcription initiation, repressors have traditionally been defined as quenchers of activated transcription. This is thought to be achieved by precluding the activity of an activator, either directly by steric hindrance through a protein interaction or indirectly by occupying its DNA binding site, whilst leaving the basal transcriptional initiation intact. However, as indicated for the NCoR mediated repression, repressors can also actively inhibit basal transcription. Therefore, the concept of activatory predominance in the regulation of transcription is being challenged due to the identification of an increasing number of non-quenching repressors (Levine and Manley 1989).

Many transcriptional regulators contain domains which are required to facilitate direct interactions with components of the basal transcription complex. In addition, their ability to form homodimeric and heterodimeric complexes can have a number of regulatory consequences as is evident for members of the Myc-Max regulatory network. The DNA binding capacity of a transcription factor can be influenced by its dimeric partner; the Myc homodimer binds DNA weakly, if at all and requires heterodimerisation with Max for efficient binding (Blackwood and Eisenham 1991). The different dimeric conformations of transcription factors can confer opposing transcriptional activities. The Myc-Max complex activates transcription, whereas the Max-Max, Max-Mad and Max-Mxi1 heterodimers all bind to the same response element but are all inhibitory to varying degrees (Kretzner, Blackwood et al. 1992; Ayer, Kretzner et al. 1993; Zervos, Gyuris et al. 1993). As discussed above the Max-Max interaction is associated with a quenching

effect as it reduces transcription back to a basal level, however both the Mxi1 and Mad proteins in complex with Max can actively repress basal transcription by recruiting the repressive co-factor mSin-3 (Ayer, Lawrence et al. 1995; Schreiber, Chin et al. 1995). The heterodimeric interaction can confer a dominant negative effect; a splice variant of the Max protein, dMax, has been characterised that lacks the DNA binding capacity of Max but is still able to heterodimerise with Myc. However, the resultant complex is unable to bind DNA and activate transcription (Arsura, Deshpande et al. 1995). A similar effect is apparent for SHP, a novel member of the nuclear receptor superfamily which lacks a DNA binding domain but can dimerise with RXR and RAR and inhibit their transcriptional activity (Soel, Choi et al. 1996). An additional consequence of protein dimerisation, which is not evident for the Myc-Max network is that the different protein complexes can confer different DNA binding specificities. For the Jun protein, its heterodimerisation with Fos results in the recognition of the AP-1 DNA binding site however the Jun/CREB complex preferentially binds to the related cyclic AMP responsive element, CRE (Benbrook and Jones 1990; Macgregor, Abate et al. 1990).

1.1.3 Developmental transcriptional regulation

During any process of development, cells must control their pattern of gene expression in response to extracellular stimuli, such as growth factors and cytokines, and express thousands of genes in both a temporally and a spatially regulated fashion. To activate or repress transcription, regulators must be localised to the nucleus and be in the appropriate conformation to bind to DNA and interact with the basal transcription machinery. Accordingly, these processes are the targets of extracellular stimuli mediated transcriptional regulatory signals. Most commonly, this regulation is achieved by reversible, protein kinase mediated, phosphorylation signals (Hunter and Karin 1992).

The mechanisms underlying developmental gene expression are most well established for limb morphogenesis. This is not a reflection of the comparative simplicity of limb development but of the ease with which the avian limb bud can be manipulated *in vivo*. Limb bud development involves the coordinated control of cell differentiation, proliferation and survival to ensure the asymmetric patterning along the dorso-ventral, proximo-distal

and postero-anterior axes (Tabin 1991). Patterning across the antero-posterior, thumb to finger, axis is specified by the polarising region in the posterior mesenchyme. This patterning is mediated by the retinoic acid induced, RAR mediated, production of sonic hedgehog (SHH) which activates the expression of the BMP2 (bone morphogenic protein) as well as members of the HoxD complex. A region of undifferentiated proliferating mesenchyme cells at the tip of the bud is known as the progress zone. Rapidly dividing cells beneath the progress zone determine the proximo-distal axis; the longer the cells remain in the undifferentiating progress zone the more distal the structures they will produce. Both the polarising region and the progress zone are thought to be maintained by signals mediated by the fibroblast growth factors (FGF 4 and FGF 8) from the apical ectodermal ridge (Duprez, Kostakopoulou et al. 1996).

A number of transcriptional regulators involved in limb development, such as Hox genes and BMPs as well as RARs, figure prominently in haematopoiesis. This overlap in transcription factor usage is thought to be a reflection of the derivation of both blood and limb buds from embryonic mesoderm (Zon 1995). The tightly regulated developmental process of haematopoiesis is responsible for the generation of ten morphologically distinct blood cell types, most with limited life spans and all derived from the same multipotential precursor cells (Godin, Garcia Porrero et al. 1993). The transition from immature undifferentiated precursor cells to lineage restricted blood cells, as well as the maintenance of a constant rate of output for each lineage, in analogy to limb development, involves the balanced regulation of differentiation, proliferation and cell survival (Shivdasani and Orkin 1996).

Haematopoiesis has been traditionally subdivided into myeloid and lymphoid lineages, reflecting the earliest point of divergence in the program of differentiation. The myeloid lineage gives rise to eosinophils, basophils, monocytes, neutrophils, erythrocytes and platelets (Eaves 1996). Whereas the lymphoid lineage generates the B cells, T cells and NK (natural killer) cells of the immune system (LeBien 1996). During murine embryogenesis, haematopoietic precursors cells are transiently identifiable in the yolk sac and intraembryonic AGM region (Aorta, Gonad and Mesonephros) and later in development, the site of haematopoietic initiation moves to the fetal liver (Muller, Medvinsky et al. 1994). In the adult, the process is restricted to the bone marrow and whilst this is the site of complete myeloid differentiation

(Eaves 1996), the generation of mature lymphoid cells is completed in the secondary lymphoid organs; either the thymus, lymph nodes or spleen (Cooper 1987). These secondary lymphoid organs are the sites of lymphoid cell exposure to antigens (see section 1.2.2 for details).

In the normal course of events, haematopoietic progression towards mature blood cells involves the sequential activation and silencing of a number of transcriptional regulators as well as changes in the expression and signalling capacity of the haematopoietic cell surface receptors for different growth factors. Growth factor bound receptors will trigger a tyrosine kinase mediated cascade of signal transduction culminating in the initiation of DNA synthesis, the activation of a differentiation response or the regulation of an apoptotic response (Varmus and Lowell 1994). Transcriptional regulators mediating this developmental process are initially implicated as a result of their disruption in chromosomal translocations associated with malignant haematopoiesis or based on their specific expression patterns. For example, RAR α , which is involved in a leukaemogenic translocation (see section 1.2.1), is the predominant RAR expressed in haematopoietic cells and is thought to function in the control of myeloid differentiation in the presence of its obligate heterodimerisation partner, RXR (Dawson, Elstner et al. 1994; Onodera, Kunisada et al. 1995).

1.1.4 POZ domain transcriptional regulators

The POZ domain is a 120 amino acid, highly conserved (see Figure 1), hydrophobic motif (Albagli, Dhordain et al. 1995) and is present in two

(See next page)

Fig. 1. Alignment of the entire POZ domain of mPLZF and 19 other POZ-domain containing Zn-finger and non-Zn-finger protein regions selected from the data base. Sequences, which are numbered on the left with respect to their Met initiation codons, were aligned using Microgenie software (Beckmann). Black and grey backgrounds are used to indicate identical and/or conserved residues found in at least ten times at a given position. Conserved amino acid substitutions are defined according to the scheme (A,S,T), (Q,N), (E,D), (F,Y), (H,K,R) and (I,L,M,V). Consensus sequence is shown at the top of the alignment. The numbered arrows indicate the position of the degenerate primers used in the RT-PCR screening strategy, see section 2.1.5 for nucleotide sequence.

	H	S/T	LØØ	ØN	ØR	2 -/G LØC D V TØL Ø Ø Ø G										4 FØ + A H R K Ø V L A S A C S Y FØ L F																																										
PLZF	16	H	P	G	L	L	C	K	A	N	Q	M	R	L	L	G	T	L	C	D	V	V	I	M	V	D	-	S	Q	E	F	H	A	H	R	T	V	L	A	C	T	S	K	M	F	E	I	L	F	H	R	N	S	Q				
Z13	6	H	S	O	R	V	L	E	Q	L	N	Q	Q	R	S	R	D	L	L	T	D	V	V	I	V	V	D	-	G	V	D	F	K	A	H	K	T	V	L	A	A	C	C	S	G	L	F	Y	S	L	F	-	-	T	D	Q	L	
BCL-6/LAZ-3	14	H	A	S	D	V	L	L	N	L	N	R	L	R	S	R	D	L	L	T	D	V	V	I	V	V	S	-	R	E	Q	F	K	A	H	K	T	V	L	A	A	C	C	S	G	L	F	Y	S	L	F	-	-	T	D	Q	L	
z15	18	H	K	T	L	F	L	K	T	L	N	E	Q	R	R	L	E	G	F	C	D	I	A	I	V	V	E	-	D	V	K	F	K	A	H	R	E	C	V	L	A	A	C	C	S	T	Y	F	K	K	L	F	K	K	L	E	V	
kup	6	H	S	L	V	L	Q	Q	L	N	M	Q	R	R	E	F	G	F	L	C	D	C	T	V	A	I	G	-	D	V	Y	F	K	A	H	R	A	V	L	A	A	C	C	S	N	Y	F	K	M	L	F	-	-	I	H	Q	T	
HFBCM81	6	H	P	N	N	L	L	K	E	L	N	K	C	R	L	S	E	T	M	C	D	A	T	I	V	V	G	-	S	R	S	F	K	A	H	R	A	V	L	A	A	C	C	S	G	Y	F	Q	N	L	F	L	N	T	G	L		
ZFPJS	8	H	S	V	R	V	L	Q	E	L	N	K	Q	R	E	K	G	Q	Y	C	D	A	T	L	D	V	G	-	G	L	V	F	K	A	H	R	W	S	V	L	A	A	C	C	S	H	F	F	Q	S	L	Y	-	-	-	G	D	
ZID	15	Q	G	D	V	V	L	Q	K	M	N	L	L	R	Q	Q	N	L	F	C	D	V	S	I	V	I	N	-	D	T	E	F	Q	E	H	K	V	L	A	A	C	C	S	Y	F	M	R	D	Q	F	L	L	T	Q	-			
FDP-B	6	H	S	R	Q	L	L	L	Q	L	N	T	Q	R	T	K	G	F	L	C	D	V	I	V	V	Q	-	N	A	L	F	H	K	W	L	L	A	A	C	C	S	A	Y	L	K	S	L	-	-	-	-	-	-	-	-			
vmt-8	1				M	S	Y	P	L	Y	K	L	R	K	G	K	L	C	D	V	E	V	I	V	A	E	-	G	K	S	I	F	A	H	R	L	V	L	S	A	Y	K	Y	F	S	N	L	L	F	S	I	L	F	S	I			
va55	3	N	S	S	E	L	I	A	V	I	N	G	F	R	N	S	G	R	F	C	D	I	S	I	V	I	-	D	K	S	I	F	A	H	R	L	V	L	S	A	Y	K	Y	F	S	I	L	L	F	S	I	L	F	S	I			
vp65	8	H	N	R	R	V	V	S	N	I	S	S	L	L	D	N	D	I	L	C	D	V	I	I	T	I	G	D	G	E	E	L	K	A	H	K	T	L	A	A	C	C	S	K	Y	F	R	T	L	L	F	T	T	P	M	I		
GAGA	16	Y	G	T	S	L	V	S	A	I	O	L	L	R	C	H	G	D	L	V	D	C	T	L	A	A	G	-	G	R	S	F	P	A	H	K	I	V	L	C	A	A	C	C	S	P	F	L	L	D	L	L	K	N	T	P	-	
ttk	15	H	Q	S	N	L	I	S	V	F	D	Q	L	L	H	A	E	T	F	T	D	V	T	L	A	V	E	-	G	Q	H	L	K	A	H	K	M	V	L	S	A	C	C	S	P	Y	F	M	T	L	L	F	V	S	H	P	-	
lola	14	H	Q	S	T	L	I	S	V	F	D	T	L	L	E	N	E	T	L	V	D	C	T	L	A	A	E	-	G	K	F	L	K	A	H	K	M	V	L	S	A	C	C	S	P	Y	F	A	T	L	L	F	V	S	E	Q	Y	-
BR-C	14	Y	Q	S	S	I	T	S	A	F	E	N	L	R	D	D	E	A	F	V	D	V	T	L	A	C	E	-	G	R	S	I	K	A	H	R	V	V	L	S	A	C	C	S	P	Y	F	R	E	L	L	K	S	T	P	-		
E(var)3-9D	14	F	N	T	N	L	S	A	G	F	H	E	S	L	C	R	G	D	L	V	D	V	S	L	A	A	E	-	G	Q	I	V	K	A	H	L	V	L	S	A	C	C	S	P	F	F	R	K	M	F	T	Q	M	P	S			
bab	11	Y	Q	T	N	L	T	T	I	F	D	Q	L	L	Q	N	E	C	F	V	D	V	T	L	A	C	D	-	G	R	S	M	K	A	H	K	M	V	L	S	A	C	C	S	P	Y	F	Q	T	L	L	A	T	S	P	C		
kelch	138	H	T	A	R	S	F	D	A	M	N	E	M	R	K	Q	K	Q	L	C	D	V	I	L	V	A	D	-	D	V	E	I	K	A	H	R	L	V	L	A	A	C	C	S	P	Y	F	Y	A	M	F	L	A	T	S	P	C	
c08c3	129	T	G	F	S	K	D	V	L	R	S	F	D	E	S	E	K	R	F	S	D	V	I	L	V	V	G	-	D	E	K	F	Y	V	L	K	L	F	L	A	A	C	C	S	S	Y	F	N	A	L	F	L	G	K	F	K		

	H	ØØ	L	Ø	E/D	Ø	3 LØLØEDFY Ø Y T G Ø Ø										Ø	Ø	L	A/T	L	Ø	L																																		
PLZF	69	-	-	H	Y	T	-	-	-	L	D	F	L	-	S	P	K	T	F	Q	Q	L	E	F	A	Y	T	A	T	L	Q	A	K	A	E	D	L	D	D	L	L	Y	A	A	E	I	L	E	I	E	L						
Z13	57	-	K	D	-	-	V	V	H	-	L	D	-	L	S	-	N	A	A	G	L	Q	Q	L	E	F	A	Y	T	A	K	L	S	L	S	P	E	N	V	D	D	V	L	A	V	A	S	F	L	Q	M	Q	D	I			
BCL-6/LAZ-3	66	K	C	N	L	S	V	I	N	-	L	D	P	E	I	-	N	P	E	G	F	C	I	L	L	D	F	A	Y	T	S	K	R	L	N	L	R	E	G	N	I	M	A	V	M	A	T	A	M	Y	L	L	Q	M	E	H	V
z15	71	D	S	S	-	S	V	I	E	-	I	D	F	L	R	-	S	D	I	F	E	E	L	L	H	Y	M	Y	T	A	K	T	S	V	K	K	E	D	V	N	L	M	M	S	S	G	Q	I	L	G	I	R	F	L			
kup	58	S	E	C	I	K	I	Q	P	-	T	D	-	I	-	Q	P	D	I	F	S	Y	L	L	H	I	M	Y	T	G	E	K	G	P	K	Q	I	V	D	H	S	R	L	E	E	G	I	R	F	L	L	H	A	D	Y		
HFBCM81	6	D	A	A	R	T	Y	V	-	-	V	D	F	I	-	P	A	N	-	F	? K	L	L	S	F	Y	T	G	E	L	F	T	D	L	I	N	V	G	V	I	Y	E	V	A	E	R	L	L	G	I							
ZFPJS	58	G	S	G	S	V	V	-	-	L	P	A	G	-	F	A	E	I	F	G	L	L	L	S	F	Y	T	G	E	H	L	A	L	T	S	G	N	R	D	Q	V	L	L	A	A	E	R	L	L	R	V	P	E	A			
ZID	67	S	K	H	-	V	R	I	T	I	Q	S	A	E	-	V	G	R	-	-	-	K	L	L	L	S	C	Y	T	G	A	L	E	V	K	R	K	E	L	L	K	Y	L	L	A	A	S	Y	L	L	Q	M	V	H	I		
FDP-B	53	V	V	H	D	N	L	L	N	-	L	D	H	E	M	V	S	P	G	I	F	R	L	L	D	F	Y	T	G	E	L	G	E	C	E	P	G	G	E	Q	S	L	G	A	V	L	A	A	A	S	Y	L	O				
vmt-8	51	E	K	N	V	D	V	I	D	-	L	E	A	D	-	Y	K	T	-	V	F	D	L	I	Y	T	Y	T	G	E	S	I	E	L	H	K	G	N	T	E	S	I	F	S	L	V	H	Y	L	L	Q	I	K	P	L		
va55	56	D	S	N	E	Y	E	V	N	-	L	E	H	L	D	-	Y	Q	S	-	V	N	D	L	L	Y	Y	T	G	E	I	P	L	S	L	T	N	D	N	V	K	Y	I	L	L	A	D	F	L	L	Q	I	G	S	A		
vp65	62	I	R	D	L	V	T	R	V	N	L	Q	M	F	D	-	K	D	A	-	V	K	N	Y	Q	Y	L	Y	T	G	E	H	H	I	S	S	-	-	M	N	V	I	D	V	L	K	C	A	D	Y	L	L	I	D	D	L	
GAGA	68	C	K	H	P	V	V	M	-	-	L	L	A	G	V	N	-	A	N	D	-	L	E	A	L	L	E	F	Y	Y	H	G	E	V	S	V	D	M	A	O	L	P	S	L	L	Q	A	A	O	C	L	L	W	I	Q	G	L
ttk	67	E	K	H	P	I	V	I	-	-	L	K	D	V	P	-	Y	S	D	-	M	K	S	L	L	D	F	Y	Y	H	G	E	V	S	V	D	Q	E	R	L	T	A	F	L	L	R	V	A	A	E	S	L	L	Q	I	K	G

distinct classes of transcriptional regulators: the POK (POZ and *kriippel*) proteins and the Bach proteins. The high degree of evolutionary conservation of the POZ domain for a given protein implies that it is an important functional region. As will be discussed, the structure of the POZ domain implicates it as a potential protein interaction motif which could clearly have a number of transcriptional regulatory consequences.

POK proteins are a newly identified family of potential transcriptional regulators that contain an N-terminal POZ domain and a C-terminal *Kriippel*-like zinc finger DNA binding motif (Schuh, Aicher et al. 1986). POK proteins represent the largest class of POZ domain proteins and the function of both the Mammalian and *Drosophila* members are summarised in Table 1. As will be discussed in section 1.2, two members of this class of proteins, BCL-6 and PLZF, are involved in oncogenic translocations. The majority of *Drosophila* members of this class also contain zinc finger domains with the exception of E(Var) 3-93D-B which contains stretches of charged amino acids thought to mediate DNA binding. Like the POZ domain, the *Kriippel* associated box (KRAB) is a conserved structural domain, which lies N-terminal to the *Kriippel*-like zinc finger region of a number of proteins (Bellefroid, Poncelet et al. 1991; Witzgall, O'Leary et al. 1994); therefore this first class of POZ proteins also defines a new subfamily of *Kriippel*-like proteins.

The recent identification of the POZ proteins, Bach1 and Bach2, has led to the realisation of a second class of DNA binding POZ proteins which are potential transcription factors but contain a C-terminal basic leucine zipper motif instead of a zinc finger region (Oyake, Itoh et al. 1996). Basic leucine zipper motifs mediate both the dimerisation and DNA binding of a protein (Vinson, Sigler et al. 1989). Bach 1 and Bach 2 were identified by a yeast-two hybrid screen as a result of their leucine zipper mediated interaction with the small MafK transcription factor. Whereas both Bach proteins display transcriptional repression, Bach1 can also act as an activator. Their interaction with MafK and their expression patterns suggest an involvement in hind brain development and haematopoiesis. The *Drosophila* protein pipsqueak (psq) can also be tentatively assigned to this fourth class owing to the presence of two protein interaction motifs. Downstream of the N-terminal POZ domain there are 34 alternating histidine residues that are proposed to mediate protein interactions by the formation of a histidine-metal zipper. This protein also contains another newly identified potential

Table 1. POZ domain proteins, class 1

A) Mammalian POK proteins

Protein	No. of zinc fingers	Transcriptional activity	Proposed function
PLZF (Promyelocytic leukaemia zinc finger)	9	Repressor	Myeloid development and hind brain development. Implicated in the pathogenesis of acute promyelocytic leukaemia. 2 transcripts (Chen, Brand et al. 1993).
BCL-6 (B cell Lymphoma 6)	6	Repressor	B cell development and germinal centre formation. Implicated in the pathogenesis of non hodgkins lymphoma (Ye, Lista et al. 1993).
ZID	4	?	Unknown. Binds a actin promoter (Bardwell and Treisman 1994).
ZF PJS (clone 18)	11	Activator	Regulates MHC II DPA promoter (Sugawara, Scholl et al. 1994).
KUP	2	?	Unknown. Predominantly expressed in testes and haematopoietic cells (Chardin, Courtois et al. 1991).
ZF5	5	Repressor	Regulates the c-myc promoter (Numoto, Niwa et al. 1993).
gFBP	5	Repressor	Lens development and somite differentiation. 3 transcripts (Liu, Shalaby et al. 1994).
HIC-1	5	?	Tumour suppresser gene (Wales, Biel et al. 1995).
Z13	13	?	Unknown. Ubiquitous expression (Schulz, Hopwood et al. 1995).

b) Drosophila POK proteins

Tramtrack (TTK)	2	Repressor	Embryonic segmentation. Cell fate determination; photoreceptor development. 2 transcripts (Brown and Wu 1993; Xiong and Montell 1993)
GAGA	1	Activator	PEV (position effect variegation). Pattern formation. Regulation of nuclear cleavage. 2 transcripts (Soeller, Eukoh et al. 1993).
E(VAR)/mdg4	-	Activator	PEV. Pattern formation. 2 transcripts (Dorn, Krauss et al. 1993).
LOLA (longitudinal lacking gene)	2	Repressor	Axon growth and guidance. Regulation of the copia retrotransposon. 2 transcripts (Giniger, Tietje et al. 1994; Cavarec, Jenson et al. 1997)
BAB (bric a brac)	?	?	Limb and ovary development. Pattern formation (Zollman, Godt et al. 1994).
Br-C (Broad complex)	2	Activator and repressor	Primary response regulator of metamorphosis. 4 transcripts (DiBello, Withers et al. 1991; Karim, Guild et al. 1993).
Fru (Fruitless)	2	?	Determination of male sexual orientation. Multiple transcripts (Ryner, Goodwin et al. 1996).
Abrupt	2	?	Limb morphogenesis. Neuromuscular connections (Hu, Fambrough et al. 1995).

DNA binding psq motif at its C-terminus (Horowitz and Berg 1996). There are at least six different splice variants of psq and the protein products are thought to play a role in oogenesis and like the POK protein Tramtrack (TTK), also photoreceptor development (Weber, Siegel et al. 1995).

The POZ domain is present in two other classes of proteins which lack any obvious DNA binding motifs and are not implicated as transcriptional regulators. Nevertheless, the activities of these proteins might eventually shed light on the shared function of the POZ domain. The third class of POZ proteins all share a common structural organisation in which the N-terminal POZ domain is linked to six imperfect repeats terminating at Gly-Gly doublets. Members of this class include a *Caenorhabditis elegans* protein, the *Drosophila* Kelch protein and a number of viral proteins (Albagli, Dhordain et al. 1995). The Gly-Gly doublets have been shown to mediate actin binding in the POZ-less homologue, scruin (Way, Sanders et al. 1995). The Kelch protein has been shown to co-localise with intercellular bridges or ring canals. Kelch has been proposed to mediate the cytoplasmic flow from supporting nurse cells to developing oocytes through these ring canals by a direct interaction with the actin cytoskeleton (Xue and Cooley 1993; Cooley and Theurkauf 1994). Along the same lines, it is proposed that the viral POZ proteins in this class are involved in the infectious cycle by mediating the assembly of microvilli through their interactions with the actin cytoskeleton of the host cell (Koonin, Senkevich et al. 1992). The fourth class comprises POZ proteins which contain no other obvious structural motifs. Two putative *C.elegans* proteins have been assigned to this group as well as the *Drosophila* germ cell-less (*gcl*) gene product which has been associated with germ cell precursor development (Jongens, Ackerman et al. 1994; Albagli, Dhordain et al. 1995). As discussed later, these POZ only proteins might be expected to interact with and affect the function of, some of the other DNA binding POZ domain proteins.

The POZ domain has been shown to be sufficient to mediate the homodimerisation of a number of POZ domain proteins *in vitro* and the BCL-6 protein *in vivo* (Bardwell and Treisman 1994; Chen, Zollman et al. 1995; Dhordain, Albagli et al. 1995). In addition, steps have been taken to identify specific residues mediating this protein interaction (Chen, Zollman et al. 1995). Mutational analysis of the *bric a brac* (BAB) *Drosophila* protein has revealed that the N-terminal half of the POZ domain is essential for

mediating homo-dimerisation *in vitro*. The less highly conserved C terminal half of the domain may be only required to stabilise the interaction. Using a computer program this N-terminal region was predicted to be an α -helix with a hydrophobic face comprising 50% leucine residues. Mutation of the leucines within this face reduced its binding to wildtype BAB whereas mutations outside of this face had no effect. An ionic interaction was neatly demonstrated by mutating the two oppositely charged residues flanking the hydrophobic face. When either or both of these residues had its charge reversed, binding to wild type BAB was abolished. However a double mutant (43K to D/35D to K) could bind to another double mutant indicating that these residues could interact *in trans* during dimerisation. This mutational analysis identified residues essential for interaction which are only conserved within the TTK group of proteins and therefore this may not be applicable to POZ proteins in general. In addition, although the results of the analysis support the computer generated structure, different structural predictions can be made depending on the parameters of the program. This is particularly salient in light of the fact that the structure of this N-terminal half of the POZ domain is more commonly predicted to exist as two alternating α -helices and β -sheets (Albagli, Dhordain et al. 1995).

The first heterodimeric interaction between POZ domain proteins was demonstrated by the *in vitro* interaction of the *Drosophila* proteins TTK and GAGA (Bardwell and Treisman 1994). This potential interaction is supported by the existence of a common developmental target gene, engrailed. Perhaps TTK and GAGA bind certain response elements together as a heterodimer but bind their unique target genes as either homodimers or as heterodimeric complexes with other POZ proteins. The apparent diversity of function of individual POZ domain proteins can therefore be accounted for by the formation of heterodimers which potentially allows for the recognition of a wider spectrum of DNA sequences by a limited number of regulators.

The identification of target genes or potential binding sites, for a number of POK proteins as well as the Bach proteins has confirmed their potential role in transcriptional regulation. As indicated in Table 1 transient transfection experiments give differing results; some POK proteins activate transcription whilst others are inhibitory. As for the Bach proteins, depending on the assay system used, Bach 2 is either an activator or a repressor (Oyake, Itoh et al. 1996). This variation in transcriptional activity lends support to the

possibility that POZ domain proteins mediate transcriptional regulation by dimerisation. In different dimeric complexes the POZ proteins might act as activators or repressors of transcription. This regulation of transcription by dimerisation is further supported by the observation that for BCL-6, ZF5 and PLZF, the respective POZ domains have been shown to be partially responsible for their repressive effects (Numoto, Niwa et al. 1993; Albagli, Dhordain et al. 1996 and unpublished data). However this apparent POZ domain mediated repression may be a consequence of an altered nuclear localisation (see later).

Bach 1 and Bach 2 have been implicated in the small Maf network mediated transcriptional regulation. The heterodimeric interaction between either of the bach proteins and Mafk is mediated entirely through the leucine zipper motifs and the complexes can bind to NFE2 sites and repress transcription (Oyake, Itoh et al. 1996). Although not suggested by the investigators, this repressive effect could be mediated directly through the POZ domain of the Bach proteins. NFE2-like binding sites have been shown to be occupied by regulatory proteins during a variety of developmental processes; small Maf proteins can heterodimerise with the P45 protein to form NFE2 protein complexes which are critical for regulating erythroid specific gene expression. Whereas the P45/Maf complex binds the NFE2 binding site and activates transcription, the small Maf proteins have also been shown to bind to the NFE2 site as homodimers and actively repress transcription (Igarashi, Kataoka et al. 1994). With the absence of any obvious regulatory domains within small Maf proteins, the repression has been ascribed to the displacement of endogenous activatory AP1 (Fos/Jun) complexes which can also bind NFE2 sites (Kataoka, Igarashi et al. 1995). However, this apparent, homodimeric, small Maf protein mediated repression can equally be explained as a consequence of their heterodimeric interaction with endogenous Bach proteins, in the assay system, which could actively repress transcription through their POZ domains.

The GAGA POK protein behaves as a transcriptional activator in transient transfections but has also recently been shown to mediate transcriptional activation indirectly by inducing local alterations in chromatin structure thereby creating hypersensitive sites which allow other transcriptional regulators access to DNA (Soeller, Eukoh et al. 1993; Tsukiyama, Becker et al. 1994). The ability of GAGA to dislodge the intrinsically repressive

nucleosomes from their DNA template might occur as a result of their destabilisation due to torsional constraints mediated through protein-protein interactions between GAGA proteins bound at several adjacent sites. Such protein interactions would be expected to be mediated through the GAGA POZ domains. However, despite using an elaborate *in vivo* nucleosome assembly system to demonstrate this indirect anti-repression by GAGA, the effects of a POZ less GAGA mutant were not examined (Tsukiyama, Becker et al. 1994). It is difficult to reconcile this activatory co-factor role for GAGA with its potential interaction with TTK (Bardwell and Treisman 1994). Perhaps the interaction serves to recruit TTK to the site of transcription once GAGA has mediated its open conformation.

In addition to its role in regulating the chromatin structure of individual genes, GAGA has also been proposed to participate in the packaging of large segments of the chromosome into active, euchromatic, and inactive, heterochromatic, domains. This is supported by the association of GAGA with specific heterochromatic regions of the chromosome through out the cell cycle (Raff, Kellum et al. 1995). This may have structural consequences or simply be part of a mechanism to ensure that GAGA is evenly distributed when cells divide. However, mutations in the trithorax-like gene that encodes GAGA have been shown to enhance position effect variegation (PEV) which is characterised by the variegated expression of a gene due to its inactivation as a result of the spreading of inhibitory heterochromatin into neighbouring euchromatin (Henikoff 1990; Farkas, Gausz et al. 1994). The presence of GAGA on the heterochromatin might normally act to inhibit this spreading. Interestingly, mutations in another POZ domain protein, E(Var)3-93D, strongly enhance PEV (Dorn, Krauss et al. 1993). This would suggest that the POZ domains of both E(Var)3-93D and GAGA might be involved in protein interactions to establish an open chromatin conformation.

Studies on the POK protein, ZID, suggest an involvement of the POZ domain in the inhibition of DNA binding (Bardwell and Treisman 1994). This inhibitory effect of the POZ domain was shown to be transferable, as the same effect was seen if the ZID POZ domain was replaced with a POZ domain from a viral POZ protein. This inhibition was shown not to require direct interaction between the POZ domain and zinc finger region as the ZID POZ domain could inhibit DNA binding when fused to heterologous DNA binding motifs. As the ZID protein was shown, *in vitro*, to interact with itself

through the POZ domain, it was proposed that the inhibitory effect of this domain was indirect and due to its mediation of homodimerisation. Therefore the inhibition might be overcome by a heterodimeric interaction which would increase the access of the zinc fingers to DNA. This was supported by the finding that the dimerisation of full length ZID protein with just its POZ domain increased its DNA binding. The inhibition of DNA binding by the POZ domain may, however, not be a general property of all POZ proteins (Albagli, Dhordain et al. 1995). Unpublished results suggest that the presence of the POZ domain actually enhances the DNA binding capacity of PLZF. The DNA binding site of ZID was selected, *in vitro*, using only the zinc finger portion of the protein (Bardwell and Treisman 1994) and it may not reflect its *in vivo* response element; hence this inhibition of DNA binding maybe an *in vitro* phenomenon.

Nevertheless, by demonstrating an increase in DNA binding capacity of ZID in a heterodimeric complex with a POZ only mutant, Bardwell and Triesmann highlighted a potential dominant negative transcriptional regulatory effect of POZ domain proteins (Bardwell and Treisman 1994). Independent evidence of this comes from experiments with the *Drosophila* protein, psq. As stated, this protein is involved in oogenesis and *in vivo* expression of a psq mutant containing only the POZ domain leads to a reduction in egg laying. The presence of two copies of the mutant transgene further enhances this inhibitory effect on egg production. It was suggested that the interaction of the psq POZ only mutant with a heterologous partner POZ protein might increase the access of the partner protein's zinc fingers to DNA resulting in the inappropriate activation of downstream target genes (Horowitz and Berg 1996). This dominant negative effect could prove to be a physiological phenomenon as the fourth class of POZ only proteins, discussed above, may represent a wild type equivalent of this POZ only mutant. In addition, as indicated in Table 1, a number of POZ domain proteins can produce multiple transcripts by alternative splicing and this represents an attractive method of potentially generating alternative, POZ only, variants.

In the case of the POK proteins PLZF, BCL-6 and ZID, the POZ domain has been shown to mediate the speckled nuclear localisation of the protein. In the absence of the POZ domain the nuclear localisation is more diffuse and is some times cytoplasmic (Bardwell and Treisman 1994; Dhordain, Albagli et

al. 1995; Reid, Gould et al. 1995). The BCL-6 POZ domain has been shown to be dispensable for the targeting of the protein to the nucleus however its POZ domain is able to impose the same discrete nuclear localisation when fused to heterologous nuclear proteins (Dhordain, Albagli et al. 1995). This does call into question the proposed role of the same domain in mediating the transcriptional repression of BCL-6, ZF5 and PLZF, discussed above. In the reporter systems used, the POZ-less constructs may not have exerted the same repressive effects as the full length protein because of their altered nuclear localisation rather than because of the absence of any POZ domain mediated repressor activity.

Whilst the POZ domain has been implicated in dimer formation, the nuclear speckles may reflect an additional or alternative function of this domain in mediating multimeric complex formation. This is a particularly attractive role for the Bach proteins which possess two different protein interaction motifs. It has been suggested that the interaction between Bach1 and Bach2 and MafK could result in the formation of heterogeneous multimeric regulatory structures, as MafK is itself part of a network of Maf proteins (Oyake, Itoh et al. 1996). It will be of interest to see if either Bach 1 or Bach 2 and MafK colocalise to the same nuclear speckles.

1.2 Leukaemias and lymphomas

The uncontrolled malignant expansion of haematopoietic cells with the concomitant block in their differentiation, at varying stages of development towards mature blood cells, is the hallmark of leukaemic progression. The disease was originally classified into acute and chronic leukaemia on the basis of the duration of survival of those afflicted. The same classification is used today to distinguish 'acute' leukaemias, resulting from the proliferation of immature cells, from 'chronic' forms of the disease which are the result of the proliferation of more differentiated haematopoietic cells (Sawyers, Denny et al. 1991). Whereas the leukaemic phenotype is identifiable in myeloid and lymphoid haematopoietic cells, in contrast, lymphomas are restricted to the lymphoid lineage and are characterised by the malignant proliferation of more mature T and B lymphocytes within the secondary lymphoid organs (Magrath 1990). Both leukaemias and lymphomas are monoclonal in origin as they arise from a single cell as a result of the accumulation of a number of

genetic changes which are sufficient to subvert the normal haematopoietic regulatory control mechanisms, see section 1.1.3. These genetic alterations arise as a result of genetic deletion, amplification or recombination, the latter including chromosomal translocation, insertion or inversion. The initial genetic change must confer a growth advantage to ensure the survival of its subclonal descendants. In some instances, this first 'hit' may be sufficient for disease progression; alternatively the resultant expansion of the genetically aberrant cell will increase the chances of a second, cooperative, mutation occurring in the same cell (Greaves 1996).

Chromosomal translocations are commonly associated with leukaemias and lymphomas and they frequently affect the function of transcriptional regulators. These chromosomal translocations can be classified into two types (Rabbitts 1994). A classical example of a type I translocation, which results in the generation of a novel fusion protein comprising part of the coding region of the genes from each chromosome, is the chronic myelogenous leukaemia (CML) associated Philadelphia translocation. This translocation fuses the BCR and c-ABL genes (Heisterkamp 1983) and the resultant chimera has constitutive tyrosine kinase activity resulting in the uncontrolled expansion of the myeloid compartment. This expansion is thought to be due to the suppression of apoptosis rather than the deregulation of cellular proliferation (Kabarowski, Allen et al. 1994). As the development of leukaemia involves both enhanced expansion of haematopoietic cells and a block in differentiation, additional genetic changes often accompany chromosomal translocations; for example, mutations or rearrangements in the p53 tumour suppresser gene have been associated with CML (Ahuja, Bar-Eli et al. 1986).

Type II chromosomal translocations result in the inappropriate expression of a gene as a result of its transactivation by the factors regulating the gene juxtaposed to it. These translocations frequently involve the fusion of a transcriptional regulator to the T cell receptor or immunoglobulin genes. The most frequent translocations associated with the B cell malignancy, Burkitts lymphoma, juxtapose the c-Myc and immunoglobulin heavy-chain genes leading to the aberrant expression of c-Myc (Rabbitts and Boehm 1991). The resultant enhanced proliferation of B lymphocytes is a consequence of the deleterious effect of the over expression of c-Myc on the highly concentration sensitive Myc transcriptional regulatory network (see section 1.1.2). In

addition, a number of Type II translocations, which activate transcriptional regulators, such as LMO2, by their fusion with T cell receptor genes, have been associated with T cell acute lymphoblastic leukaemia (T-ALL) (Sanchez-Garcia and Rabbitts 1993). Like c-Myc, LMO2 contains a protein interaction motif; the LIM domain (Sanchez-Garcia and Rabbitts 1994). As will be discussed, two POK proteins, PLZF and BCL-6, are present in Type I and Type II chromosomal translocations, respectively. This supports an important role for transcriptional deregulation, caused by disturbed or inappropriate protein interactions (see section 1.1.2), in the progression of haematopoietic malignancies.

1.2.1 Type I translocations of the RAR α locus in APL

Acute promyelocytic leukaemia (APL) is the M3 subtype of acute myelogenous leukaemia (AML), accounting for 10% of all adult AML cases (Stone and Mayer 1990) and is of particular interest owing to the recent development of a highly successful treatment strategy. APL has been recognised as a clinical entity since the 1950s and the disease is characterised by the malignant proliferation of immature myeloid cells blocked at the promyelocytic stage of development (Hillstead 1957; Warrell, de The et al. 1993). The successful treatment of APL patients stemmed from initial in vitro experiments that demonstrated that the addition of all-trans retinoic acid (ATRA) to cultured APL cells led to their complete differentiation into mature granulocytic cells (Breitman and collins 1981). Although the exact mechanism of this differentiation induction was not established, these experiments led to clinical trials which confirmed that APL patients could be treated with ATRA alone, achieving complete remission in up to 90% of patients (Huang, Ye et al. 1988; Castaigne, Chomienne et al. 1990; Grignani, Fagioli et al. 1994; Warrell, Maslak et al. 1994). However, the high relapse rate in these patients has led to the development of a combined chemotherapy/ATRA treatment approach (Fenaux, Wattel et al. 1994).

To date, four Type I chromosomal translocations have been associated with APL; the t(15;17), t(5;17) and two variant t(11;17) translocations (Rowley, Golomb et al. 1977; Chen, Zelent et al. 1993; Corey, Locker et al. 1994; Wells and Kamel-Reid 1996). In all forms of these translocations the same portion of the RAR α gene on chromosome 17 is fused to the PML, NPM (Nucleophosmin), PLZF or NuMa (Nuclear mitotic apparatus) genes,

respectively (de The, Chomienne et al. 1990; Chen, Brand et al. 1993; Redner, Rush et al. 1996; Wells and Kamel-Reid 1996). The t(15;17) translocation is most commonly associated with APL and gives rise to the expression of a number of novel reciprocal chimeric proteins. In 20% of t(15;17) APL patients, only the PML-RAR α fusion protein and not the reciprocal fusion product is detected suggesting that this chimera is responsible for the pathogenesis of APL (Kakizuka, Miller et al. 1991; Pandolfi, Grignani et al. 1991; Alcalay, Zangrilli et al. 1992). The PML-RAR α fusion replaces the RAR α activation (A) domain with PML sequences whilst the critical functional domains of RAR α , that mediate DNA binding, dimerisation and retinoic acid binding, remain (Brand, Petkovich et al. 1988; Leroy, Krust et al. 1991; Kastner, Perez et al. 1992).

The disruption of the RAR α gene in all APL associated translocations implicates an aberration of wild type RAR α mediated myeloid differentiation in the pathogenesis of APL. This is supported by the fact that a block in promyelocytic differentiation can be caused by a dominant negative RAR α mutant (Tsai and Collins 1993). However, given the predominance of the t(15;17) translocation in APL patients, a wealth of publications have addressed the potential role of PML disruption in the pathogenesis of APL particularly with respect to enhanced proliferation. The PML gene is expressed ubiquitously (Fagioli, Alcalay et al. 1992) and structural elements suggest it may function as a transcription factor; its protein product is a member of the 'ring finger' family of zinc finger proteins whose functions include the regulation of gene expression in development and DNA repair (Lovering, Hanson et al. 1993). In addition to the proposed DNA binding ring fingers, PML also contains a coiled coil protein interaction motif (Kastner, Perez et al. 1992). Although its function is not well characterised, PML has been shown to be a growth suppresser, inhibiting the transformation of both NIH 3T3 cells and rat embryonic fibroblasts by cooperative oncogenes (Mu, Chin et al. 1994). PML is normally localised in the nucleus in large macromolecular aggregates known as PML oncogenic domains (PODS) or nuclear bodies (Dyck, Maul et al. 1994) and the ability of PML to suppress growth is dependent on the formation of PODS (Le, Yang et al. 1996).

Compared to wild type PML, PML-RAR α has a more diffuse nuclear localisation (Kastner, Perez et al. 1992). The expression of PML together with

PML-RAR α causes PML to be delocalised from the PODs to a pattern similar to PML-RAR α and upon treatment with ATRA, PML relocalises to the PODs (Koken, Puvion-Dutilleul et al. 1994; Weis, Rambaud et al. 1994). In addition, co-transfection of PML-RAR α with PML into transformed NIH3T3 cells inhibits the growth suppressive capacity of PML (Mu, Chin et al. 1994). Together these observations have led to the suggestion that the PML-RAR α chimera acts as a dominant negative inhibitor of the wild type PML growth suppresser function which increases the proliferative capacity of the affected cells (Weis, Rambaud et al. 1994). With respect to the additional block in differentiation, required for a full leukaemic phenotype, discussed above, co-localisation studies have demonstrated that PML/RAR α sequesters RXR (Weis, Rambaud et al. 1994). This sequestration is thought to prevent the heterodimeric interaction of RXR with RAR α , or with other heterodimeric partners such as the vitamin D3 receptor (Kliwer, Umesono et al. 1992; Yu, Naar et al. 1992), thus preventing the normal induction of myeloid differentiation (Perez, Kastner et al. 1993).

In conflict with the proposed inhibition of PML mediated growth suppression, as a mechanism of disease progression, is the existence of three other partner proteins for RAR α in APL associated translocations. All four partner proteins possess putative protein interaction motifs. Therefore, the abrogation of RAR α function might be mediated through a common mechanism by the protein interaction motifs of the heterologous partner proteins within the respective RAR α chimeras. A common, protein interaction mediated, interference with retinoid signalling, like the sequestration of RXR for example, could be sufficient for APL disease progression. In this respect, in addition to mediating myeloid differentiation (see above and section 1.1.3), retinoid receptors have been implicated in the induction of apoptosis in myeloid HL-60 precursor cells (Nagy, Thomazy et al. 1995).

Despite this possible 'passive' role for RAR α partner proteins in APL disease progression, the t(11;17)(q23;q21) translocation is of great interest because, unlike 'classical' t(15;17) associated APL, patients harbouring this translocation do not respond to ATRA and respond poorly to chemotherapy (Guidez, Huang et al. 1994; Licht, Chomienne et al. 1995). An understanding of the cause of this resistance will contribute towards the elucidation of the molecular mechanisms mediating the ATRA induced differentiation within

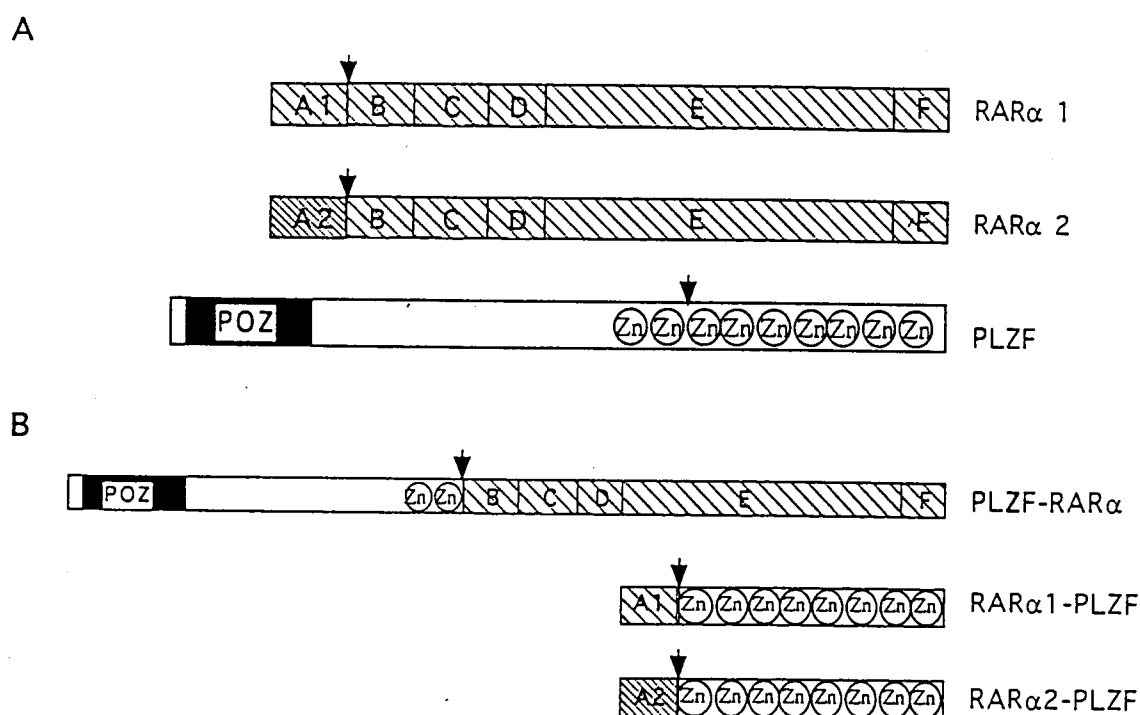


Fig. 2. A. Schematic representation of the two RAR α isoforms and PLZF. The RAR α 1 and RAR α 2 are designated with hatched rectangles subdivided into their conserved (B-F) and divergent A1 and A2 functional domains. The POZ domain (POZ) and nine zinc fingers (circled Zn) of PLZF are indicated.

B. Schematic representation of the PLZF-RAR α and two RAR α -PLZF isoforms. The positions corresponding to the translocation breakpoint are indicated with black triangles.

responsive APL cases. As indicated in Figure 2, unlike the t(15;17) translocation, as well as the PLZF-RAR α chimera, all patients with the t(11;17)(q23;q21) translocation possess the reciprocal RAR α -PLZF chimeric transcripts (Chen, Brand et al. 1993). These transcripts contain fusions of either the A1 or A2 transactivation domains of RAR α to the C-terminal sequences of PLZF including seven of the nine zinc fingers (RAR α /PLZF) (Licht, Chomienne et al. 1995). It is apparent, therefore, that either the disruption of PLZF wildtype function or the presence of PLZF sequences fused to the RAR α sequences may be the cause of the unique ATRA resistance of t(11;17)(q23;q21) patients (Sitterlin, Tiollais et al. 1997).

Wild type PLZF has nine *kriippel*-like zinc fingers and an N-terminal protein-protein interaction POZ domain (Chen, Brand et al. 1993). Its homology to other *kriippel*-related transcription factors (El-Baradi and Pieler 1991) supports its role as a potential transcriptional regulator. In contrast to PML, the PLZF gene has a tissue specific and evolutionary conserved pattern of expression which implies an important role in a number of developmental processes. During normal haematopoiesis, PLZF expression is primarily restricted to the immature progenitor cells but does appear to be down regulated during growth factor or RA induced myeloid differentiation (Chen, Zelent et al. 1993; Reid, Gould et al. 1995). This implies an important function for PLZF in the regulation of biological processes in both progenitor cells and myelopoiesis. The down regulation of PLZF may be necessary to allow for the subsequent differentiation of myeloid cells. However a block in the differentiation of myeloid precursors by the aberrant expression of PLZF has yet to be demonstrated experimentally. During embryonic development PLZF is expressed in the AGM region (Reid, Gould et al. 1995) which gives rise to haematopoietic progenitors and also in the developing limb buds which, as discussed in Section 1.1.4, reflects the overlap in transcription factor usage in these two developmental processes. In addition, in the developing nervous system PLZF has a very restricted but dynamic expression pattern within the rhombomeres of the hindbrain suggesting a role for PLZF in the regionalisation of the vertebrate hindbrain during development (Cook, Gould et al. 1995). Recent PLZF knock out experiments have revealed abnormalities in hindlimb development and spermatogenesis. The hindlimb abnormalities reflect the PLZF expression pattern as they affect the more distal structures of the hindlimb and result in the duplication, fusion or absence of digits (Hawe, Soares et al. 1996).

The speckled nuclear localisation of PLZF (see section 1.1.3) is altered upon fusion with RAR α . Recent experiments have shown that this altered nuclear localisation is mediated by the POZ domain of PLZF (Reid, Gould et al. 1995). As for the PML-RAR α chimera, PLZF-RAR α can form heterodimers with RXR as well as POZ domain mediated homodimers and heterodimers with wild type PLZF. The latter interaction could prevent homodimerisation of wild type PLZF or its heterodimerisation with, as yet, unidentified partner POZ domain proteins (Dong, Zhu et al. 1996). Given the absence of any other obvious functional motifs within PLZF apart from the DNA binding zinc fingers, it remains to be seen what role, if any, the POZ domain plays in the un-responsiveness of t(11;17)(q23;q21) APL to ATRA treatment.

1.2.2 Type II translocations and mutations of the BCL-6 gene in NHL

Non-Hodgkin's lymphomas (NHL) include a group of neoplasms which share a common target tissue, lymphoid cells, but are characterised by a high degree of biological and clinical heterogeneity. Most NHLs however derive from mature B cells and are classified according to their degree of clinical aggressiveness, their stage of differentiation and their pattern of growth (Magrath 1990). Structural alterations disrupting the BCL-6 gene have been identified in two subtypes of NHL; diffuse large cell lymphomas (DLCL) and follicular lymphomas (FL) (LoCoco, Ye et al. 1994) DLCL is the most frequent and the most lethal human lymphoma and accounts for ~40% of initial NHL diagnoses and is often the final stage of progression of FL (Magrath 1990).

The BCL-6 gene was initially cloned from the breakpoint of a type II, t(3;14), translocation associated with DLCL (Ye, Lista et al. 1993; Kawamata, Miki et al. 1994). The 3q27 BCL-6 locus has subsequently been shown to have over 13 different translocation partners including all three immunoglobulin (Ig) loci (Bastard, Deweindt et al. 1994). In each translocation, the break points in the BCL-6 gene map to the same 4 kb region, spanning the 5' flanking region, the first exon and the first intron, resulting in the truncation or complete removal of the putative regulatory region of BCL-6 whereas its coding region is always left intact. These type II translocations are therefore thought to contribute to NHL disease progression by the deregulated over expression of BCL-6 (Ye, Chaganti et al. 1995). However, more than 73% of DLCL patients

and 47% of FL patients display structural alterations of the BCL-6 gene, some of which are independent of chromosomal rearrangements. In these cases, the BCL-6 gene is altered by somatic point mutations in the 5' non coding region. Again the coding region is unaffected and the mutations are thought to lead to the inappropriate regulation and expression of the gene (Migliazza, Martinotti et al. 1995).

BCL-6 is a 95 kilodalton phosphoprotein comprising six C-terminal *Krüppel*-like zinc fingers and an N-terminal POZ domain (Ye, Lista et al. 1993) In adult tissues, BCL-6 is expressed ubiquitously, predominantly at low levels. However, in haematopoiesis, BCL-6 is primarily expressed within the B cell lineage. There is a very specific window of expression of BCL-6 in the B cells of the germinal centres (see below) which is absent from both precursor (pre) B cells and the more differentiated plasma cells and memory cells (Onizuka, Moriyama et al. 1995, Cattoretti, Chang et al. 1995). This expression pattern has been confirmed at the protein level and knock-out studies have shown that BCL-6 is essential for germinal centre formation (Dent, Shaffer et al. 1997).

BCL-6 has therefore been implicated in B cell development. Prior to mediating the critical antibody response to a foreign antigen, B cells undergo an antigen independent developmental stage in the bone marrow which involves the generation of cytoplasmic immunoglobulins (Ig) by gene rearrangement. Whilst progenitor (pro) B cells make functional heavy chain rearrangements, pre-B cells are associated with the subsequent light chain rearrangements. Upon completion of these rearrangement, the pre-B cells differentiate into immature B cells expressing surface Ig receptors which migrate to the secondary lymphoid organs where they are exposed to antigen. The germinal centres are dynamic structures in which antigen primed immature B cells undergo complex changes including rapid proliferation, hypermutation of the Ig variable region and Ig isotype class switching. Upon completion of these changes, those B cells with inappropriate or self reacting surface Igs will undergo apoptosis whilst the foreign antigen specific B cells will differentiate into either mature plasma cells or memory cells (Cooper 1987).

The identification of a response element has led to the demonstration that BCL-6 is a site specific repressor. Subtle variations on the same BCL-6

consensus binding site sequence have been identified by three independent groups (Deweindt, Albagli et al. 1995; Chang, Ye et al. 1996; Seyfert, Allman et al. 1996). One of these studies used the entire BCL-6 open reading frame to select this binding site which supports the view that this consensus sequence is the *in vivo* response element (Chang, Ye et al. 1996). The repressive activity of BCL-6 has been mapped to two non-overlapping regions of the protein which includes the N-terminal POZ domain and a highly charged proline rich region; both regions appear to be necessary for maximal repression (Cattoretti, Chang et al. 1995). All investigators have experienced an appreciable endogenous repressive effect on the BCL-6 response element in the absence of transfected BCL-6. This could well be due to endogenous BCL-6, however the possibility of an additional endogenous protein binding to the same response element has not been ruled out. If this were the case, the addition of BCL-6 might augment the repressive effects of the endogenous protein through a heterodimeric interaction. A repressive effect could not be demonstrated in a yeast system (Seyfert, Allman et al. 1996) which further supports the view that repression requires additional protein interactions, possibly with other POK family members, which are not conserved in yeast.

As a result of this repressive transcriptional activity for BCL-6, it has been suggested that BCL-6 acts as a silencer within the germinal centres which needs to be subsequently down regulated to permit further, post germinal centre, B cell development. The differentiation stage of DLCL cells corresponds to that of the germinal centre B cells and it is proposed that the aberrant expression of BCL-6, occurring as a result of chromosomal translocations or mutation, could contribute to disease by preventing the subsequent apoptosis or differentiation of B cells (Ye, Lista et al. 1993; Cattoretti, Chang et al. 1995; Chang, Ye et al. 1996). However, the proposed silencing role for BCL-6, particularly with respect to apoptosis, has yet to be confirmed experimentally. Given the complete absence of germinal centres in BCL-6 knockout (-/-) mice (Dent, Shaffer et al. 1997), rather than acting as a silencer, it is conceivable that BCL-6 actively promotes some of the complex changes associated with germinal centre development, such as proliferation. This could be achieved through the repression of negative regulators of proliferation; alternatively BCL-6 may prove to exert an activatory effect, *in vivo*, through its heterodimerisation with other as yet unidentified POK proteins.

Clearly the identification of a potential target gene for BCL-6 would shed some light on its role in germinal centre activities. The putative binding site for BCL-6 has been found to be very homologous to the GAS sites that are the target sequences for the cytokine induced STAT (Signal Transducers and Activators of Transcription) DNA binding proteins. BCL-6 has recently been shown to bind and repress the transcription of a downstream target gene of the, IL-4 induced, STAT-6 protein, in vivo (Dent, Shaffer et al. 1997). IL-4 is produced by activated T cells, basophils and mast cells and has been shown to mediate the proliferation of B cells within the germinal centres which have IL-4 receptors on their surface. IL-4 has also been implicated in isotype class switching, the production of MHC class II molecules and the auto regulatory production of IL-4 receptors in B cells (Paul and Ohara 1987). IL-4 stimulation is thought to be mediated through at least two signalling pathways; one of which involves the STAT-6 protein (Kotanides and Reich 1993; Hou, Schindler et al. 1994) and the other the IRS-2 (Insulin Receptor Substrate) protein (Keegan, Nelms et al. 1994). It is conceivable, therefore, that the presence of BCL-6, in the B cells of the germinal centres, may act to inhibit the STAT-6 mediated responses in favour of those mediated by IRS-2, such as the induction of proliferation (Keegan, Nelms et al. 1994). However, recent knock out studies, have demonstrated that the STAT-6 protein mediates all of the IL-4 responses in B cells (Kaplain, Schindler et al. 1996). At present, therefore, it is hard to reconcile a regulatory role for BCL-6 within the germinal centre B cells with its repressive effects on STAT-6 responsive genes.

1.3 The identification of regulatory heterodimeric POK protein complexes

Haematopoiesis is a very tightly regulated process; to ensure the appropriate restriction and commitment of cells, the exact timing of the activity of regulatory proteins is critical. Therefore the differential formation of protein complexes, mediated through the POZ domain, could contribute to the tight regulation of BCL-6 and PLZF POK protein activities. As discussed in section 1.2, the over expression of c-Myc in Burkitts lymphoma will have dramatic effects on the Myc-Max regulatory network. By the same means, the POZ domain of both PLZF and BCL-6 could contribute to APL and NHL disease progression, respectively. The heterodimeric interaction between PLZF-RAR α and PLZF might prevent the latter from interacting with its, as yet

uncharacterised, normal heterodimeric partners, leading to the disruption of highly concentration sensitive regulatory processes. Likewise, the over expression of BCL-6 could lead to its inappropriate interaction with partner POZ proteins. With this in mind, the unpublished reports of an *in vivo* heterodimeric interaction between PLZF and BCL-6 are intriguing. However, within the haematopoietic system, these two proteins have non overlapping patterns of expression which calls into question the functional significance of this *in vivo* interaction.

The object of this investigation was to identify both biologically significant heterodimeric POZ domain partner proteins for PLZF and closely related POK proteins which might be important developmental regulators or potential targets for oncogenic transformation. It was hoped that this would contribute towards an understanding of the regulatory networks that might be disrupted as a result of the t(11;17) translocation in APL. This study has led to the isolation of four POZ proteins and the characterisation of a heterodimeric partner for BCL-6; a novel POK protein termed LRF. The critical regions mediating this *in vivo* heterodimeric interaction have been identified in this study and a model for POK protein mediated transcriptional regulation has been proposed.

Chapter 2: Materials and methods

2.1 The identification of PLZF related POZ domain proteins

2.1.1 Zoo blot procedure

To identify potential cross species counterparts of mPLZF it was necessary to perform a low stringency screen of a southern blot containing genomic DNA from different species. The stringency of a screen can be controlled by varying the temperature of hybridisation of the radioactive probe as well as the extent of post hybridisation washing; the higher the temperature and the more intensive the washes, the more stringent the screen. In addition, as utilised in this screen, lowering the formamide content of the hybridisation solution from the normal 50% value can have the same effect as lowering the temperature of hybridisation.

Genomic DNA preparation

Both BamHI and HindIII (40Units [U], New England Biolabs [NEB]) restriction enzyme [RE] digestions were performed on genomic DNA (10µg each) derived from human, mouse, chicken, drosophila and sarcomyces [S] cervisiae (Clonetech) in the presence of 1x RE buffer (NEB) in a 200µl total volume overnight [o/n] at 37°C. To check for complete digestion 10µl of each digest was added to 1x gel loading buffer (6x: 0.25% bromophenol blue, 0.25% xylene cyanol FF, 30% glycerol) and run out on a 1.4% agarose electrophoresis gel (1.4% agarose [gibco], 1x TAE [50x: 242mg/ml Tris base, 57.1% glacial acetic acid, 18.61mg/ml EDTA, pH 8.0]) in 1x TAE running buffer together with a DNA ladder (Stratagene) as a size marker. The gel was run at 80-100 volts until the faster migrating bromophenol blue dye had run 3/4 of the gel's length, then stained in 1µg/ml ethidium bromide for 10min and the DNA visualised under long wave ultra violet [UV] light. If digestion was incomplete, the genomic DNA was digested for longer in a higher concentration of the appropriate RE. The completely digested genomic DNA solutions were then extracted twice with an equal volume of phenol/chloroform isoamyl alcohol (25:24:1 ratio, Sigma). The DNA was precipitated by first mixing with 0.1 volumes of 3M sodium acetate and 3 volumes of absolute EtOH, then incubating for 30min at -80°C and finally sedimenting by centrifugation at 13000rpm for 20min at 4°C. The resultant pellets were washed in 75% EtOH, dried under vacuum and resuspended in 20µl of TE pH 8.0 (0.5M EDTA, pH8.0, 1M Tris-HCl, pH 8.0).

Southern blot

The digested genomic DNA and a DNA ladder marker were individually added to 1x gel loading buffer and run out on a large scale (200ml) lower percentage agarose gel (0.8%, see above) at 80 volts until the bromophenol blue dye had run 3/4 of the gel's length (~3hr). The gel was then stained as described above and visualised under UV. To allow for efficient transfer of the high molecular weight genomic DNA, the gel was incubated in 200ml of 0.1N HCl for 10min to partially depurinate the DNA. To denature the double stranded DNA, the gel was washed twice for 15min each in 200ml of denaturing solution (1.5M NaCl, 0.5M NaOH). This was followed by two 15min washes in 200ml of neutralising solution (1.5M NaCl, 0.5M Tris-HCl pH 7.4). Finally the gel was washed once in 20xSSC (3M NaCl, 0.3M sodium citrate, pH 7.0).

To transfer the DNA from the gel to BAS85 reinforced nitrocellulose membrane (Schleicher and Schuell) a standard capillary blot was set up. 4 pieces of Whatman 3MM paper slightly larger than the area of the gel soaked in 20xSSC were placed on saran wrap, the gel was added on top followed by the nitrocellulose membrane (pre-soaked first in H₂O for 10min then in 20xSSC). To prevent any capillary by-pass the edges of the gel were sealed with saran wrap. Three more 20x SSC soaked pieces of Whatman 3MM paper were placed on top of the of the nitrocellulose followed by a 5-7 cm thick stack of paper towels and a light weight. The DNA was transferred o/n and then the DNA bound nitrocellulose was air dried and baked in an oven under vacuum at 80°C for 2hr.

Low stringency hybridisation with radio-labelled DNA probes

The dried nitrocellulose filter was incubated in 30ml of prehybridisation solution (30% formamide (molecular biology grade, NEB), 5x SSPE [20x: 3.6M NaCl, 0.2M sodium phosphate, pH 7.7, 20mM EDTA], 5x Denhardt's [50x: 10mg/ml Pharmacia Ficoll Type 400, 10mg/ml Sigma Bovine Serum Albumin Fraction V, 10mg/ml polyvinyl pyrrolidone], 1mg/ml salmon sperm (ss)DNA [10mg/ml: 1 g of Sigma ssDNA Type III sodium salt dissolved in 100ml of H₂O and autoclaved], 0.1% sodium dodecyl sulphate solution [SDS]) in a hybridisation bag placed flat on rocking platform in an incubator at 42°C for 2hr. The prehybridisation solution was then replaced with 30ml of hybridisation solution (30% formamide, 5x SSPE, 1x Denhardt's solution,

0.2mg/ml ssDNA) containing 2×10^6 cpm/ml of a purified ^{32}P -labelled DNA probe (see below) that had been denatured for 10min at 100°C and the membrane was incubated at 42°C , with rocking, o/n.

Probe labelling

Linear DNA without zinc finger sequences was prepared by digesting mPLZF BsK (Cook, Gould et al. 1995) with both EcoRI and SphI as described above for the genomic DNA digestion. The digested DNA was run out on a gel and the appropriate DNA fragment excised from the gel by gene cleaning (see section 2.1.3). The DNA was dissolved in H_2O to a concentration of $25\mu\text{g}/\text{ml}$, denatured by heating to 100°C for 10min and then chilled on ice. Using the rapid multiprime DNA labelling kit reagents (Amersham), a ^{32}P -labelling reaction was set up on ice: $10\mu\text{l}$ mPLZF digested DNA, $10\mu\text{l}$ Buffer (solution 1), $5\mu\text{l}$ primer (solution 2), $5\mu\text{l}$ $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$ at 8000 ci/mmol and 2U klenow enzyme in a total volume of $50\mu\text{l}$. The reaction mix was incubated at rm temp for 3hr for optimal labelling. The probe was then purified through sephadex G50 beads (Pharmacia). Autoclaved glass wool was placed at the neck of a 10ml pipette and sephadex beads were added from the top of the pipette alternating with TE buffer pH 8.0 until the beads were compacted to the top of the column. $100\mu\text{l}$ of diluted radiolabelled probe was then added drop by drop to the top of the beads and then chased through the column with TE buffer pH 8.0. The first peak of radioactivity was collected as purified probe and its activity determined by automated scintillation counting.

Post hybridisation washes

It was important that the stringency of these washes was sufficiently high to remove most of the background signal without washing away the cross hybridising signals. The blot was removed from the hybridisation bag and rinsed in solution A (2xSSPE, 0.03% sodium pyrophosphate solution NaPPi, 0.1% SDS). This solution was then replaced with fresh solution A and left shaking for 15min at rm temp. This was repeated if the radioactive signal was still sufficiently high and then the blot was washed in pre-heated solution B (1x SSPE, 0.03% NaPPi, 0.1% SDS) and incubated, with shaking, at 55°C for 15min. Subsequently, the blot was damped dry on Whatman 3MM paper, wrapped in saran wrap and exposed to X-OMAT™ AR Kodak scientific imaging film between two intensifying screens inside an exposure cassette, o/n at -80°C . Upon developing the film (automated), owing to a high degree

of cross reactivity, the blot was washed again in solution B at 60°C for 30min to remove the high background signal and reexposed to Kodak film for 1 week at -80°C.

2.1.2 cDNA Phage Library screen

Both a chicken heart cDNA λ gt10 library (Clontech) and a mouse 10 dpc embryonic cDNA λ gt10 library (A gift from P. Chambon, Institut de Genetique et de Biologie Moleculaire et Cellulaire, Illkirch, France) were screened using the same 32 P-labelled PLZF and the same conditions of stringency as described in section 2.1.1.

Plating the phage libraries

To ensure that the cDNAs within each library were fully represented in this screen, 24 plates (150mm) each containing approximately 40,000 plaques were required for both λ gt10 libraries. Each titred library was appropriately diluted in λ dil solution (10mM Tris-HCl pH 7.5, 10mM $MgSO_4$) added to 300 μ l of plating bacteria (C600hfl bacteria grown o/n at 37°C in 100ml of L broth, pelleted, resuspended in 50ml λ dil solution and stored at 4°C) and incubated in a 37°C water bath for 15min. 7ml of λ B top agar (0.025% NaOH 10N, 2.5mg/ml NaCl, 2.6mg/ml $MgSO_4$, 5mg/ml yeast extract, 10mg/ml bactotryptone, 0.6% agarose), incubated in a 47°C water bath, was added individually to each aliquot of phage DNA in the 37°C water bath, mixed by shaking and poured immediately on to prewarmed 150mm LB plates (0.025% NaOH 10N, 10mg/ml NaCl, 5mg/ml yeast extract, 10mg/ml bactotryptone, 15mg/ml Bactoagar; 50ml/plate). When the plates were dry, they were placed up side down in a 37°C incubator until plaques started to form (approx. 4hr) and then they were removed from the incubator and placed up side down in a sealed polystyrene box, at rm temp o/n.

Replicating plaques on to nitrocellulose

Once the plates were removed from the polystyrene box, appropriately numbered 137mm BAS85 nitrocellulose filters (Schleicher and Schuell) were placed on to the bacteriophage plaques, orientated with a blank ink soaked needle, peeled off immediately from the plate and placed plaque side up on to Whatman 3MM paper. A replica lift was performed by placing another labelled nitrocellulose filter on to the same plate, orientating it to the same ink

spots as the first filter and leaving it for 20 sec before peeling off and placing plaque side up on to Whatman 3MM paper.

Denaturing plaques

Three baths were made up containing Whatman 3MM paper saturated with denaturing, neutralising and 2xSSC solutions respectively. The filter papers were placed plaque side up on to each of the three saturated 3MM papers successively, for 5min each, and left to dry, plaque side up, on Whatman 3MM paper for 30min. All the filters were then placed between sheets of Whatman 3MM paper and dried, under vacuum, in an oven at 80°C for 2hr. The dried filters were placed into a plastic hybridisation bag and washed with 6xSSC solution prior to hybridisation and subsequent washes which were performed under the same low stringency conditions described in section 2.1.1.

Secondary screen

A secondary screen was necessary in order to isolate individual positive plaques. The plaques that were positive on both the original and replica blots were picked from the corresponding area on the plate. The picked top agar was placed in 400µl of λdil and left for 2 hr at rm temp; then 2µl of chloroform was added to kill the plating bacteria and the phage stocks were stored at 4°C. The appropriate concentration of phage stock was then plated out to give discrete individual plaques and the plaques were replicated on to nitrocellulose, hybridised and washed as described in section 2.1.1. The positive plaques were picked as described above and stored at 4°C prior to DNA isolation.

Isolation of phage DNA

100µl of secondary picked phage stocks were incubated with 50µl of plating bacteria for 15min at 37°C then 5ml of NZYCM medium (10mg/ml NZ amine, 5mg/ml NaCl, 5mg/ml yeast extract, 1mg/ml casamino acids, 2mg/ml MgSO₄, 0.025% NaOH 10N) was added and the phage stocks were grown o/n at 37°C. 300µl of chloroform was added and the culture was agitated for 15min to kill off the plating bacteria. The supernatant was then removed from the chloroform and centrifuged for 10min at 5000rpm to pellet the bacterial debris. To destroy any remaining bacterial nucleic acids, without disrupting the phage DNA, the supernatant was poured into a fresh tube and after adding 5µl of RNase A (10mg/ml, Boehringer Mannheim) and

12.5µl of DNase1 (10U/µl Boehringer Mannheim), was incubated at 37°C for 30min . The phage was then precipitated from solution by the addition of 5ml of PEG solution (20% PEG 6000, 2M NaCl, 10mM Tris-HCl pH 7.5, 10mM MgSO₄), incubation on ice for 1hr and centrifugation at 5000rpm for 20min at 4°C. To disrupt the phage protein-nucleotide complexes, the pellet was resuspended in 0.5ml λdil, containing 0.1% SDS and 3mM EDTA, and heated to 68°C for 15min. The phage DNA was then extracted with Tris saturated phenol, phenol/chloroform isoamyl alcohol and finally chloroform alone. To precipitate the DNA an equal volume of isopropanol was added and mixed and the DNA pelleted by centrifugation at 13000rpm for 5min at 4°C. After washing with 75% EtOH the phage DNA was resuspended in 100µl of TE buffer pH 8.0 supplemented with 20µg/ml of RNase A.

2.1.3 Subcloning phage DNA into Bluescript (Bsk)

The phage DNA was subcloned into BsK (SK⁺, Stratagene) in order to facilitate both the initial sequencing of the clones Using T7 and T3 primers and the subsequent construct design (see section 2.3.5)

Phage DNA digestion

50µl of phage DNA was digested with 5µl of EcoRI RE in a 100µl volume at 37°C for 1hr and then run out on a 2% agarose gel, as described in section 2.1.1. The resultant digested insert was cut out of the gel and gene cleaned (see below). In addition, to confirm that the phage DNA was positive, a proportion of the digest was southern blotted and reprobbed under the same conditions of stringency as for the library screen (see section 2.1.1).

Gene cleaning

The DNA fragment was excised from the gel by using the GeneClean^R II kit. The cut out band was incubated in three volumes of Sodium Iodide at 50°C for 5min. 5µl of glassmilk was then added to the DNA solution and left at 4°C for 10min to allow the DNA to bind efficiently to the glass beads. The glass beads were then pelleted by centrifugation at 13000rpm for 5sec and washed three times in New Wash. The DNA was eluted from the glass beads by adding 10µl of H₂O and incubating at 50°C for 3min. The glass beads were pelleted by centrifugation at 13000rpm for 30sec and the supernatant containing the DNA removed from the pellet. The concentration of the eluted

DNA was determined by absorbance analysis at 260nm and its integrity verified by a 1.4% agarose gel electrophoresis.

Bsk vector preparation

10µg of Bsk plasmid DNA was digested with 40U of EcoRI in a 200µl volume for 1hr at 37°C. The EcoRI enzyme was inactivated by heating at 70°C for 5min. The digested plasmid was then extracted with phenol/chloroform isoamyl alcohol and precipitated in 0.1 volume of sodium acetate and 3 volumes of EtOH by centrifugation at 13000rpm. The resultant pellet was resuspended in 10µl of TE pH 8.0. Owing to the single restriction enzyme digestion the ends of the linearised vector were dephosphatated to prevent self ligation by incubation with 2U of CIAP (calf intestinal alkaline phosphatase, Clontech) and 10µl of 10x CIAP buffer (Clontech) in a 100µl volume for 1hr at 37°C. The vector solution was then extracted with phenol/chloroform isoamyl alcohol and EtOH precipitated, as above. The vector DNA concentration was determined by absorbance at 260nm and its integrity verified by a 1.4% agarose gel electrophoresis.

Ligation reaction

25ng of Bsk vector, 1x ligase buffer (NEB), 4U ligase enzyme (NEB), 30-50ng of gene cleaned insert were combined in a 20µl reaction volume and incubated for 3hr at rm temp. In addition, both a self ligation control reaction which did not include insert DNA and an uncut vector control reaction, without ligase, were setup.

Transformation

For each transformation reaction, 100µl of supercompetent XL-1 Blue E. coli cells (Stratagene) were aliquoted into pre chilled 15ml polypropylene tubes. To increase the transformation efficiency, 1.7µl of 14.2M β-mercaptoethanol was added to each tube and swirled gently at 2min intervals for 10min. 1-5µl of the respective ligation reactions were then added to the cells and incubated on ice for 30min prior to being heat shocked at 42°C for 45 sec followed by a further incubation on ice for 2min. 0.9ml of preheated L broth (10mg/ml bactotryptone, 5mg/ml NaCl, 5mg/ml yeast extract) was added to each tube which were then incubated at 37°C for 1hr with shaking. 5-200µl of the transformed DNA was then plated out, using a sterile spreader, onto LB-ampicillin plates (L broth, 1.5% bactoagar, 50µg/ml ampicillin) and grown o/n at 37°C.

Plasmid DNA preparation

Individual colonies were picked from the transformed plates and grown up o/n at 37°C in 5ml of L broth containing 50µg/ml ampicillin (Sigma). To miniprep the DNA, 1.5ml of each of the bacterial cultures was centrifuged at 13000rpm for 30 sec and the medium removed. Each pellet was resuspended in 100µl of solution A (50mM glucose, 10mM EDTA, 25mM Tris-HCl, pH 8.0) and incubated at rm temp for 5min and then 200µl of lysis solution B (0.2 N NaOH, 1% SDS) was added and the tubes mixed thoroughly and incubated on ice for a further 5min. 150µl of ice cold potassium acetate solution (0.6mg/ml potassium acetate adjusted to pH 5.5 with glacial acetic acid) was then added and the lysates neutralised by incubating on ice for 5min. The tubes were then centrifuged at 13000rpm for 5min at 4°C and the supernatants transferred to new tubes and extracted with an equal volume of phenol chloroform isoamyl alcohol. The DNA was precipitated with 1ml of EtOH per tube and centrifuged for 10min at 13000rpm. The DNA pellets were washed in 75% EtOH, air dried and resuspended in 100µl of TE buffer, pH 8.0, supplemented with 10µg/ml RNaseA.

The minipreps were digested with EcoRI, as described for the phage DNA digestions (see above) and the presence of the appropriate insert corresponding to the size of the gene cleaned product was determined by running the digested products on a 1.4% agarose gel. Minipreps containing the correct size insert were then maxipreped using the Qiagen plasmid purification kit (Qiagen inc.). The original corresponding bacterial cultures were each added to 200ml of L Broth containing 50µg/ml ampicillin and large scale cultures were grown up o/n. The bacteria was then pelleted by centrifugation at 6000rpm and lysed in the same way as the minipreps by using 10ml of each of the appropriate Qiagen kit solutions. The DNA solution was purified by passing through equilibrated Qiagen 500 tip columns. The DNA, which is retained in the columns, was washed twice with 30ml of the kit wash buffer and then eluted from the columns with 15ml of the kit elution buffer. The DNA was precipitated in 0.7 volumes of isopropanol and pelleted by centrifugation at 13000rpm. The resultant pellet was washed in 75% EtOH, air dried and resuspended in sterile TE buffer, pH 8.0. Subsequently, the concentration of the isolated DNA was determined by absorbance analysis at 260nm and 1µg was digested and run on a 1.4% electrophoresis gel to confirm the existence of the appropriately sized insert.

Sequencing

The sequence of the newly isolated clones was determined by automated sequencing using the Bsk specific T3 and T7 primers. 0.5µg of plasmid DNA was combined with 3.2pmole of either of the two primers and 8µl of terminator ready reaction mix (Perkin Elmer) in a 20µl reaction volume. The sequence was then amplified by polymerase chain reaction (PCR). The PCR conditions included a 5min denaturation at 96°C, followed by 25 cycles of 96°C for 30 sec, 50°C for 15 sec and 60°C for 4min. The PCR products were precipitated in 95% EtOH for 10min on ice and pelleted by centrifugation at 13000rpm for 20min at 4°C. The pellets were washed in 75% EtOH and air dried before being loaded on to an automated sequencing gel (core facility). Subsequent sequencing of the isolated clones involved the use of sequence specific primers derived from the initial T7 and T3 sequencing.

2.1.4 POZ protein isolation by degenerate RT-PCR

RT-PCR involves the generation of cDNA by reverse transcription, from a source of total or poly adenylated [pA⁺] RNA. The resultant cDNA is then used as a PCR template to amplify cDNAs using specific primers. In this study, degenerate oligonucleotide primers derived from the POZ domain were used and the resultant PCR products were then subcloned into a TA cloning vector (see below) and screened using an alternative degenerate, POZ specific, probe.

The positions of the degenerate primers used to identify PLZF related clones are indicated on Figure 1. The sequence of these primers are indicated below (N = A/G/C/T, V = A/G/C, B = T/G/C, H = A/C/T, D = A/T/G, Y = C/T, S = C/G, R = G/A, W = A/T, K = G/T, M = A/C). The primers were designed with the highest degree of degeneracy corresponding to the unconserved amino acids (indicated in white boxes in Figure 1, section 1.1.4). Selected amino acids are indicated below each mutated nucleotide triplet, in a single letter code.

Primer 1: 5' NRS NNN NAR NNN NSC NGT RTA
 V X L X A T Y

Primer 2: 5' GGN NNN YTS TGY GAY GTB GTB AT
 G X L C D V V I

Primer 3: 5' C NGT RTA NGC RWA YTC CAR VA
 T Y A Y E L I

Primer 4: 5' TTY MRN GCY CAY MRR WNN GTS CTK GC
 F H A H R T V L A

Primer 1 was derived from the most C-terminal end of the POZ domain and was used to prime the generation of multiple reverse transcribed RNAs from the pooled 13.5 and 8.5 mouse pA⁺ RNA isolated as described in section 2.2.4. Primer 2 was derived from the most N-terminal region of conservation of the POZ domain and primer 3 was nested with primer 1. Primer 2 and 3 are the respective 5' and 3' primers used for the PCR reaction on the reverse transcribed RNA template. Primer 4 was used to screen the resultant PCR products. It was hoped that, at this level of degeneracy, most of the PCR products generated would contain a POZ domain.

Reverse transcription

1µg of pA⁺ RNA was added to 100ng of primer 1 and made up to a final volume of 27.6µl in diethyl pyrocarbonate (DEPC) treated H₂O. The tube was incubated at 70°C for 5min and then transferred to ice. Following this, 1x reverse transcription buffer (5x: 250mM Tris-HCl, pH 8.3, 375mM KCl, 15mM MgCl₂), 0.5mM dNTPs (Pharmacia), 1mM DTT (Gibco-BRL), 40U of RNasin (Promega) and 200U of Molony Murine Leukaemia virus reverse transcriptase (Gibco-BRL) were added to the cooled RNA solution in a final volume of 40µl. The solution was mixed by tapping the side of the tube and incubated at 37°C for 45min. The resultant cDNA was stored at -20°C, prior to PCR.

PCR

2.5µl of the reverse transcribed cDNA was used as template in a 100µl reaction volume containing 0.2mM dNTPs (Pharmacia), 1x PCR buffer (10x: 500mM KCl, 15mM MgCl₂ 200µg/ml bovine serum albumin), 800ng of both primer 2 and primer 3, and 2U of Taq DNA polymerase (Perkin Elmer), added in the order listed. The reaction was transferred immediately to a pre-

heated Perkin Elmer 2400 thermal cycler. The cDNAs were amplified using a program of 25 cycles: 95°C for 30 sec (denaturation), 54°C for 1min 30sec (annealing), 72°C for 3min (extension). The last cycle was followed by a final extension at 72°C for 15min and the reaction was then stored at 20°C.

Screening with an oligonucleotide probe

To confirm that the PCR strategy had amplified POZ domain sequences, 10 μ l of the PCR reaction was run out on a 1% agarose electrophoresis gel and southern blotted, as described in section 2.1.1. The membrane was incubated for 2hr in 10ml of oligonucleotide prehybridisation solution (5x SSPE, 5x Denhardt's, 1mg/ml ssDNA, 0.1% SDS) at 50°C, then the solution was replaced with hybridisation solution (5x SSPE, 1x Denhardt's, 1mg/ml ssDNA) containing 2 μ l of ³²P-labelled primer 4 and incubated with rocking o/n at 50°C. The oligonucleotide ³²P-labelling reaction included 5pmole of primer 4, 1x kination buffer (10x: 500mM Tris-HCl, pH 7.6, 1mM spermidine, 1mM EDTA, 100mM MgCl₂, 50mM DTT), 5 μ l ³²P- γ -ATP (Specific activity > 5000 Ci/mmole: Amersham) and 10U of T4 polynucleotide kinase (NEB) in a final volume of 10ml and was incubated at 37°C for 30min. Post hybridisation, the membrane was washed twice for 15min with agitation in 2x SSPE, 0.1% SDS. The blot was then exposed, between two intensifying screens, to Xray film (X-OMAT™ AR, Kodak). Upon developing the film, as high back grounds were apparent, the blot was washed more extensively, in the same wash solution as above but at 45°C, before being exposed to film again for 4hr at -80°C. A successful RT-PCR procedure was confirmed by the existence of hybridised bands of ~190bp corresponding to the anticipated size of an amplified POZ domain.

Subcloning of PCR fragments

As described in section 2.1.3, to subclone the PCR products, the remainder of the PCR reaction was run out on a 1.4% agarose gel (after the reaction volume had been decreased by EtOH precipitation and resuspension in 20 μ l of H₂O). The region of gel corresponding to ~190bp was cut out from the gel, gene cleaned and the quantity of DNA determined. 30ng of the gene cleaned PCR product was ligated into 50ng of the Invitrogen TA cloning vector, (pCR™II). This 'TA' ligation reaction takes advantage of the non-template dependant activity of Taq DNA polymerase which adds single deoxyadenosines to the 3' end of all duplex PCR products; therefore, without any restriction enzyme digestion, these A-overhangs can be ligated into the pCR™II vector which

contains single 3' T overhangs. 1µl of the ligation reaction was then transformed into competent bacteria.

The resultant colonies were picked (with autoclaved tooth-picks) in duplicate on to orientated LB agar plates (containing 50µg/ml ampicillin). The colonies were grown o/n and transferred on to nitrocellulose membranes and denatured as for the replication of phage plaques, see section 2.1.2. For each replicated filter, one filter was probed with ³²P-labelled primer 4 and the other with a nondegenerate, mPLZF POZ domain specific, ³²P-labelled primer (T_m of 60°C, hybridised at 55°C, washed at 50°C: high stringency) as described above. The resultant, primer 4 positive, mPLZF negative, colonies were picked, grown up in 5ml of LB (containing 50µg/ml ampicillin), miniprep and sequenced as described in section 2.1.3.

2.1.5 Whole Mount In Situ Hybridisation

This non radioactive method of detecting the expression pattern of a gene of interest during embryonic development requires the generation of a digoxigenin labelled anti-sense RNA probe which is hybridised directly to the permeabilised whole embryos. The embryos are then exposed to a digoxigenin specific antibody which is conjugated to an alkaline phosphatase (AP). Subsequent exposure to the appropriate substrates, 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and nitroblue tetrazolium salt (NBT), results in an AP catalysed oxidation reaction which generates a blue precipitate at the sites where the antibody has cross reacted with the digoxigenin labelled probe. Therefore, provided the embryos have been extensively washed, to prevent probe and antibody trapping, the blue colouration will correspond to the site of expression of the gene. The degree of trapping, as compared to a genuine signal, can be determined by comparing the signal generated from a digoxigenin labelled sense RNA probe (negative control) which should, ideally, be completely devoid of a blue signal.

A number of modifications have been made to the original protocol (Wilkinson 1992). To minimise high background, the digoxigenin labelled probes were purified through spin columns to ensure the complete removal of unincorporated digoxigenin UTP. In addition the antibody was incubated in blocking agent (Boehringer Mannheim) instead of sheep serum. To

minimise non specific signals, due to trapping, more extensive washes were performed post both hybridisation and antibody incubation.

Probe synthesis

Antisense and sense digoxigenin labelled probes were synthesised from linearised Bsk plasmids of clones 1-4. (Transcription mix: 1x transcription buffer [Promega], 0.01M DTT, 10% ATP/CTP/GTP/dig-UTP nucleotide mix [Boehringer Mannheim], 1µg linearised plasmid, 50U ribonuclease inhibitor [Promega], 10U T7/T3 RNA polymerase [Promega]). Post synthesis, RNA probes were diluted to 100µl and purified through Pharmacia microspin S400 HR columns and EtOH precipitated twice. The RNA probes were then resuspended in 100µl of HPLC purified H₂O. Once the integrity of the probes had been confirmed by running 10µl of each out on a 0.8% TAE agarose gel they were stored at -70°C.

Pretreatments

The 10 dpc embryos were dissected out into Tyrode's salt solution (Sigma) to provide a similar osmolarity as in the uterus and thus prevent the embryos from invaginating. The embryos were then transferred to Costar™ netwell 74µm meshes (15mm diameter). All subsequent manipulations were performed by immersing the meshes into 3ml of the appropriate solution in Costar 12 well plates (16mm diameter wells). This ensured minimum contact with the embryos thus reducing the risk of RNase contamination. For the same reason, all solutions were made from autoclaved HPLC treated H₂O. Unless otherwise stated, all washes were performed with shaking at rm temp for 5min.

Embryos were fixed in 4% paraformaldehyde in PBSA (10mg/ml NaCl, 0.25mg/ml KCl, 1.437mg/ml KH₂PO₄, 0.25mg/ml Na₂HPO₄, pH 7.2.) with shaking at 4°C o/n. The embryos were then washed three times in PBT (PBSA and 0.1% Tween-20 [Sigma]) at 4°C. The embryos were dehydrated by washing once with 25%, 50%, 75% EtOH/PBT then twice in 100% EtOH at 4°C followed by the reverse order of washes at 4°C to rehydrate them. The embryos were washed three times in PBT at 4°C prior to treatment with 10 µg/ml proteinase K in PBT, for 6min, to permeabilise them and consequently enhance the probe penetration. To neutralise the proteinase K, the embryos were subsequently washed in freshly prepared 2mg/ml glycine in PBT and then washed three times in PBT at 4°C. To cross link the embryos, in order to

re-establish their integrity after the proteinase K treatment, the embryos were refixed in 4% paraformaldehyde/0.2% glutaraldehyde in PBT for 20min at 4°C. The embryos were washed a further three times in PBT at 4°C and bleached in 3% hydrogen peroxide in PBT for 1hr. They were then washed three times in PBT at 4°C and finally immersed in prewarmed hybridisation buffer (50% formamide (molecular biology grade, NEB), 5xSSC pH 4.5, 50mg/ml yeast RNA [Sigma], 1% SDS, 50mg/ml heparin [Sigma]) and incubated inside a formamide saturated sealed perspex box (to prevent evaporation) in a 70°C hybridisation oven for 2hr. The solution was then replaced with hybridisation buffer containing 8µl of the appropriate digoxigenin-labelled probe and placed, as for the pre-hybridisation, in a hybridisation oven at 70°C for 16hr.

Post hybridisation washes

Embryos were washed three times in solution 1 (50% formamide, 5x SSC pH 4.5, 1% SDS) at 70°C for 30min each, then once in a 1:1 mix of solution 1 and solution 2 (0.5M NaCl, 10mM Tris-HCl pH 7.5, 0.1% Tween [Sigma]) at 70°C for 10min followed by one wash in solution 2 for 10min. To remove all unhybridised RNA probe, the embryos were washed once in solution 2 containing 20µg/ml RNaseA for 1hr at 37°C. This was followed by another wash in solution 2 for 10min and a wash in solution 3 (50% formamide, 2x SSC pH 4.5) for 10min. The embryos were then washed three times in solution 3 at 65°C for 30min each and this was followed by three washes in TBST (0.15M NaCl, 0.1M Tris-HCl pH 7.5, 0.1% Tween-20). The embryos were subsequently pre-blocked in blocking solution (1.5% blocking powder [Boehringer Mannheim] in TBST) for 2hr.

Meanwhile, to preadsorb sufficient digoxigenin antibody (Boehringer Mannheim) for 5 netwell meshes, 18mg of embryo powder (see below) was first incubated in 3ml of TBST at 70°C for 30min. After the solution had cooled down on ice, 30µl of blocking solution and 9µl of digoxigenin antibody were added and the solution was incubated at 4°C, with gentle shaking, for 1hr. The preadsorbed antibody was recovered from the embryo powder by centrifugation at 13000rpm for 10min, at 4°C. The resultant, antibody containing, supernatant was added to 12ml of blocking solution and the pre-blocked embryos (see above) were incubated in this pre-absorbed dig-antibody solution for 16hr at 4°C, with shaking.

To prepare the embryo powder, 12.5-14.5 dpc mouse embryos were homogenised in a minimum volume of PBSA, incubated on ice in 4 volumes of ice-cold acetone and then pelleted at 10000g for 10min. The pellet was washed in ice-cold acetone, centrifuged again and then spread out on Whatman 3MM paper and ground into a fine powder. The dried powder was stored at 4°C prior to use.

Post antibody washes

To remove the majority of the unbound antibody, the embryos were washed three times in TBST, then five times in TBST for 1hr each. Prior to incubating with the AP substrates, the pH of the embryos was equilibrated by washing four times in NTMT (100mM NaCl, 100mM Tris pH 9.5, 50mM MgCl₂ 0.1% Tween-20) for 10min. The embryos were then incubated in the dark with shaking in NTMT containing 4.5ml NBT and 3.5ml BCIP per ml until a blue colour had developed sufficiently (about 12hr). To stop the reaction, the embryos were washed 3 times in NTMT and once in PBT, 10min each time. The stained embryos were bleached, if necessary, by washing once in 50% EtOH in PBT for 10min, once in 100% EtOH for up to 1hr, once in 50% EtOH in PBT for 10min and once in PBT for 10min. Prior to taking photographs, the embryos were preserved by washing once in 25% glycerol in PBT and stored at 4°C in 50% glycerol in PBT.

2.2 A novel POK protein: LRF

2.2.1 High stringency screen of mouse heart λ ZapII library

An adult mouse heart λ Zap II cDNA library (A gift from P. Chambon, Institut de Genetique et de Biologie Moleculaire et Cellulaire, Illkirch, France) was screened as described in section 2.1.1 but under high stringency conditions (50% formamide) and the phage were grown in XL1-Blue MRF' host bacteria (Stratagene). The ³²P-labelled probe was derived by the EcoRI digestion of the partial embryonic mLRF Bsk (clone 1) plasmid. Using a λ Zap II phage library avoided the need to subclone the phage DNA into Bsk as was necessary for the λ gt10 libraries (see section 2.1.1); using the ExAssist helper phage, the Bsk phagmids were excised directly from the λ Zap II phage vector (Altingmees, Kretz et al. 1993).

In vivo excision of Bsk phagemid

50ml of L Broth (supplemented with 0.2% maltose and 10mM MgSO₄) was inoculated in a sterile flask with a single colony of the XL1-Blue MRF' cells. In addition, 50ml of LB broth was inoculated with a single colony of SOLR cells (Stratagene). The cells were grown o/n with shaking at 37°C then pelleted at 1000g for 5min and resuspended in 10mM MgSO₄ to an optical density of 1.0 at 600nm. In a 50ml tube, 200µl of XL1-Blue MRF' cells were added to 250µl of plaque purified, mLRF positive, phage stock (derived from the secondary library screen; see section 2.1.2.) in the presence of 1µl of ExAssist helper phage (Stratagene). The mixture was incubated at 37°C for 15min then 3ml of LB broth was added and the culture incubated for 2hr at 37°C with shaking. To destroy the XL1-blue MRF' cells the tube was heated to 70°C for 15min and then pelleted at 4000g for 15min. The supernatant, containing the excised phagemid BsK packaged as filamentous phage particles, was decanted into a sterile tube. Both 10µl and 100µl of this phage supernatant were added to two 100µl aliquots of the SOLR host bacteria and incubated at 37°C for 15min. 200µl of each culture was plated out on LB plates (containing 50µg/ml ampicillin), inverted and incubated, with shaking, o/n at 37°C. Colonies appearing on the plate contained only the BsK phagemid with the cloned DNA inserts as the ExAssist helper phage is unable to replicate in SOLR cells and these cells are also resistant to λZap II phage infection. The colonies were picked and the plasmid DNA isolated and sequenced as described in section 2.1.3.

2.2.2 Isolation of Genomic LRF clones (Genome systems incorporated)

Mouse LRF amplicon

To screen for the mouse LRF genomic clone it was necessary to derive a PCR amplicon with mLRF specific primers which, upon PCR amplification of mouse genomic DNA, produced an amplicon of 294bp. This amplicon was then used by Genome Systems to screen their mouse ES P1 phage library which is comprised of plasmids containing inserts of up to 100Kb in size (Sternberg 1992). Given the low degree of sequence conservation within the P-Z region of mLRF (see section 3.2.2), an amplicon was generated from this region to ensure a high specificity screen. A 5' primer was derived from a region just up stream of the unique NheI site within the P-Z region of mouse LRF and the 3' primer was derived from a region just downstream of the

unique *StuI* site (see section 2.3.5, for schematic representation of mLRF). PCR amplification from both the mLRF cDNA (10ng) and mouse genomic DNA (100ng) templates produced a PCR product of 294bp which confirmed the absence of intronic sequence within this region. The PCR reactions were set up as described in section 2.1.4 and the conditions included a 95°C denaturation for 5min then 30 cycles of 95°C for 30 sec, 61°C for 1min 30 sec and 72°C for 3min followed by a final 10 min extension at 72°C.

Human LRF amplicon

Repeated searches of the databases led to the identification of a recently entered human expressed sequence clone (accession number N77854) which proved to be the human homologue of mLRF. From the available 316bp of this expressed sequence clone, which spanned part of the POZ domain, it was possible to design human (h)LRF specific primers which could amplify a 179bp PCR product from human genomic DNA (PCR conditions as for the mLRF amplicon). These primers were used by Genome Systems to successfully screen for the hLRF genomic P1 clone.

Mouse and human P1 plasmid DNA preparation

A modified protocol was used to prepare the DNA from the LRF positive P1 (pAD10SacBII) glycerol stocks provided by genome systems. The P1 glycerol stock was streaked on to kanamycin (25µg/ml) agar plates and grown o/n at 37°C. An individual bacterial colony was seeded in 16ml of terrific broth and incubated with shaking at 37°C o/n. The o/n culture was added to 500ml of terrific broth (12mg/ml bacto tryptone, 24mg/ml yeast extract, 0.4% glycerol, 2.31.mg/ml KH₂PO₄, 12.54mg/ml K₂HPO₄) and incubated with shaking at 37°C for 90min. IPTG was added to the culture to a final concentration of 0.5mM and the culture was then incubated o/n at 37°C with shaking. The resultant bacterial culture was pelleted at 6,000rpm and the DNA isolated using the Qiagen plasmid purification kit, as described in section 2.1.3.

Sequencing of P1 plasmid DNA

Owing to the large size of the P1 plasmid inserts, a modified sequencing reaction was necessary for effective results. Prior to sequencing, the P1 clone was purified by spot dialysis on VSWP 02500 Millipore filters floated on a bath of H₂O for 30min. The sequencing reaction was set up as for the standard protocol, see section 2.1.3., using 0.5µg of P1 plasmid DNA as template. The PCR conditions included a 95°C denaturation for 3min then 20

cycles of 95°C for 30 sec, 50°C for 30 sec and 70°C for 1min followed by 10 cycles of 95°C for 30 sec and 70°C for 1min). Best results were achieved using a sequencing primer 22 nucleotides in length with a T_m ($2x[A + T] + 4x[G + C]$) of greater than 70°C.

2.2.3 mLRF Zooblot

A zoo blot was prepared as in section 2.1.1 but also contained *Sarcomyces pombie* genomic DNA (Clontech) digested with both HindIII and BamHI. The conditions of stringency were identical to the mPLZF screen in section 2.1.1 but the ³²P-labelled probe was derived from an EcoRI digest of the partial embryonic mLRF plasmid.

2.2.4 Northern blot analysis

Isolation of the murine adult and embryonic tissue RNA was as described below. The murine and human cell line RNA were obtained from previously made frozen stocks provided by Dr Arthur Zelent.

Total RNA Isolation (mouse) (Chirgwin, Przybyla et al. 1979)

1-4gm of tissue was frozen in liquid nitrogen and ground to a fine powder with a mortar and pestle. The frozen powder was suspended in 15-20ml of solution I (4.5M Guanidinium thiocyanate (GnSCN), 0.5% sodium N-lauroylsarcosine, 25mM sodium citrate, pH 7.0, 0.1M 2-mercaptoethanol, 25mM EDTA, pH 7.0 and 0.1% Antifoam A) in a 50ml polypropylene tube and homogenised with an ultra-turrax for 1-2min at maximum speed. The homogenate was then transferred to a 30ml COREX tube and centrifuged for 10min at room temp and 8000rpm in a sorvall HB-4 rotor. The supernatant was saved and the pellet resuspended in 5-10ml of solution I, homogenised again and any remaining insoluble material removed by centrifugation. The supernatants were combined and caesium chloride (CsCl, optical grade, Gibco-BRL) added to a final concentration of 0.2g/ml and layered over a 5.7M CsCl cushion (5.7M CsCl, 0.1M EDTA, pH 7.0) occupying one-sixth of a total Beckman centrifuge tube volume. The supernatant was then centrifuged in a Beckman SW28 rotor at 25000rpm and 20°C for 20hr with slow deceleration. The supernatant was carefully removed and the RNA pellet

resuspended in 2-5ml of solution II (7.5M Guanidine hydrochloride [GnHCl], 25mM sodium citrate, pH 7.0 and 5mM dithiotriethanol [DTT]). The solution was heated to 60°C for 30 sec and subsequently vortexed to ensure complete resuspension of the pellet. Any remaining insoluble material was removed by centrifugation at rm temp and 8000rpm in a Sorvall HB-4 rotor. The RNA was then precipitated by adding to the supernatant 0.025 volumes of 1M acetic acid and 0.5 volumes of absolute EtOH, vortexing and incubating o/n at -20°C. The RNA was sedimented by centrifugation in a sorvall HB-4 rotor at 6000rpm and -10°C. The RNA was then washed with 80% EtOH to remove any remaining GnHCl, centrifuged at 8000rpm and dried under a vacuum. The RNA pellet was resuspended in DEPC treated H₂O and stored at -80°C.

Total RNA was derived from cell lines as described above except that approximately 10⁹ cells were lysed directly in GnSCN solution without freezing. Where appropriate, to induce differentiation, cells were treated for 24 hrs with 0.5µM RA (retinoic acid) prior to lysis.

Poly (A)⁺ RNA isolation (mouse) (Aviv and Leder 1972)

The re-precipitated total RNA pellet was dissolved in ETS solution (10mM Tris-HCl, pH 7.5, 1mM EDTA, 0.5% SDS) to a concentration of 0.65mg/ml, was heated to 85°C for 5min and cooled rapidly to rm temp. 5M NaCl was added to a final concentration of 0.5M and passed over an oligo dT chromatography column (see below). The poly A⁺ RNA bound column was washed with 15 volumes of ETS/0.5M NaCl and the pA⁺ RNA was eluted with 4 volumes of ET/0.5S (10mM Tris-HCl, pH 7.5, 1mM EDTA, 0.05% SDS) into a 15ml siliconised COREX centrifuge tube. The NaCl concentration of the eluate was adjusted to 1M and the pA⁺ RNA was precipitated by adding 3 volumes of absolute EtOH, incubating o/n at -20°C and sedimenting by centrifugation for 15min at 4°C and 8000rpm. The pellet was then washed in 80% EtOH, recentrifuged, dried under vacuum and resuspended in DEPC treated H₂O and stored at -80°C.

To prepare an oligo-dT chromatography column, 100mg of oligo-dT cellulose, type III (Collaborative Biomedica Products) suspended in 1ml of ETS solution was poured into a 12ml disposable poly-prep chromatography column (Biorad). Once the oligo-dT had settled it was washed with 10 volumes of ETS/0.1 NaOH and equilibrated with 10 volumes of ETS/0.5M NaCl.

Cell lines

Murine

P19 : neuronal cells that differentiate on treatment with retinoic acid (RA) (Jones-Villeneuve, McBurney et al. 1982); Wehi 3BD: differentiated myeloid cells (Warner, Moore et al. 1969); 32D: myeloid progenitors (Valtieri, Tweardy et al. 1987); 18.8 cells: B cell precursors (Yancopoulos, Desiderio et al. 1984); Mel: erythroid precursor cells (Friend, Patuleia et al. 1965); A4: multipotential cells (Heyworth, Dexter et al. 1990); B6 cells: multipotential cells (Greenberger, Sakakeeny et al. 1983); Lyd9: B cell precursors (Palacios, Karasuyama et al. 1987); J774: mature macrophage cells (Ralph and Nakoinz 1977); EL-4: mature transformed T cells (Ralph and Nakoinz 1973); ES: embryonic stem cells.

Human

Jurkat: acute T cell leukaemic cells (Weiss 1984); Nalm-6: pre-B child hood acute lymphoblastic leukaemic cells (Minowada, Janossy et al. 1978); Daudi: Burkitts lymphoma cells (Klein, Klein et al. 1967); K562: chronic myelogenous leukaemic cells (Lozzio and Lozzio 1975); KG1: acute myelogenous leukaemic cells (Koeffler and Golde 1978); HL-60: promyelocytic leukaemic cells lacking an RAR α gene translocation (Collins 1987); APL: t(15;17) promyelocytic leukaemic bone marrow cells (Chen, Brand et al. 1993); 293: adenovirus transformed 293 cells (Graham, Smiley et al. 1977); HepG2: hepatocellular carcinoma cells (Knowles, Howe et al. 1980); HSG: salivary gland adenocarcinoma cells (Shirasuna, Sato et al. 1981); T₂Cl₁₃: teratocarcinoma cells (Weima, van Rooijen et al. 1988); T47D: breast carcinoma cells (Keydar, Chen et al. 1979)

RNA gel electrophoresis

The RNA was reprecipitated (3-5 μ g of pA⁺ RNA or 10-15 μ g of total RNA) and the pellet resuspended in 6.25 μ l of DEPC treated H₂O. 2 μ l of 10x MOPS (0.4M MOPS, 0.1M NaAcetate, 10mM EDTA), 10 μ l of formamide (Molecular biology grade, NEB), 1.75 μ l of formaldehyde (Fluka) and 2 μ l of gel loading buffer (25% Ficoll 400, 1mM EDTA, 0.4% bromophenol blue and Xylene cyanol) were then added to the RNA. This RNA solution was denatured for 3min at 65^oC, chilled quickly on ice, loaded immediately on to a 1% RNA agarose gel and run by electrophoresis at 90mA for 3hr in circulating 1x MOPS buffer. To make the RNA agarose gel, 1% SEAKEM GTG agarose

(FMC) in DEPC treated H₂O was melted in a microwave oven and cooled to 60°C then 1xMOPS and 3.4% formaldehyde was added and the gel poured.

RNA Transfer

The gel was washed twice in 20x SSC for 10min each and transferred o/n by standard capillary action (as described in section 2.1.1) onto BAS85 nitrocellulose membrane (Schleicher and Schuell) pre-washed in H₂O and 20xSSC. The blot was then air dried and baked for 2hr at 80°C under vacuum.

Hybridisation

All the northern blots were hybridised under high stringency conditions, see section 2.2.1. Post hybridisation washes were as described in section 2.1.1. but the blots were more extensively washed in solution B for 15min at 65°C and subsequently washed in solution C (0.1x SSPE, 0.03% NaPPi, 1% SDS) for 15-30min at 65°C and exposed to film for a week at -80°C. For the mouse RNA blots, the ³²P-labelled probe was the same as that used in the mouse heart λZapII library screen, see section 2.2.1. For the human RNA blots a probe was derived by PCR from human genomic DNA. The PCR primers used for amplification corresponded to the N-terminal and C-terminal ends of hLRF POZ domain sequence (equivalent to nucleotides 43 to 395 of mLRF in Figure 9, section 3.2 and obtained by sequencing the human P1 clone, see section 2.2.2) and generated a 352bp fragment upon amplification. This fragment was ³²P-labelled as described in section 2.1.1. The PCR reactions were set up as described in section 2.1.4 and the conditions included a 95°C denaturation for 5min then 30 cycles of 95°C for 30 sec, 58°C for 1min 30 sec and 72°C for 3min followed by a final extension at 72°C for 10 min.

2.2.5 Chromosomal localisation of LRF (by Genome systems incorp.)

mLRF

DNA from the mouse LRF P1 clone (see section 2.2.2) was labelled with digoxigenin dUTP by nick translation. Labelled probe was combined with sheared mouse DNA and hybridised to metaphase chromosomes derived from mouse embryonic fibroblasts in a solution containing 50% formamide, 10% dextran sulphate and 2x SSC. Specific hybridisation signals were detected by incubating the hybridised slides in fluoresceinated

antidigoxygenin antibodies followed by counter staining with DAPI. The initial experiment resulted in specific labelling of a medium sized chromosome which was believed to be chromosome 10 on the basis of DAPI staining. A second experiment was conducted in which a probe which is specific for the centromeric region of chromosome 10 was cohybridised with LRF. This experiment demonstrated that LRF is located on chromosome 10. Measurements of 10 specifically hybridised chromosomes 10 demonstrated that LRF is located at a position which is 52% of the distance from the heterochromatic-euchromatic boundary to the telomere of chromosome 10, an area that corresponds approximately to band 10B5.3. A total of 80 metaphase cells were analysed and 68 exhibited specific labelling.

hLRF

DNA from the hLRF P1 clone was labelled as described above, combined with sheared human DNA and hybridised, under the same conditions as above, to normal metaphase chromosomes derived from PHA stimulated peripheral blood lymphocytes. Specific signals were detected as described above. The initial experiment resulted in specific labelling of the distal short arm of a group F chromosome. A second experiment was conducted in which a genomic clone previously mapped to 19q13.4 by Genome Systems (confirmed by co hybridisation with a genomic probe from the E2A locus on 19p13.3) was co hybridised with the hLRF probe. This experiment resulted in the specific labelling of the long and short arm of chromosome 19. Observations of specifically labelled chromosomes 19 demonstrated that hLRF is located at the terminus of chromosome arm 19p, an area which corresponds to band 19p13.3. A total of 80 metaphase cells were analysed with 69 exhibiting specific labelling.

2.3 The characterisation of a heterodimeric partner of LRF

2.3.1 Yeast two hybrid assay (Fields and Song 1989)

This assay involves the high efficiency co-transformation of two expression vectors (pGBT9/pAS I-CYH2 and pGAD424/pACTII) into the yeast strain Y190 (Harper, Adami et al. 1993). Both the pGBT9 and pAS1 hybrid vectors can be used to generate a fusion of the GAL4 DNA-binding domain with a whole or partial protein of interest (DBD). Likewise both the pGAD and pACTII hybrid vectors can be used to generate a fusion of the GAL4 activation domain with a whole or partial protein (AD). Both the DBD and

AD hybrid proteins are targeted to the yeast nucleus by nuclear localisation signals which are intrinsic to the GAL4 DNA-binding domain and have been incorporated into the GAL4 activation domain. To select for yeast colonies containing both the DBD and AD hybrid plasmids, transformed yeast are plated onto agar devoid of tryptophan and leucine which are essential amino acids for the growth of the Y190 strain. The DBD plasmid carries the wild type gene for tryptophan and the AD plasmid carries the wild type gene for leucine so that only co-transformed yeast cells will grow on the selective agar. If the two proteins within the respective GAL4 hybrids interact with each other, the normal function of the GAL4 transcriptional activator is restored because the GAL4 activation domain is tethered to its DNA binding domain. Transcription of a LacZ reporter construct containing upstream GAL4 binding sites can then be used as an indicator of protein interaction. The resultant β -galactosidase activity of the co-transformed yeast colonies can be assayed on X-gal plates; colonies containing interacting proteins will turn blue (Guarente 1983).

Growth of yeast

Y190 yeast were streaked out on a YPAD agar plate (20mg/ml Difco peptone, 10mg/ml yeast extract, 20mg/ml bactoagar and 2% dextrose/glucose solution. pH 5.8) and grown for 2 days at 32°C. 5ml of liquid YPAD (as for YPAD agar but without bactoagar) was inoculated with a freshly grown Y190 yeast colony and incubated with shaking at 200rpm o/n at 32°C. 50ml of warm YPAD was inoculated with the o/n culture of Y190 at a cell density of 5×10^6 /ml and grown with shaking at 32°C to a density of $1-2 \times 10^7$ cells/ml. This culture was sufficient for 10 transformations.

Competent cells

The Y190 cells grown as described above were harvested by centrifugation at 1500rpm for 5min. The Y190 medium was poured off and the cells resuspended in 25ml of sterile H₂O and centrifuged again. The H₂O was poured off and the cells resuspended in 1ml of 100mM LiAc. The cells were then pelleted in a microcentrifuge at 13000rpm for 5sec, the LiAc removed with a micropipette and the cells resuspended to a final volume of 0.5ml in 100mM LiAc to give a cell titre of $\sim 2 \times 10^9$ cells/ml. These competent cells were incubated at 30°C for 15min.

High efficiency transformation (Gietz and Woods 1994)

1ml of ssDNA, at 2mg/ml, was boiled for 10min and chilled on ice. The competent cells were vortexed briefly and 50 μ l were aliquoted into labelled eppendorf tubes. The cells were pelleted in a microcentrifuge at 13000rpm for 5sec and the LiAc removed with a micropipette. 240 μ l of 50% PEG 3500, 36 μ l of 1M LiAc, 25 μ l of ssDNA and 1 μ g of the appropriate plasmid DNA (both DBD and AD hybrid expressing vectors diluted to a volume of 50 μ l in sterile H₂O) were added, in the order listed, to each yeast pellet. Each eppendorf tube was vortexed vigorously until the cell pellet was completely mixed and then incubated at 30°C for 30min. The cells were heat shocked at 42°C for 20min and then pelleted in a microcentrifuge at 8000rpm for 15sec and the supernatant was removed with a micropipette. The yeast pellets were gently resuspended in 200 μ l of sterile H₂O and then plated out on selective agar plates (6.7mg/ml Difco yeast nitrogen base without amino acids, 20mg/ml dextrose/glucose, 20mg/ml Bacto agar, 0.64mg/ml yeast drop out amino acid medium without leucine and tryptophan) and grown for 2-4 days at 32°C to recover the transformants.

β -galactosidase assay (Breedon and Nasmyth 1985)

Individual colonies that grew on the selective agar plates were streaked onto replica selective agar plates and grown for 2 days at 32°C. Sterile 85mm cellulose filter papers (Whatman) were pre-soaked in 1.8ml of Z buffer/X-gal solution: 1.67ml of 10mg/ml X-gal stock solution, 0.27ml of 1.8M β -mercaptoethanol and 100ml of Z buffer (16.1mg/ml Na₂HPO₄·7H₂O 5.5mg/ml NaH₂PO₄·H₂O, 0.75mg/ml KCl and 0.246mg/ml MgSO₄·7H₂O, pH 7.0) in clean 100mm plates (Guarente 1983). Colonies were lifted by laying down another orientated 85mm cellulose filter paper on to the agar plate containing the transformed colonies. The cellulose filter paper was carefully peeled away from the agar plate and the adhered colonies permeabilised by immersing in liquid nitrogen for 10sec. Once thawed, this cellulose filter paper was placed colony side up onto the pre-soaked cellulose filter paper. These β -galactosidase assay plates were then wrapped in nescofilm and incubated at rm temp until the appearance of blue colonies. To stop the reaction the filters were dried in a fume hood.

2.3.2 Co-immunoprecipitation

An implied in vivo interaction between two proteins can be confirmed by demonstrating the co-precipitation of both transfected proteins from a cell lysate using an antibody specific to one of them. The lysis conditions are critical for this process to ensure that the proteins are isolated from the nucleus whilst maintaining the integrity and specificity of the interaction. With this in mind, a high salt, non ionic detergent lysis buffer was utilised.

Transient transfection of COS-1 cells by lipofection

Adherent COS-1 cells (Gluzman 1981) were propagated on 60mm plates in 5ml of D10 medium (500ml Dulbecco's modification eagle medium (Gibco), 2.4mM L-Glutamine, 2% Streptomycin sulphate, 1.2% benzyl penicillin sodium, 10% Fetal Calf Serum [FCS, Sigma]) in a 37°C incubator at 5% carbon dioxide. 1hr prior to transfection the COS-1 cells were fed with fresh D10 medium. 10ml of lipofectamine reagent (Gibco) was added to 2µg of plasmid DNA (1µg of BCL-6 and/or LRF expressing pSG5 plasmids were transfected and all quantities of DNA were made up to 2µg with empty pSG5) in a final volume of 64µl in PBSA and left at rm temp for 15min. Meanwhile the COS-1 medium was replaced with D10-FCS (D10 medium without FCS). 400µl of D10-FCS was mixed with the lipofectamine/DNA solution and added to the COS-1 cells. The cells were left at 37°C for 5hr and then the medium was replaced with D10 medium and the cells left to grow at 37°C.

Radioactive methionine labelling

After 16hr, the above transfected cells were washed three times in D10 medium minus methionine/cysteine and left in this medium for 30min at 37°C. The medium was then replaced with D10 medium plus 200µci ³⁵S methionine/cysteine (14.3mci/ml specific activity, Amersham) and left for 5hr at 37°C.

Lysis

Cells were washed three times in ice cold PBSA and 800µl of lysis buffer: 250mM NaCl, 10mM Tris pH 7.4, 1mM EDTA, 0.1mM Na₃VO₄, 0.5% Nonidet P-40 [Sigma], 2% phenylmethyl sulphonyl fluoride saturated in isopropanol and 3% proteinase inhibitor stock (100%: 2.5mg/ml each of arrotonin, soybean trypsin inhibitor, leupeptin and bestatin [Sigma]) was then added directly to the plates and left shaking on ice for 30min. The lysed cells were

pelleted in eppendorf tubes for 2min at 14000rpm and 4°C and the supernatants were collected and kept on ice.

Immunoprecipitation

To preadsorb the antibody on to the sepharose beads, for each sample, 2µg of Kodak™ anti-FLAG M2 mouse mono-clonal antibody was added to 50µl of 50% protein G sepharose slurry (Pharmacia) in lysis buffer and incubated with rotation for 2hr at 4°C. The antibody/sepharose slurry was then added to the above isolated cell lysates and incubated with rotation for 2hr at 4°C.

Post immunoprecipitation washes

The protein bound sepharose beads were sedimented briefly in a microcentrifuge and washed three times in lysis wash buffer (250mM NaCl, 10mM Tris pH 7.4, 1mM EDTA, 0.5% Nonidet-P40). After the final wash all remaining lysis wash buffer was removed with a drawn out pipette and the beads were resuspended in 100µl of denaturing loading buffer (0.125M Tris-HCl pH 6.8, 4% SDS, 20% glycerol, 0.2M DTT, 0.02% bromophenol blue) and denatured at 100°C for 10min to detach the proteins from the beads.

Western blot procedure

After denaturation of the proteins, the beads were sedimented in a microcentrifuge and the supernatant, containing the co-precipitated proteins, was run through a stacking gel: 0.125M Tris-HCl pH 6.8, 5% polyacrylamide (29% acrylamide, 1% N,N'-methylene bis-acrylamide stock) and 1% SDS, (polymerised with 0.1% ammonium persulphate and 0.05% N,N,N',N'-tetra methylethylenediamine) followed by an 8% denaturing running gel: 0.375M Tris-HCl pH 8.8, 8% polyacrylamide and 1% SDS, polymerised as for the stacking gel. The gel running buffer comprised 0.2M glycine, 25mM Tris base and 0.1% SDS. The proteins were then electro-transferred on to Immobilon™-P (PVDF) protein membrane (Millipore) at 120mA for 2hr in transfer buffer (48mM Tris base, 40mM glycine, 20% Methanol, 0.0375% SDS).

For immunodetection of the transferred proteins, the membrane was washed succesively, once in methanol, three times in H₂O, once in TBST (0.5% Tween-20 in TBS: 8mg/ml NaCl, 0.38mg/ml KCl, 3mg/ml Tris base pH 7.4 and then for 1hr with shaking in TBST plus 0.5% non fat milk. The membrane was then incubated for 1hr, with shaking, in TBST plus 5% non fat milk containing a 1:1000 dilution of rabbit polyclonal anti-BCL-6 antibody (a gift from D. Le

Prince Dhordain, Albagali et al. 1995). Subsequently the membrane was washed three times in TBST plus 5% non fat milk. In order to detect the primary antibody, the membrane was washed for 1hr, with shaking, in TBST plus 5% non fat milk containing a 1:1000 dilution of donkey anti rabbit immunoglobulin horse radish peroxidase conjugate (Amersham). Finally the membrane was washed, with shaking, three times for 10min each in TBST. The signal was detected using an ECL (enhanced chemiluminescence) kit (Amersham) then exposed to BIOMAX™ Kodak scientific imaging film for 30 sec and developed immediately (automated).

2.3.3 Co-immunofluorescence

Immunofluorescence involves the visualisation of the cellular localisation of a protein by incubating with a primary antibody specific to the protein of interest followed by a secondary antibody specific to the species of the primary antibody which is conjugated to a fluorophore dye. Upon excitation by the absorption of light at a particular wavelength, the dye will emit a characteristic fluorescence (Titus, Haugland et al. 1982). To examine the localisation of two proteins of interest within a cell, the two primary antibodies must have been derived from different species, such as rabbit and mouse, and each of the two secondary antibodies must fluoresce at a different wave length and be specific to a different primary antibody species. It is then possible to determine, by confocal microscopy, if the two proteins have overlapping localisations by examining their fluorescence at the two specific wave lengths. In isolation, the co-localisation of two proteins cannot be interpreted as an indication of their physical interaction but it is a good means of corroborating a previously demonstrated interaction.

Transfections

Chinese hamster ovarian (CHO) cells (Puck, Cieciura et al. 1958) were grown on glass cover slips in 60mm tissue culture plates. 2µg of plasmid DNA (1µg of BCL-6 and/or LRF or [-POZ]LRF expressing pSG5 plasmids; all quantities of DNA were made up to 2µg with empty pSG5) was transfected as described in section 2.3.2 and 24hr post transfection the cover slips with adhered transfected CHO cells were transferred to multi well dishes.

Immunostaining

Transfected cells were washed twice in PBSA (unless otherwise stated, all washes were performed at rm temp) and then fixed in 4% formaldehyde (Fluka) for 15min. The cells were then washed four times in PBSA, leaving the last wash on for 10min. Cells were permeabilised with 0.1% Triton X-100 (Sigma) in PBSA for 10min and were then washed four times in PBSA again leaving the last wash for 10min. 200µl of a primary antibody mix of anti-FLAG M2 antibody (1:1000 dilution) and anti BCL-6 polyclonal antibody (1:200 dilution) in PBSA was added on top of the single and double transfected CHO cells. The cells were placed in humidifying conditions and left for 1hr at rm temp. Excess primary antibody was removed by washing four times in PBSA, leaving the last wash for 10min. Cells were incubated with the secondary antibody mix of a 1:40 dilution of both fluorescein isothiocyanate (FITC) labelled goat anti mouse immunoglobulin and tetramethyl rodamine (TRITC) labelled goat anti rabbit immunoglobulin (Southern biotechnology associated inc.) in PBSA as for the primary antibody incubation. Excess secondary antibody was removed by washing four times in PBSA, leaving the last wash for 10min. Each cover slip was placed on a glass slide and as an anti-quenching agent, 10µl of moviol (Calbiochem) containing p-Phenylenediamine (Sigma) was added to each cover slip which was then covered with another glass cover slip. The corresponding fluorescence from each antibody bound protein was visualised by confocal microscopy (Biorad MRC 1000).

2.3.4 Luciferase reporter assay

The firefly luciferase protein can catalyse the oxidation of beetle luciferin with the concomitant production of a photon. In the presence of co-enzyme A, the kinetics of this oxidative reaction are sufficiently stabilised to allow the intensity of this flash of light to be measured by scintillation counting (Wood 1991). In this reporter system, a luciferase expressing reporter plasmid (PT109) under the control of a TK promoter, containing an upstream response element for BCL-6 is co-transfected with BCL-6 and/or LRF expressing plasmids. The intensity of photon emission from the lysed cells is a measure of the amount of transcribed and subsequently translated luciferase which correlates with the transcriptional activity of the co transfected gene products.

Transfections

293T cells (DuBridge, Tang et al. 1987) were propagated on 60mm plates, 4 μ g of DNA (See Figure 19 legend for plasmid concentrations, DNA content was made up with empty pSG5) was transfected, as described in section 2.3.2 and 24hr post transfection, the cells were lysed.

Lysis

Cells were washed twice with PBSA and 800 μ l of Promega 1x reporter lysis buffer (40mM TrisHCl, pH 7.5, 1mM EDTA, 15mM NaCl) added directly to the plates and left for 15min at rm temp with gentle agitation. Cell lysates were then collected into eppendorfs, mixed thoroughly and pelleted at 13000rpm for 2min at 4 $^{\circ}$ C. The supernatant was collected and placed at 4 $^{\circ}$ C for β -galactosidase and luciferase assays or stored at -70 $^{\circ}$ C.

β -galactosidase assay

In conjunction with the reporter and protein expression vectors, 0.25 μ g of a β -galactosidase expressing plasmid (PMC1871, Pharmacia) was transfected into the 293T cells. β -galactosidase will hydrolyse its colourless substrate, o-nitrophenyl- β -D-galactopyranoside (ONPG), to o-nitrophenol, which is yellow. The intensity of colouration is a measure of the quantity of β -galactosidase in the lysate which provides a measure of transfection efficiency.

100 μ l of the respective cell lysates were diluted in 50 μ l of 1x reporter lysis buffer and 150 μ l of Promega 2x assay buffer (200mM Na₂PO₃, pH 7.3, 2mM MgCl₂, 100mM β -mercaptoethanol and 1.33mg/ml ONPG) was added. This lysate solution was mixed thoroughly and incubated at 37 $^{\circ}$ C until the appearance of a yellow colour. The reaction was then stopped by adding 500 μ l of 1M sodium carbonate. The absorbance at 420nm for each reaction was measured using untransfected cell lysate as a blank.

Luciferase assay

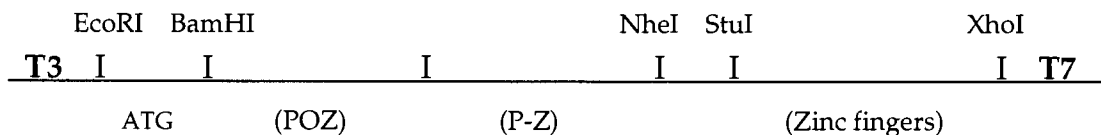
10 μ l of each cell lysate was added to 100 μ l of Promega luciferase assay reagent [20mM Tricine, 1.07mM (MgCO₃)₄Mg(OH)₂·5H₂O, 2.67mM MgSO₄, 0.1mM EDTA, 33.3mM DTT, 270 μ M coenzyme A, 470 μ M luciferin and 530 μ M ATP, pH 7.8] at the same time intervals for each sample. The scintillation count was then measured and the luciferase activities normalised for

transfection efficiency by dividing the scintillation values by the above determined β -galactosidase values.

2.3.5 Construct design

All PCR based cloning strategies involved the use of Native PFU DNA polymerase (Stratagene) which reduces the error rate 12 fold compared to Taq DNA polymerase. A 100 μ l reaction mix comprised 1U PFU polymerase, 0.2mM dNTPs, 50mM KCl, 10mM Tris-HCl pH 8.3, 1.5mM MgCl₂ and 0.001% gelatin. All digestions, gene cleaning, precipitations, ligations and sequencing were performed as described in section 2.1.3. The resultant plasmids were sequenced to check for mismatches within the PCR amplified section of the construct and, where necessary, to check that the sequence was in frame.

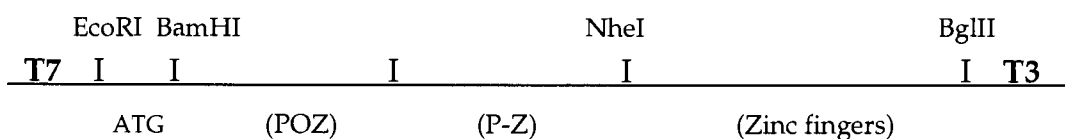
Schematic representation of full length LRF Bsk (see section 3.2)



LRF pSG5

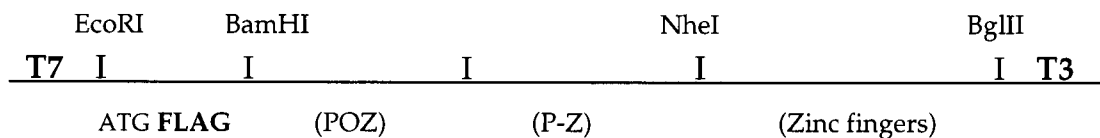
The LRF coding sequence was incorporated into an EcoRI/BglII cut mammalian expression vector pSG5 (Green and Isseman 1988) by the ligation of both an EcoRI/NheI digested fragment of LRF Bsk and an NheI/BglII digested PCR fragment. The PCR fragment was generated using a 5' primer from upstream of the NheI site of LRF Bsk and a 3' degenerate primer derived from the T7 end of Bsk with a novel BglII site incorporated into it. The PCR conditions, using 0.5 μ g of LRF Bsk as template, included a 5min denaturation at 94^oC then 10 cycles of 94^oC for 30sec, 58^oC for 1min 30sec and 72^oC for 3min followed by a final 10min extension at 72^oC.

Schematic representation of LRF pSG5

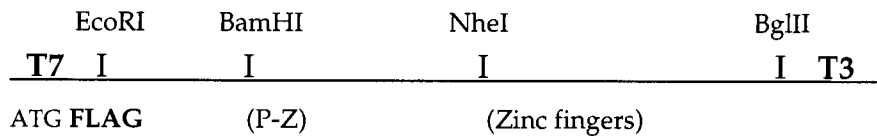


Flag-LRF pSG5

The FLAG sequence (N -Asp Tyr Lys Asp Asp Asp Lys- C) was incorporated into the N-terminus of LRF pSG5. PCR was performed using a long 5' primer containing an EcoRI site, followed by an ATG site, the FLAG nucleotide sequence and 18 nucleotides of sequence specific to the N-terminus of LRF with the indigenous ATG site mutated and a 3' primer derived from LRF Bsk sequence downstream of the NheI site. PCR conditions, using 0.5µg of LRF Bsk as template, included a 5min denaturation at 95⁰C then 10 cycles of 94⁰C for 25 sec, 59⁰C for 1min 30 sec and 72⁰C for 3min followed by a final extension of 10min at 72⁰C. The resultant PCR product was digested with EcoRI/NheI and an NheI/BglII fragment was digested from LRF pSG5. The two fragments were ligated into a pSG5 vector cut with EcoRI and BglII.

Schematic representation of Flag-LRF pSG5**POZ less Flag-LRF pSG5**

The POZ domain of Flag-LRF was deleted by a PCR strategy which was possible owing to the internal BamHI site downstream of the ATG of mLRF. A 5' degenerate primer was derived from a region just downstream of the POZ domain and it incorporated a novel BamHI site at its 5' end, the 3' primer was a, pSG5 specific, T3 sequence. PCR conditions, using 0.5µg of LRF pSG5, included a 5min denaturation at 95⁰C, then 10 cycles of 95⁰C for 25 sec, 54⁰C for 1min 30 sec and 72⁰C for 3min followed by a final extension for 10min at 72⁰C. The PCR product was digested with BamHI/BglII and ligated into a vector which was constructed by digesting Flag-LRF pSG5 (see above) with BamHI/BglII and gene cleaning it twice to ensure the complete removal of all of the original insert.

Schematic representation of POZ less Flag-LRF pSG5**LRF pACTII**

Full length LRF was cloned in frame into pACTII by the ligation of the resultant cDNA fragment of a BamHI/XhoI digestion of LRF Bsk into a BamHI/XhoI restricted pACTII vector.

LRF [POZ] pACTII

LRF without the zinc finger region was cloned in frame into pACTII by the ligation of the resultant cDNA fragment of a BamHI/XhoI digestion of embryonic partial mouse LRF Bsk plasmid (clone 1, see section 3.1.1) into a BamHI/XhoI restricted pACTII vector.

LRF[POZ] pASI

LRF without the zinc finger region was cloned in frame into pASI by the ligation of the resultant cDNA fragment of a BamHI digestion of embryonic mLRF Bsk plasmid (see above) into a BamHI restricted, phosphatased, pASI vector. The resultant construct was checked for the correct orientation by digesting with NcoI.

[-POZ]LRF pACTII

LRF without the POZ domain was cloned into pACTII by the ligation of the resultant cDNA fragment of a BamHI/BglII digestion of the POZ less Flag-LRF pSG5 construct (see above) into a BamHI restricted, phosphatased pACTII vector. This strategy was possible owing to the fact that BamHI and BglII sites are compatible. The digestion of POZ less Flag-LRF pSG5 with BamHI/BglII resulted in the removal of the FLAG region of the fragment owing to the internal BamHI site. The resultant plasmid was checked for the correct orientation by digesting with BamHI and XhoI.

[-POZ]LRF pGBT9

LRF without the POZ domain was cloned into pGBT9 by the ligation of the resultant cDNA fragment of a BamHI/BglII digestion of the POZ less FLAG-

LRF pSG5 construct (see above) into a BamHI restricted, phosphatased pGBT9 vector. The resultant plasmid was checked for the correct orientation by digesting with BamHI and Sall.

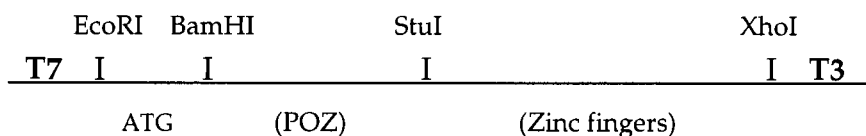
LRF pGEX

In order to facilitate the future production of a glutathione-s-transferase (GST) LRF fusion protein, to enable the generation of polyclonal antibodies, LRF was cloned in frame into the pGEX-5X-2 vector (Pharmacia) (Smith and Johnson 1988). LRF Bsk was digested with EcoRI/XhoI and the resultant cDNA fragment ligated into EcoRI/XhoI cut pGEX vector.

LRF [P-Z] pACTII

The P-Z region was deleted from LRF and the resultant mutant cloned into pACTII by a two step cloning strategy which utilised the existence of a StuI site just down stream of the P-Z region. An EcoRI/StuI digested PCR fragment was first ligated into a vector derived from an EcoRI/StuI digestion of LRF pGEX (see above) to produce LRF [P-Z] pGEX. The PCR fragment was generated using a 5' primer corresponding to the T3 Bsk sequence and a 3' degenerate primer derived from a region just downstream of the POZ domain of LRF Bsk which had a novel StuI site incorporated into it. PCR conditions, using 0.5µg of mLRF Bsk as a template, included a 5min denaturation at 95°C then 10 cycles of 95°C for 25 sec, 56°C for 1min 30 sec and 72°C for 3min followed by a final extension of 10 min at 72°C.

Schematic representation of LRF [P-Z] pGEX



The LRF P-Z mutant was then subcloned in frame into pACTII by the ligation of the resultant cDNA fragment of a BamHI/XhoI digestion of LRF [P-Z] pGEX into a BamHI/XhoI restricted PACTII vector.

BCL-6 response element luciferase reporter

Two oligonucleotides comprising the sense and anti-sense sequence of three repeats of the BCL-6 response element (5' GAA AAT TCC TAG AAA GCA TA 3') were designed so that the resultant annealed product would have a 5'

HindIII site overhang and a 3' SalI site overhang. 250pmoles of the sense and antisense oligonucleotides were annealed in 0.4M Tris-HCl pH 7.5 and 0.2M MgCl₂ at 65°C for 5min and then cooled at rm temp for 10min. 1µl of the annealed product was ligated into 25ng of the HindIII/SalI restricted TK luciferase pT109 reporter plasmid.

Chapter 3: Results

3.1 The identification of PLZF related POZ domain proteins.

In order to shed light on the transcriptional regulatory mechanisms of action of POK proteins in normal cell development and oncogenesis the aim of this study was to identify additional members of the POK gene family. It was hoped that this study would identify either potential heterodimerisation partners for PLZF or closely related POK proteins which might be important developmental regulators or potential targets for oncogenic transformation. Novel interacting partner proteins are often isolated by screening expression and yeast two-hybrid libraries. However, given that a high degree of homology between protein interaction domains can correlate with an *in vivo* interaction, in this study a cDNA homology screening strategy was used as this has the advantage of identifying closely related proteins in addition to interacting partner proteins. In parallel, an alternative approach, to isolate PLZF related POZ domain proteins, was also utilised which involved an RT-PCR strategy using degenerate oligonucleotide primers derived from the highly conserved sequences within the POZ domain.

3.1.1 The cloning of four POZ domain proteins

To investigate the potential existence of PLZF related POK proteins a low stringency screening of a zooblot using a ³²P-labelled PLZF probe was performed (see section 2.1.1). The probe used was derived from all of the coding sequence N-terminal to the zinc fingers of mouse (m)PLZF to avoid cross hybridisation with numerous unrelated zinc finger proteins. As seen in Figure 3, the presence of multiple bands in all species suggests that under these conditions of stringency the probe detected a number of related genes. It was therefore decided to use the same conditions to probe both a chicken heart cDNA library and a mouse 10 days post coitus (dpc) embryonic library (see section 2.1.2)

This cross species screen led to the isolation of a partial chicken (c)PLZF cDNA, see Figure 4 (Cook, Gould et al. 1995) and four additional POZ domain encoding genes; initially termed clones 1 to 4. The presumed full

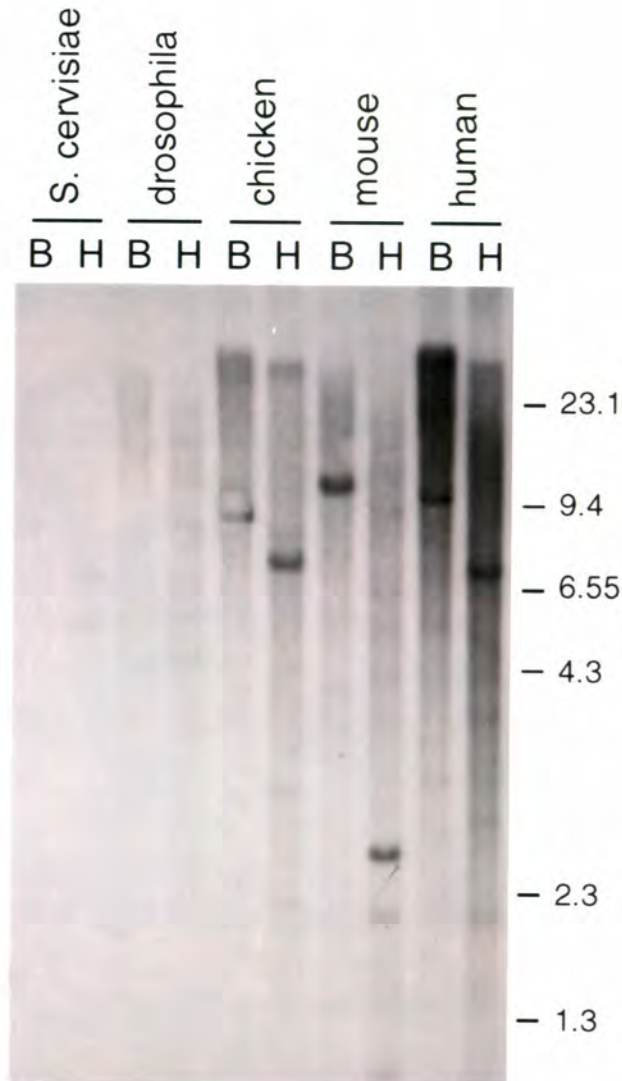


Fig. 3. Zooblot analysis. HindIII (H) and BamHI (B) digested genomic DNA derived from human, mouse, chicken, drosophila, *saccharomyces cerevisiae*, hybridised at low stringency (see page 30) with a ^{32}P -labelled probe comprising the POZ domain of murine PLZF.

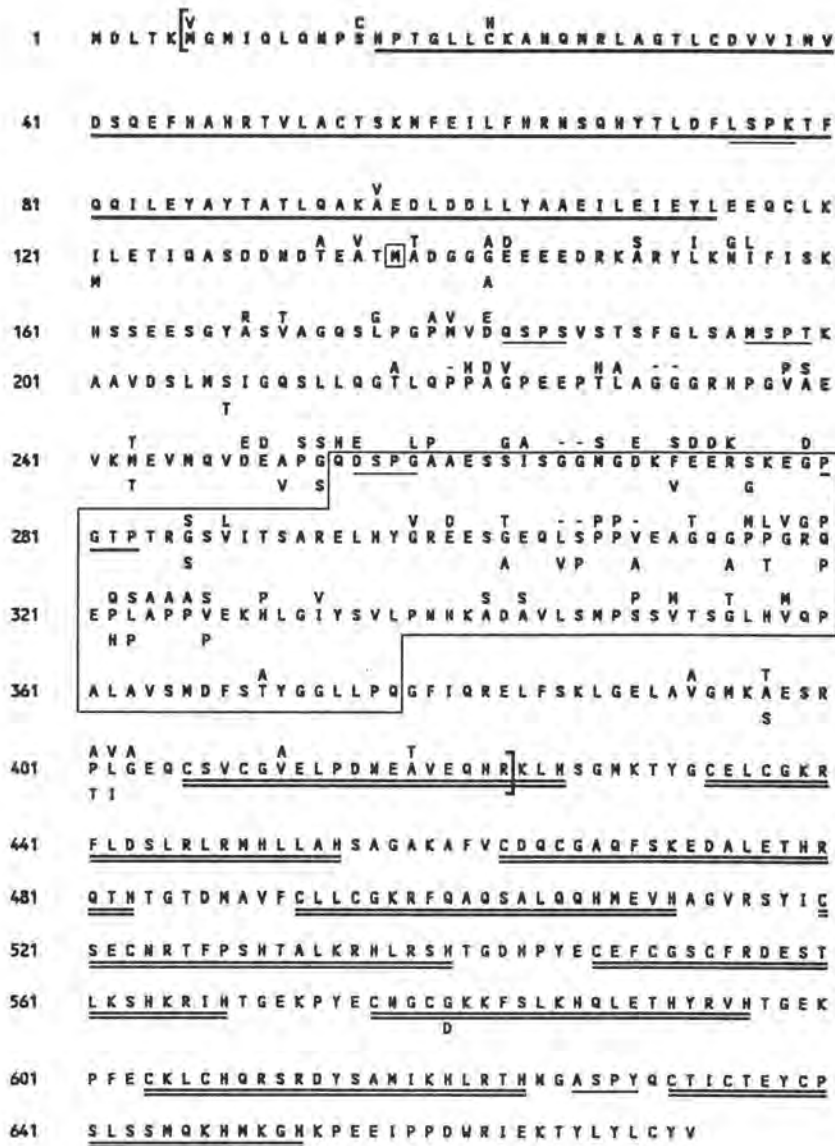


Fig. 4. Deduced amino acid sequence of the mPLZF(B) protein and its conservation. The murine sequence is numbered on the left. The residues that differ from the mouse PLZF(B) sequences are indicated below (human) and above (chicken) their murine counterparts. Dashes indicate an absence of the corresponding amino acids. Brackets indicate the extent of the amino acid sequence derived from the cPLZF(B) cDNA clone. Sequences that are encoded in the putative alternatively spliced exon are boxed. The POZ domain and Zinc-finger sequences are underlined with heavy single and double lines, respectively.

length open reading frame (ORF) of clone 1 was isolated from the chicken heart cDNA library and the partial cDNA of clone 1 was also isolated from the mouse embryonic cDNA library. The partial cDNAs of clones 2, 3 and 4 were isolated from the mouse embryonic cDNA library. Figure 5 shows the nucleotide and amino acid sequence of the full length clone 1 chicken sequence and the partial nucleotide and amino acid sequences of the mouse clones 2-4 are indicated in Figure 6. A putative ATG start site has been indicated for all four sequences but this is discussed in more detail with respect to clone 1 in section 3.2.

Although not present in the data base at the time of this screen, clone 2 and clone 3 have subsequently been published (Oyake, Itoh et al. 1996) and correspond to the Bach 2 and Bach 1 POZ proteins, respectively. Contrary to the initial belief, these two clones do not appear to represent partial POK proteins; they belong to the 4th group of POZ domain proteins (see section 1.1.4) which contain both an N-terminal POZ domain interaction motif as well as a basic leucine zipper motif. Just as clones 2 and 3 were pulled out together from a mouse 10 day library screen, the partial sequences of these same proteins (Bach 2 and Bach 1) were both pulled out of a yeast two-hybrid screen of a mouse 17.5 day expression library. A point of sequence divergence between both clones and their respective Bach proteins at the start of the second half of the POZ domain is indicated in Figure 6A and 6B (*). This site of divergence is the same for both clones and the nucleotide sequence conforms to an intron/exon boundary consensus (MacCumber and Ornstein 1984).

Clones 1 and 4 therefore represent the two novel POZ proteins identified as a result of this cDNA library screen. Clone 1 is a POK protein as it contains four C-terminal *krippel*-like zinc fingers. A search of the data base reveals that both these novel proteins are most closely related to the POK proteins Z13, ZF5 and PLZF (see section 1.1.4).

The alignment of the POZ domains from different proteins (Figure 1, section 1.1.4) indicates three highly conserved regions of amino acid sequence separated by non-conserved sequences. As described in section 2.1.4, in addition to the cDNA library screening approach, an RT-PCR strategy was performed on a pooled source of 8.5 and 13.5 dpc mouse embryonic RNA,

-49 CAATTCGGCGGAGGGACAGCGCGGCACGGCCCGCCCTTCGCAGCCACG
1 ATGGCGGGCGGCGTGGACGGCCCCATAGGGATCCCATTCCCGATCACAGCAGCGACATC
(1) M A G G V D G P I G I P F P D H S S D I
61 CTCAGCAGCCTGAACGAGCAGAGAAACAACGGGCTGCTGTGCGACGTGGTCATCTTGGTG
(21) L S S L N E Q R N N G L L C D V V I L V
121 GAAGGCCAGGAGTCCCCACCCACCGCTCCGTCTGGCGGCTGCAGCCAGTACTTCAAG
(41) E G Q E F P T H R S V L A A C S Q Y F K
181 AAGCTCTTCACCTCAGGGTTAGTGGTGGACCAGCAGAACGTGTATGAGATTGACTTTGTG
(61) K L F T S G L V V D Q Q N V Y E I D F V
241 AGCGCCGACGCCCTGTCGGCCCTGCTGGAGTTCGCCTACACCGCGACCCTCACAGTCAGC
(81) S A D A L S A L L E F A Y T A T L T V S
301 ACTTCCAACGTCAACGACATCCTCAACGCCGCAAACTGCTGGAGATCCCGCGGTGAGG
(101) T S N V N D I L N A A K L L E I P A V R
361 GATGTTTGACGGATCTCCTGGACAGGAAGATTCTGGCCAAAAATGACCAGATGGATTTA
(121) D V C T D L L D R K I L A K N D Q M D L
421 GTAGATCAAATTGATCAAAGGAACCATCTCAGAGCAAAAGAGTACCTGGAGTCTTCCAG
(141) V D Q I D Q R N H L R A K E Y L E F F Q
481 AGCAACCCCGTCAACGGCCCCAAGGCAGCTTTCCTGGACCAACCCGGAGTTGAAAAAC
(161) S N P V N G P Q G S F P W T N P *E L K N
541 CTTCAAAGGTCAACTTCCGCGGCCAGGAGGACGAGGAGGAGCCGGACTGCAACGGCGTGG
(181) L Q R S T S A A R R T R R S R T A T A W
601 ACTTCTACTTGCAAGCCCCCTCAACGAAAGACCAAGGCCAATGACTGTGATCCCGACA
(201) T S T C K P P S T K D Q R P M T V I P T
661 GCAACCCGGCCATGTGGCTGGACAGAGAGGAGGAGGAGGCGGGCGCCCGGCTCCTTG
(221) A T R P C G W T E R R R R R R R R P G S L
721 TTCTCGCTTCTCAGAACGGACATTACAGCGCCGAGGCTGGGCACGCCAGGAGAAGAG
(241) F S P S Q N G H Y S G R G W A R Q G E E
781 GAGGGGGCTCCCGGGGCGCCCTGGACCAGCAGGACCGCGGACTCCCCGGCTTCGTC
(261) E G A P G A A L D Q Q D A G D S P G F V
841 TCCACCGGGCGGTCCGGCGGATGACGACGCCAGGGACGTGGACGGGTTGGCCGCCACCGTG
(281) S T G R S A D D D A R D V D G L A A T V
901 CTGCACCATGTGATGGGCTCGGTGGGGAGGCATCACCTCCGGGACGACTCGCGGAACGAC
(301) L H H V M G S V G R H H L R D D S R N D
961 GACGACGGAGTCGTGGATTACTTGAATACTTTGGCACTTCGAACGAGAGCGACGTG
(321) D D G V V D Y Y L K Y F G T S N E S D V
1021 TACCCGTCCTGGTCCCAAAAGGTGGAGAAGAAGATCAGGGCGAAAGCATTCCAAAAGTGC
(341) Y P S W S Q K V E K K I R A K A F Q K C
1081 CCCATCTGCGCTAAGGTGATCCAGGGGCGGCAAGCTGCCGCGCCACATCCGCACCCAC
(361) P I C A K V I Q G A G K L P R H I R T H
1141 ACCGGGGAGAAGCCCTACGAGTGCAACATCTGCAACGTGCGCTTACCAGGCAGGACGAG
(381) T G E K P Y E C N I C N V R F T R O D E
1201 CTGAAGTCCACATGCGGAAACACACGGGGAGAAGCCCTACTTGTGCCAGCAGTGCGGC
(401) L K V H M R K H T G E K P Y L C Q Q C G
1261 GCCGCTTCGCTCAACTACGACTTGAAGAACCACATGCGCGTGACACCCGGGCTGCGG
(421) A A F A H N Y D L K N H M R V H T G L R
1321 CCCTACCAGTGCAGAGCTGCTTCAAGACTTTCGTCCGCTCTGACCACTTGACAGGCAC
(441) P Y Q C D S C F K T F V R S D H L H R H

Fig 5 continued on next page

```

1381 CTTAAAAGGATGGCTGCAACGGCATCCCGTCCCGGAGGGCCGACGGCCGCGGGTGAGA
(461) L K K D G C N G I P S R R G R R P R V R
1441 GATGCCCCGGGGGCTGCCACCCACGGGGAACCCAGGATGGAGGATTTTAAGGCGGTG
(481) D A R G A A H P T G N P R M E D F K A V
1501 GGGAGAGTCAGGAATCGGAGGACACCGTGCAGGGGAACGGCCGGGAGCAGCACTTTGAGG
(501) G R V R N R R T P C R G T A G S S T L R
1561 AGAGTTCTACTCCGAAGCGGCGATTGAATGTAGCAGGAGGGCGCCGAGGGCAGCGCGC
(521) R V L L R S G G L N V A G G A P R A A R
1621 CGGGCCCCGCCTAAAAGGAAAAACCCCAACCCACCCGACCGGCAGACAACACAAAGAG
(541) R A P P K R K N P Q P H P T G R Q H K E
1681 AGGTGTCACAGCATCTAGTTAGTACTTCTGTCTCCCATTTTCTCTTTCGAGGTGTTCTC
(561) R C H S I
1741 CCCCCCCAC

```

Fig. 5. Chicken clone 1 nucleotide and predicted amino acid sequences. The nucleotides are numbered on the left with the corresponding amino acids numbering in brackets. The POZ domain and zinc fingers are underlined with heavy double and single lines, respectively. The end of the corresponding murine clone 1 is indicated by an asterisk (*). A putative proline-dependant phosphorylation site is in italics and underlined, see Figure 9.

A

```

-15                                     TATGGTTTGAACGGC
1   ATGTCTGTGGATGAGAAGCCTGGCTCCCCATGTATGTATATGAGTCCACAGTCCACTGT
(1)  M S V D E K P G S P M Y V Y E S T V H C
61   GCCAACATCCTCCTGGGCTCAATGACCAACGGAAGAAGGACATTCTCTGTGACGTGACG
(21) A N I L L G L N D Q R K K D I L C D V T
121  CTGATCGTGGAGAGGAAGGAGTCCGAGCCCACCGGGCTGTGCTGGCCGCATGCAGTGAA
(41) L I V E R K E F R A H R A V L A A C S E
181  TACTTCTGGCAAGCACTGGTGGACAGACGAAAGATGACTTGGTGGTTCAGCTTGCCTGAG
(61) Y F W Q A L V G Q T K D D L V V S L P E
241  GAGGTCTTGTCTCTTCTCAGGAAAAGGCAAAGAACTGCCTGACAGAGCCGACGACTT
(81) E V *
301  TCATAAGAAGGGAAGATGAAGAGATGGGGTTTTGCCGAGGTGAGTCACTGGACCGGAGT
361  CAGCATCATTAGTGATTGAAATGGGCACTTATCAAGCCAATCCAGCCAATCGGAGTGCTC
421  AGACTCGTCTGTGAGCAGCAGCATCTGTGATACCAATGTCACGGCAGAAAATTTAGGTT
481  AGAATAACACTGTTCTTCTACACATGTTGTCAATTATGCTTGAAAAATGC

```

B

```

-180 GCGGTGGTGGCGGAGAGGAACGGAGCGCAGGGGAGGAGACGGGCGGCTGTCCGCGCAGAG
-120 GGAGTGAGTCACTGACCGCCGCTGCCGCTGCAGTCGGGCGCGCCGGTTTCGGCTCCG
-60  GTCCGATGACAGTGAGAAGCATGCTTCCACTGCTCTCCCTGGTCCCAGTTGCCACCCAGG
1   ATGTCTGTGAGTGAGAGTGCGGTATTTGCCTACGAGTCTCTGTGCATAGCACCAACGTC
(1)  M S V S E S A V F A Y E S S V H S T N V
61   TTGCTCAGCCTCAATGACCAGCGGAAGAAGGATGTCCTGTGTGATGTGACTGTCCTGGTG
(21) L L S L N D Q R K K D V L C D V T V L V
121  GGAGGCCAGCGGTTCCGAGCCCACCGCTCGGTGCTGGCTGCGTGCAGCAGCTACTTCCAC
(41) G G Q R F R A H R S V L A A C S S Y F H
181  TCGAGAATGGTAGGCCAGACTGACGCAGAGCTCACCGTCACTGCCTGAAGAGGTGAGG
(61) S R M V G Q T D A E L T V T L P E E V *
241  GCCACGGTGGTGGCCACTACATCCTGCTG

```

C

```

-45                                     GAGCTGGCTTCTGCTGCAGCCTTGGTGACCCGCCAGGTAAGGACA
1   ATGGATTCCCTGGTCACTTTGAGCAGATCTTCCACAGTTGAACTACCAGAGACTTCA
(1)  M D F P G H F E Q I F H Q L N Y Q R L H
61   GGGCAGCTGTGTGACTGTGTCATTGTAGTGGGGAACAGGCATTTCAAAGCCCACCGCTCC
(21) G Q L C D C V I V V G N R H F K A H R S
121  GTGCTGGCAGCATGCAGTACGCATTTCCGAGCACTGCTCTCGGTGGCGGAAGGAGATCAG
(41) V L A A C S T H F R A L L S V A E G D Q
181  ACCATGAACATGATCCAGCTAGATAGTGAGGTGGTGACAGCCGAGGCCTTCGCCGCCCTG
(61) T M N M I Q L D S E V V T A E A F A A L
241  ATTGACATGATGTATACCTCCACCCTCATGCTAGGGGAG
(81) I D M M Y T S T L M L G E

```

Fig. 6. Nucleotide and predicted amino acid sequences of mouse clone 2 (**A**), mouse clone 3 (**B**) and mouse clone 4 (**C**). The POZ domain sequences are double underlined and the start of a putative intron is indicated by an asterix for clones 2 and 3.

using POZ domain derived degenerate primers (indicated on Figure 1). This led to the isolation of a large number of clones containing inserts of 180-200bp. Approximately 1000 clones were hybridised with a PLZF specific oligonucleotide probe and this indicated that more than 80% of these clones contained a PLZF POZ domain. Despite the successful hybridisation of the remaining, PLZF negative, clones with a degenerate POZ specific primer, sequencing revealed only three POZ domain containing clones all of which turned out to be identical to clone 1.

3.1.2 Embryonic expression of clones 1 to 4 at 10 dpc

All four clones identified in this study, were isolated from a 10 dpc mouse embryonic cDNA library and at this stage of development PLZF has a highly spatially restricted expression pattern (see section 1.2.1). Therefore it was of interest to examine the expression pattern of all four clones at this stage by a modified whole mount in situ hybridisation procedure (see section 2.1.5) As seen in Figure 7, at day 10 of mouse embryonic development, clones 1-4 (A-D) have distinct and sometimes overlapping expression patterns. The higher intensity staining in the brain and autic vesicle is most likely due to antibody trapping as it is present in the sense stained controls (not shown). Like PLZF, all four clones show higher levels of expression in the developing limb buds (Figure 7 a-d). Whereas PLZF expression is absent from the apical epidermal ridge (see section 1.1.3) and is restricted to the more distal mesenchyme, all four clones are ubiquitously expressed in the limb bud. In addition, clone 1 appears to have a gradient of expression which increases across the proximo-distal axis of the developing limb bud.

The expression of clone 1 (A) in the branchial arches, frontonasal mesenchyme and the developing spinal chord together with an absence of expression in the heart, corresponds to the staining pattern of PLZF, albeit at lower levels. The possibility of this overlapping expression pattern being due to a cross-hybridisation of the clone 1 probe with PLZF RNA is ruled out by the additional expression of clone 1 in the extra-embryonic tissue and an absence of staining in the developing brain. The pattern of expression for clone 2 (B) is more restricted than for clone 3 (C) which is in agreement with the northern blot expression data for Bach 1 and 2 (Oyake, Itoh et al. 1996). Whilst clone 2 is expressed primarily in the developing limb buds and the

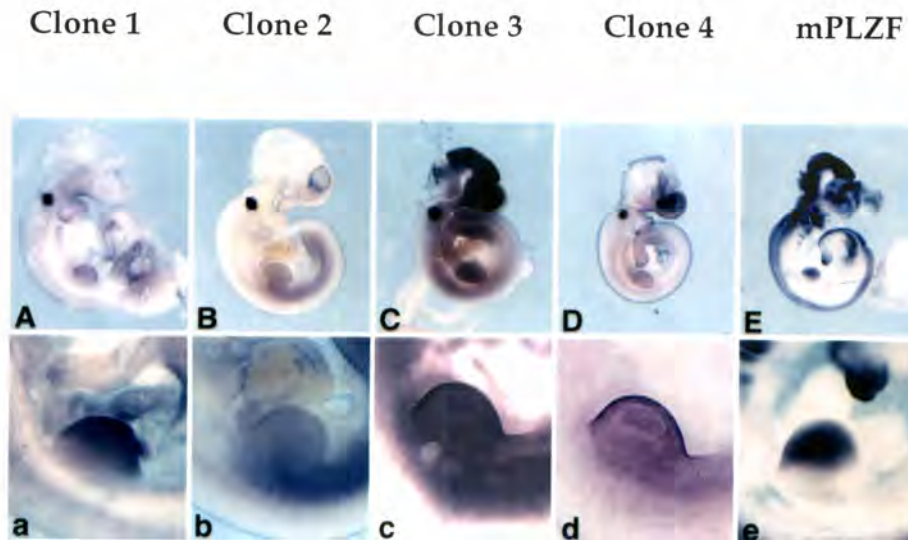


Fig. 7. In-situ hybridisation of 10 day mouse embryos. Individual embryos were hybridised under high stringency conditions (see page 40) with a digoxigenin labelled antisense RNA probe comprising all of the isolated mouse embryonic sequence of clone 1 (**A**), clone 2 (**B**), clone 3 (**C**) and clone 4 (**D**) or all of the coding sequence of mPLZF (**E**). The lower case labelled panels (**a-e**) are the corresponding high power views of the developing forelimbs. All the negative controls, comprising the respective digoxigenin labelled sense RNAs, exhibited high intensity staining corresponding to antibody or probe trapping in the autic vesicle and head folds (not shown).

intervening endoderm and ventral mesoderm, clone 3 is more widely expressed throughout the embryonic tissues with the notable exception of the spinal chord. The low level diffuse staining pattern in Figure 7D is difficult to interpret and may not reflect the complete expression pattern of clone 4 owing to the small size of the probe used in the in situ hybridisation procedure.

3.1.3. Clone 1 is the closest homologue of PLZF

Given the overlapping expression pattern of the four clones with PLZF in either the developing limb bud or the whole embryo at 10 dpc, it was important to identify the clone with the closest homology to PLZF as this might indicate a potential heterodimerisation partner for PLZF. As the partial clones did not contain a complete POZ domain, a homology comparison was made between the N-terminal halves of the respective POZ domains as this region has been implicated as being most critical in mediating protein interactions, at least among the drosophila POK proteins (see section 1.1.4). In addition, as indicated in Figure 1, this half of the POZ domain has a higher density of conserved amino acids. In Figure 8, the black shading indicates a high degree of similarity of all four clones to PLZF, whilst the grey shading indicates that the four clones are more homologous to each other than to PLZF. This is particularly apparent for clone 2 (Bach 2) and clone 3 (Bach 1) which are 74% similar. Clone 1 displayed the highest percentage of similar residues to PLZF within the N-terminal region of the POZ domain, and since the same clone was isolated by two different experimental approaches, as described in section 3.1.1, this suggests that clone 1 is the most closely related to PLZF. In addition, the overlapping pattern of embryonic expression of clone 1 and PLZF (see section 3.1.2, above) supports a potential interaction between the two proteins.

	PLZF	H	P	T	G	L	L	C	K	A	N	Q	H	R	L	A	G	T	L	C	D	V	V	I	M	V	D	S	D	E	F	H	A	R	R	T	V	L	A	C	T	S	K	M	F	E	I	L	F	-	V	D	D	-	% similarity	
clone 1		H	S	S	D	I	L	S	S	L	N	E	R	R	H	N	G	L	L	C	D	V	V	I	L	V	E	G	D	E	F	F	T	H	R	S	V	L	A	A	C	S	Q	Y	F	K	K	L	F	T	S	S	L	V	55	
clone 2		H	C	A	I	L	L	G	-	L	N	D	D	R	R	K	K	D	I	L	C	D	V	T	L	V	E	R	K	E	F	R	A	K	R	A	V	L	A	A	C	S	E	Y	F	W	Q	A	L	-	V	Q	Q	T	50	
clone 3		H	S	T	N	V	L	V	S	L	N	D	D	R	R	K	K	D	V	L	C	D	V	T	V	L	V	E	P	D	H	F	R	A	K	R	S	V	L	A	A	C	S	S	Y	F	H	S	R	M	-	V	Q	Q	D	50
clone 4		H	F	E	Q	I	F	H	Q	L	N	Y	D	R	L	H	G	Q	L	C	D	V	V	I	V	V	E	N	R	H	F	K	A	R	R	S	V	L	A	A	C	S	F	H	F	R	A	L	L	S	V	A	Q	G	52	

Fig. 8. Alignment of the N terminal half of the POZ domain of mPLZF with mouse clones 1-4. Amino acid sequences were aligned and the percentage (%) similarity to mPLZF determined using Microgenie software (Beckmann). The black background is used to indicate residues that are either identical to mPLZF or are conserved substitutions (see Figure 1 legend). The grey background indicates three or more identical residues that are different in the corresponding position in the mPLZF sequence.

3.2 A novel POK protein, LRF.

Having established that clone 1 was the most closely related to PLZF of the four POZ proteins identified, it was decided to characterise this novel protein further. To do so it was necessary to isolate the full length mouse cDNA of clone 1, henceforth named LRF for **L**eukaemia/**L**ymphoma **R**elated **F**actor, in order to facilitate more extensive functional analyses. It was hoped that further expression analysis would provide an insight into the possible function of LRF. In addition, it was important to identify the mouse and human chromosomal localisation of LRF to determine whether any reported mutations or translocation breakpoints mapped to this gene.

3.2.1 LRF structure

A high stringency screen of an adult mouse heart cDNA library using the partial mouse (m)LRF cDNA derived from the embryonic library as a ³²P-labelled probe (see section 2.2.1) was performed and led solely to the isolation of a cDNA encoding a 60 Kdalton mLRF protein; The nucleotide and predicted amino acid sequences of this clone are shown in Figure 9 and indicate a POK protein containing 4 C-terminal *krüppel*-like zinc fingers and an N-terminal POZ domain. As mentioned in section 3.1, an ATG start site has been tentatively assigned to the sequence in the absence of an upstream in frame stop codon. Nevertheless, this ATG lies within a very favourable kozak consensus sequence (Kozak 1986) and since the translational start sites of all the recorded POK proteins is of a comparable distance from the start of their respective POZ domains, see Figure 1, it is likely to be the true initiation codon of LRF.

The loss of nucleotide sequence homology between cLRF, mLRF (embryo) and mLRF (heart) upstream of this potential start site (Figure 10) suggested the existence of alternative splicing within this region. To further investigate the possibility of an intron/exon boundary it was necessary to analyse the genomic sequence of mLRF (see section 2.2.2 for P1 genomic DNA isolation) Sequencing of P1 DNA with mLRF primers specific for the POZ domain, using a modified sequencing protocol (see section 2.2.2), confirmed the existence of an intron upstream of the potential ATG start site which

```

-58      TTGCTGTCGCACGGTCGCCGCGCCCGCGCGCCCGCGGAAGGAGGTGTGCGGGAAG
1      ATGGCTGGCGGGTGGACGGCCCCATCGGGATCCCGTCCCGGACCACAGCAGCGACATC
(1)     M A G G V D G P I G I P F P D H S S D I
61      CTGAGCAGCCTGAACGAGCAGCGGACTCAGGGGCTGCTTTCGACGTGGTGATTCTTGTG
(21)    L S S L N E Q R T Q G L L C D V V I L V
121     GAAGGACGTGAGTTCCCCACGCACCGCTCGGTGCTGGCCGCTGCAGCCAGTACTTCAAG
(41)    E G R E F P T H R S V L A A C S Q Y F K
181     AAGCTGTTACGTCCGGAGCTGTAGTGGACCAGCAGAACGTGTACGAGATCGACTTCGTG
(61)    K L F T S G A V V D Q Q N V Y E I D F V
241     AGTGCCGACGACTGACGGCGCTCATGGACTTCGCTACACCGCCACGCTCACGGTCAGC
(81)    S A D A L T A L M D F A Y T A T L T V S
301     ACGGCCAATGTGGGCGACATCCTGAGTGCAGCAGCGCTGCTGGAGATCCCGCCGTGAGC
(101)   T A N V G D I L S A A R L L E I P A V S
361     CACGTGTGCGCCGACCTGCTGGAGCGTCAGATTCTGGCGGCTGATGATGTGGGCGACGCG
(121)   H V C A D L L E R Q I L A A D D V G D A
421     AGCCAGCCCGACGGGGCGGCCCCATTGACCAGCGCAACCTGCTGCGTGCCAAGGAGTAC
(141)   S Q P D G A G P I D Q R N L L R A K E Y
481     CTGGAGTTCTCCGAGTAACCCCATGAATAGCCTGCCCCCACTGCCTTCCCATGGTCT
(161)   L E F F R S N P M N [S L P P T A F P W S
541     GGCTTCGGTGCCCCGACGACGACCTGGACGCCACCAAGGAGGCTGTGGCCCGCGCTGTG
(181)   G F G A P D D D L D A T K E A V A A A V
601     GCCGCTGTGGCCGACGGCGACTGCAATGGCTTGGACTTCTATGGCCAGGGCCCCCGGCT
(201)   A A V A A G D C N G L D F Y G P G P P A
661     GATCGGCCCCCAGCCGGGATGGAGATGAGGGTGACAGTACCCAGGGCTGTGGCCTGAG
(221)   D R P P A G D G D E G D S T P G L W P E
721     AGAGATGAAGATGCCCCGCGGGAGGGCTATTTCCACCTCTACTGCCCCACCGGCCAC
(241)   R D E D A P P G G L F P T S Y C P T G H
781     CACACGAACGGCCACTATGGCCGGCAGGGGCTGGCACCGGTGAAGAAGAAGCGGGCGCT
(261)   H T E R P L W P A G A G T G E E E A A A
841     TTTTCTGAGGCCGCTCCAGAGCCGGGCGACTCCCCGGGCTTCTGTGAGGGCTGCAGAG
(281)   F S E A A P E P G D S P G F L S G A A E
901     GCGGAGGATGGGGACGCGCTGATGTGGATGGGCTAGCGGCCAGCAGCTGCTACAGCAG
(301)   G E D G ] D A A D V D G L A A S T L L Q Q
961     ATGATGTCATCGGTGGCCGGGCGGGGACAGTGATGAGGAGTCGGAACCGACGACAAG
(321)   M M S S V G R A G D S D E E S R T D D K
1021    GCGCTCATGGACTACTACCTGAAGTACTTCACTGGAGCCCACGAGGGGGATGTGTACCCA
(341)   G V M D Y Y L K Y F S G A H E G D V Y P
1081    GCCTGGTCACAGAAGGGTGAGAAGAAAATCCGGGCCAAGGCTTCCAGAAGTGTCCCATC
(361)   A W S Q K G E K K I R A K A F Q K C P I
1141    TCGGAGAAGGTGATTCAGGGTGCCGGCAAGCTGCCCGTCACATCCGCACGCACACGGG
(381)   C E K V I Q G A G K L P R H I R T H T G
1201    GAGAAGCCCTACGAGTGAACATCTGTAAAGTTCGATTACCAGACAGGACAAGCTGAAG
(401)   E K P Y E C N I C K V R F T R Q D K L K
1261    GTGCACATGCGGAAGCACACGGGTGAGAAGCCGTACTGTGCCAGCAGTGGCGGCCGCC
(421)   V H M R K H T G E K P Y L C Q Q C G A A
1321    TTCGCGCAACTACGACTGAAGAACCACATGCGGGTGACACGGGGCTGCGGCCATAC
(441)   F A H N Y D L K N H M R V H T G L R P Y
1381    CAGTGCATAGCTGCTGCAAGACCTTTGTGCGCTCCGACCATCTGCACAGACACCTTAAG
(461)   Q C D S C C K T F V R S D H L H R H L K

```

Fig. 9 continued on next page

```

1441  AAGGACGGCTGCAATGGGGTCCCCTCGCGCCGCGGCCGCAAGCCCCGTGTGCGGGGTGTG
(481)  K D G C N G V P S R R G R K P R V R [ G V
1501  CCACCCGATGTCCTGCCGGGGCCGGCGCACCCCCGGGCTCCCGGACGCCCGCGCAAT
(501)  P P D V P A G A G A P P G L P D A P R N
1561  GGCCAGGAGAAGCACTTTAAGGACGAGGAGGAGGCCAGCCCGGACGGCTCAGGCCGCCTG
(521)  G Q E K H F K D E E E A S P D G S G R L
1621  AATGTAGCGGGCAGCGGAGGAGACGATGGTGAGGTGGCCCGCGGTGGCCACCGCCGAG
(541)  N V A G S G G D D G A G G P A V A T A E
1681  GGTAACCTCGCAACCTGACTCGTATTAACAAACCAAAAAAAAAAAAAAAAAAACTCGA
(561)  G N F A T ]
1741  GGGGGCCCGGTACCCAATTGCGCCTAT

```

Fig. 9. Nucleotide and predicted amino acid sequence of murine LRF cDNA. The nucleotides are numbered on the left with the corresponding amino acids numbering in brackets. The POZ domain and zinc fingers are underlined with heavy double and single lines, respectively. The brackets indicate the sequence unconserved between mouse and chicken LRF. A single conserved proline-dependant phosphorylation site in the P-Z region (see section 3.2.2) is underlined and in bold type. An asterisk (*) indicates an intron/exon boundary in the 5' UTR, see Figure 10.

```

cLRF:  CGGCGGAGGGACAGCGCGGCACGGCCCCGCCCTTCGCAGCCACGATG
Heart mLRF:  TCGCCGGCGCCCGCGCGCCCGCCGCGAAGGAGGTGTCGCGGAAGATG
Embryo mLRF:  CGCGCGTCCTCCGGGCTCGGGTCCCCGAGAGGTGTCGCGGAAGATG
Genomic mLRF:  cccctacgctgactgctgcctctttctgcagGTGTCGCGGAAGATG

```

Fig. 10. Alignment of the 5' nucleotide sequence of chicken (c)LRF (clone 1), embryonic derived mouse (m)LRF, heart derived mLRF and genomic mLRF. The translation start site ATG is double underlined and the intronic sequence in genomic mLRF is shown in lower case. The loss of homology between heart and embryo mLRF, which corresponds to the start of the genomic intronic sequence, is indicated with an asterix.

conforms to the consensus intron/exon sequence (MacCumber and Ornstein 1984). The intron site is indicated in Figure 10 and corresponds to the point of sequence divergence for embryonic mLRF and adult heart mLRF (*). This developmental sequence divergence suggests the existence of splice variants of LRF either with alternative upstream coding sequences or with alternative 5' untranslated regions. Further investigation of this will involve the isolation of more 5' cDNA sequence of mLRF by RACE (rapid amplification of cDNA ends) (Loh, Elliot et al. 1989). Unlike the Bach proteins (clone 2 and 3), the sequence of genomic mLRF did not reveal the presence of an intron in the middle of the POZ domain coding sequence (see section 3.1.1, Figure 5A and 5B).

3.2.2 Evolutionary conservation

A low stringency screen of a zooblot using a ³²P-labelled probe derived from the embryonic mLRF cDNA is indicated in Figure 11 (see section 2.2.3). The presence of multiple bands in different lanes indicates that, under these conditions of stringency, the probe has detected specific LRF sequences in different species as well as a number of related genes. Indeed, a 6.5kb band present in the HindIII cut human genomic DNA lane in Figure 11 is also present in the corresponding lane in Figure 3 and may be due to the cross reactivity of the mLRF probe with PLZF sequences. It is unlikely that any of the related genes identified in this zooblot will prove to be drosophila and yeast LRF; results of an independent study in the laboratory on the evolutionary conservation of PLZF showed the existence of PLZF in vertebrates but not in drosophila or yeast.

As indicated in Figure 12A, a sequence alignment of the POZ domain amino acids of mouse, human and chicken LRF reveals a high degree of homology. As would be expected, the homology between mLRF and hLRF is higher at 96.6% as compared to 86.4% between mLRF and cLRF. The homology between the chicken and mouse LRF zinc finger regions is higher than within the POZ domains as indicated in Figure 12B. This degree of sequence conservation highlights the functional importance of both the POZ domain and zinc finger regions of LRF and the significance of this is investigated by deletion analysis in section 3.3.1. In contrast, the homology between cLRF

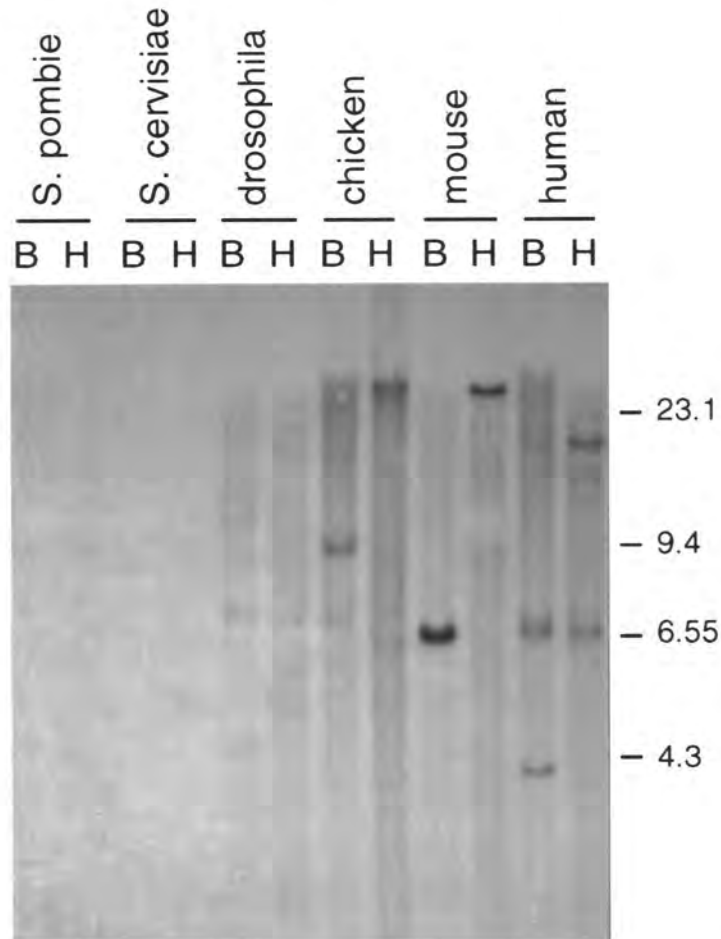


Fig. 11. Zooblot analysis. HindIII (H) and BamHI (B) digested genomic DNA derived from human, mouse, chicken, drosophila, saccharomyces cerevisiae and saccharomyces pombe probed with a ^{32}P -labelled mLRF probe.

A

```

mLRF: H S S D I L S S L N E Q R T Q G L L C D 35
      | | | | | | | | | | | | | | | |
hLRF: H S S D I L S G L N E Q R T Q G R L C D
      | | | | | | | | | | | | | | | |
cLRF: H S S D I L S S L N E Q R N N G L L C D

mLRF: V V I L V E G R E F P T H R S V L A A C 55
      | | | | | | | | | | | | | | | |
hLRF: V V I L V E G R E F P T H R S V L A A C
      | | | | | | | | | | | | | | | |
cLRF: V V I L V E G Q E F P T H R S V L A A C

mLRF: S Q Y F K K L F T S G A V V D Q Q N V Y 75
      | | | | | | | | | | | | | | | |
hLRF: S Q Y F K K L F T S G A V V D Q Q N V Y
      | | | | | | | | | | | | | | | |
cLRF: S Q Y F K K L F T S G L V V D Q Q N V Y

mLRF: E I D F V S A D A L T A L M D F A Y T A 95
      | | | | | | | | | | | | | | | |
hLRF: E I D F V S A E A L T A L M D F A Y T A
      | | | | | | | | | | | | | | | |
cLRF: E I D F V S A D A L S A L L E F A Y T A

mLRF: T L T V S T A N V G D I L S A A R L L E 115
      | | | | | | | | | | | | | | | |
hLRF: T L T V S T A N V G D I L S A A R L L E
      | | | | | | | | | | | | | | | |
cLRF: T L T V S T S N V N D I L N A A K L L E

mLRF: I P A V S H V C A D L L E R Q I L A 133
      | | | | | | | | | | | | | | | |
hLRF: I P A V S H V C A D L L D R Q I L A 96.6%
      | | | | | | | | | | | | | | | |
cLRF: I P A V R D V C T D L L D R K I L A 86.4%

```

Fig. 12A See next page for legend

B

mLRF:	C	P	I	C	E	K	V	I	Q	G	A	G	K	L	P	R	H	I	R	T	397
cLRF:	C	P	I	C	A	K	V	I	Q	G	A	G	K	L	P	R	H	I	R	T	379
mLRF:	H	T	G	E	K	P	Y	E	C	N	I	C	K	V	R	F	T	R	Q	D	417
cLRF:	H	T	G	E	K	P	Y	E	C	N	I	C	N	V	R	F	T	R	Q	D	399
mLRF:	K	L	K	V	H	M	R	K	H	T	G	E	K	P	Y	L	C	Q	Q	C	437
cLRF:	E	L	K	V	H	M	R	K	H	T	G	E	K	P	Y	L	C	Q	Q	C	419
mLRF:	G	A	A	F	A	H	N	Y	D	L	K	N	H	M	R	V	H	T	G	L	457
cLRF:	G	A	A	F	A	H	N	Y	D	L	K	N	H	M	R	V	H	T	G	L	439
mLRF:	R	P	Y	Q	C	D	S	C	C	K	T	F	V	R	S	D	H	L	H	R	477
cLRF:	R	P	Y	Q	C	D	S	C	F	K	T	F	V	R	S	D	H	L	H	R	459
mLRF:	H																				478
cLRF:	H	94.1%																			460

Fig. 12. Evolutionary sequence conservation. **A.** Alignment of the POZ domain amino acid sequence of mouse, human and chicken LRF where [|] indicates identical residues and [:] indicates conservative substitutions. **B.** Alignment of the zinc finger region of mouse and chicken LRF where [|] indicates identical residues. All percentage (%) values correspond to the degree of identity to mLRF and were determined using Microgenie software (Beckmann).

and mLRF is very low outside of the POZ domain and zinc finger regions (P-Z region). However, as is apparent from comparing Figures 5 and 9, the number of amino acids in this P-Z region is well conserved across species and in addition there is a conserved potential proline-dependant phosphorylation site. As indicated in the amino acid alignment of mPLZF and cPLZF in Figure 4, this loss of homology in the P-Z region may be a common feature for all POZ domain proteins. Therefore, it is anticipated that the sole function of this region is to serve as a spacer between the POZ domain and zinc fingers of the protein.

3.2.3. Northern blot analysis

A ^{32}P -labelled probe derived from the mLRF embryonic cDNA was used to hybridise, under high stringency conditions, northern blots containing mouse RNAs derived from adult tissues, haematopoietic cell lines and mouse embryos at different stages of gestation (Figure 13, see section 2.2.4 for procedure and cell line references). In the adult (Figure 13A), mLRF appears to exist predominantly as two isoforms, B and C, and is widely expressed. The highest levels of expression are in the adult brain (lane 3) whereas in haematopoiesis, mLRF is expressed very weakly in Lyd-9 pro-B cells (lane 15) but is detected at higher levels in the more mature 18.8 pre-B cells (lane 11). This suggests that the expression of LRF undergoes temporal variation during B cell development. During embryonic development (Figure 13B), the two predominant mLRF isoforms are A and C, as compared to B and C in adult tissues. Both isoforms A and C display a high level of expression at stage 7dpc which then declines temporarily during mid gestation. As eluded to in section 3.2.1, the variation in predominant isoforms for mLRF between embryonic and adult mouse tissues lends support to the proposed existence of alternative splicing resulting in the expression of different transcripts at different stages of development.

Northern blots representative of adult human tissues and cell lines (Figure 14) were also hybridised under the same conditions as above using a ^{32}P -labelled probe derived by PCR from the human P1 genomic clone (see section 2.2.4 for procedure and cell line references). This also revealed a wide spread expression pattern for hLRF and three predominant transcripts were

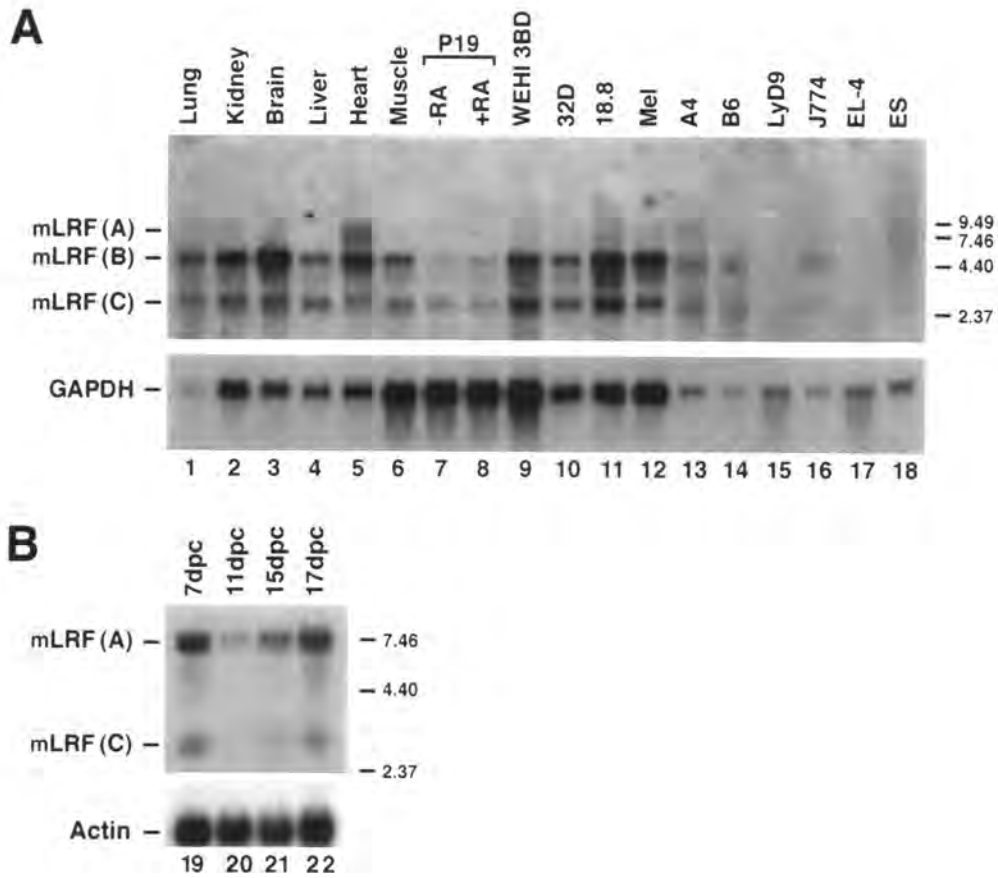


Fig. 13. Northern blots of murine RNA hybridised under conditions of high stringency (see page 49) with a ^{32}P -labelled probe comprising the POZ domain of murine LRF. Lanes 1-12: 3 μg of PolyA⁺ RNA, lanes 13-22: 10 μg of total RNA. Size markers are indicated on the right of each panel (A and B) in Kb. **A.** Expression of three mLRF transcripts (A, B and C) in selected adult tissues (lanes 1-6) and various cell lines (lanes 7-18, see section 2.2.4 for references). Hybridisation with a ^{32}P -labelled GAPDH cDNA probe (lower panel) was used as a control for mRNA integrity. **B.** Expression of two predominant mLRF transcripts (A and C) in mouse embryos at 7-17 days post coitus. The 2kb actin signal (lower panel) indicates RNA integrity.

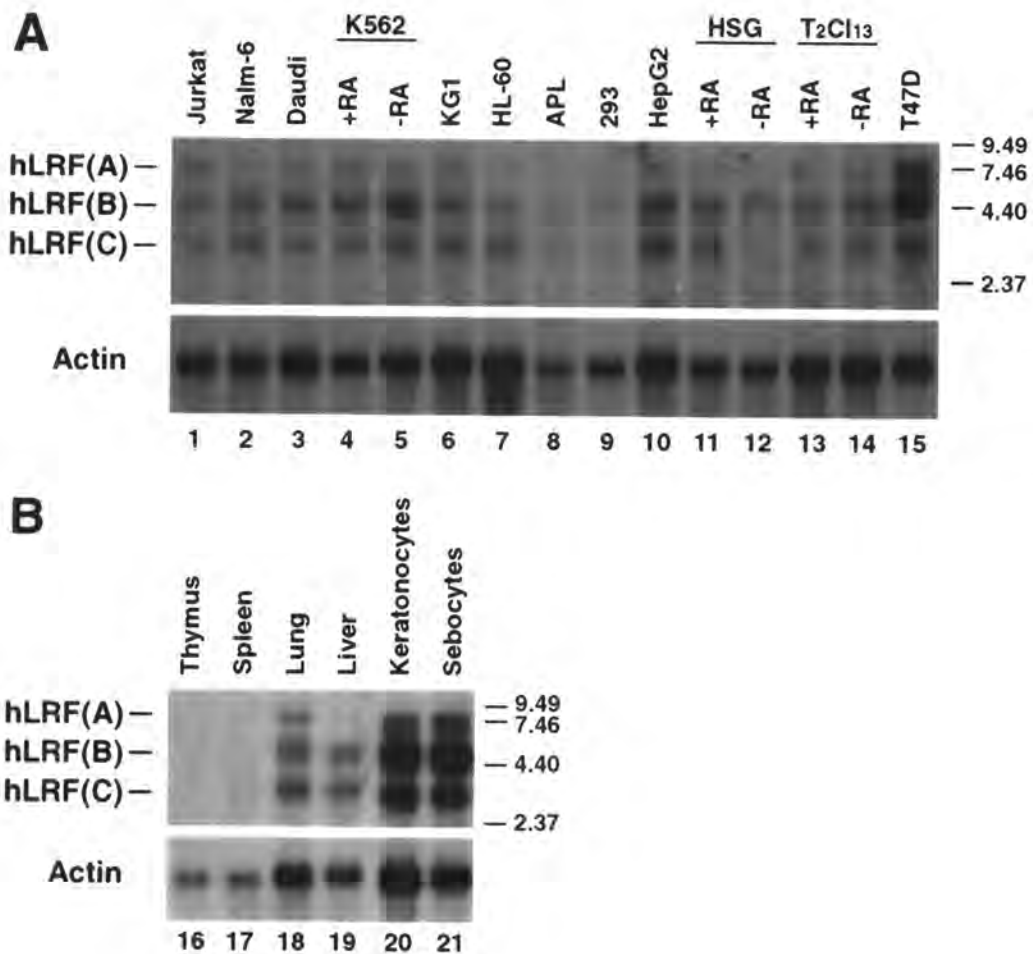


Fig. 14. Expression of human LRF in different cell lines (**A**, see section 2.2.4 for references) and tissues (**B**). Northern blots of human pA^+ RNA (3 μ g/lane) were hybridised under conditions of high stringency with a PCR derived ^{32}P -labelled probe comprising the POZ domain of human LRF (see page 49). Three predominant hLRF transcripts, corresponding in size to the murine transcripts A, B and C, are indicated. The 2kb actin signal (lower panel) indicates RNA integrity and the size markers (Kb) are positioned on the right of each blot.

detected corresponding in size to the mouse isoforms A, B and C. Figure 14A indicates that hLRF is present in all the tumour derived cells examined. Prompted by a potential role in B cell development based on the mLRF expression pattern, it was important to establish the expression pattern of LRF in lymphoid organs. Albeit at low levels, hLRF is expressed in the thymus and spleen (Figure 14B). With respect to the alternative isoforms of hLRF, transcript A is notable by its absence in the adult liver (lane 19). In addition, the expression of transcript C is upregulated upon the induced differentiation of HSG (salivary gland adenocarcinoma) cells (lanes 11 and 12) by retinoic acid treatment.

3.2.4. Chromosomal localisation

The mouse and human chromosomal localisation of LRF was examined by fluorescence in situ hybridisation (FISH) (see section 2.2.5). mLRF is localised to chromosome 10, band B 5.3 (Figure 15A) and hLRF to the short arm of chromosome 19; 19p13.3 (Figure 15B). Of the metaphase cells hybridised with digoxigenin labelled mLRF and hLRF probes, 85% and 86% exhibited specific signals, respectively. This chromosomal localisation of LRF is in accordance with the linkage homology of mouse chromosome 10 and the short arm of human chromosome 19 (see section 4.2 for details).

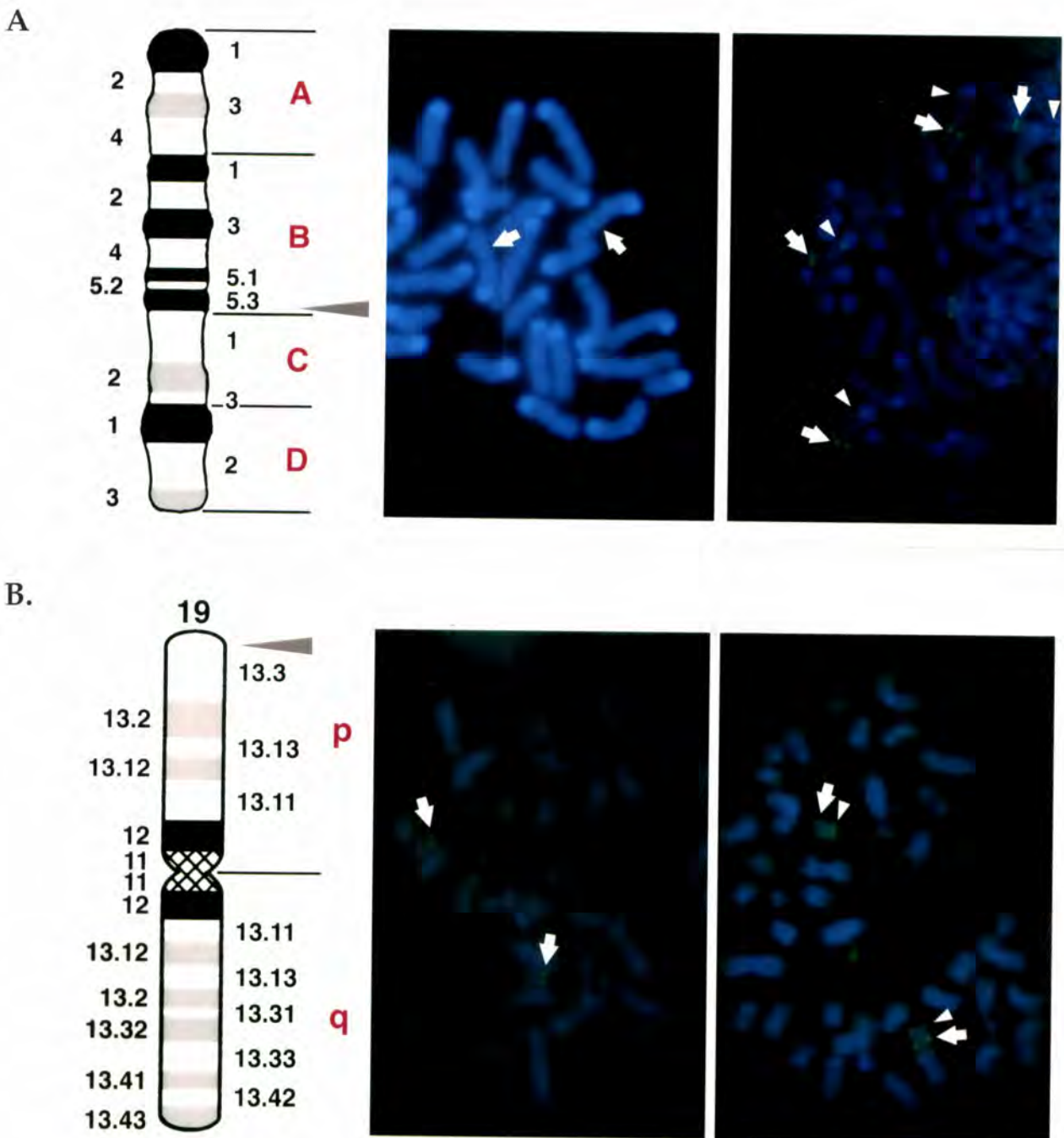


Fig. 15. Chromosomal localisation of LRF by Fluorescence In Situ Hybridisation. **A.** Specific labelling localised mLRF to mouse chromosome 10 band B5.3 (arrows, first panel) and this was confirmed by co-hybridising LRF with a probe specific for the centromeric region of chromosome 10 (triangles, second panel). This localisation of mLRF is indicated with a grey arrow on the schematic representation of mouse chromosome 10. **B.** Specific labelling localised hLRF to human chromosome 19p13.3 (arrows, first panel) and this was confirmed by co-hybridising LRF with a probe specific for the chromosomal location 19q13.4 (triangles, second panel). This localisation of hLRF is indicated with a grey arrow in the schematic representation of human chromosome 19.

3.3 The characterisation of a heterodimeric partner of LRF

Given the involvement of the POZ domain in protein-protein interactions, it was important to determine if LRF could interact with itself or any other POK proteins. It was hoped that information derived from these studies would contribute towards an understanding of POK protein function in the regulation of transcription.

3.3.1 Identification of a heterodimeric partner for LRF

An *in vivo* analysis was utilised in this study to examine LRF protein interactions. In this respect it is worth noting that all the heterodimeric interactions between POZ domain proteins, described so far, have been observed *in vitro*. A good candidate protein partner for LRF, given their degree of homology in the POZ domain and their overlapping expression pattern, was PLZF.

3.3.1.1 Yeast two-hybrid assay

Contrary to the initial premise, in the yeast-two hybrid system, LRF and PLZF did not interact. Co-transformation of the respective PLZF and LRF expressing GAL4 hybrid vectors into yeast cells did not induce the expression of the β -galactosidase gene which is indicative of a positive interaction (see section 2.3.1 for procedure and section 2.3.5 for construct design). Prior to these studies, an interaction between BCL-6 and PLZF had been demonstrated by the yeast two-hybrid system. It was decided, therefore, to determine if LRF, a protein closely related to PLZF, might also interact with BCL-6. Indeed, it was possible to demonstrate a positive interaction between full length BCL-6 and LRF (Figure 16A, lane 3). In order to characterise further this *in vivo* interaction, the effect of truncated BCL-6 and LRF proteins was analysed (Figure 16A, BCL-6 constructs were provided by D. LePrince (Dhordain, Albagli et al. 1995)). Although the POZ domain can mediate a homodimeric interaction (Figure 16, lanes 1 and 2), it is insufficient for a heterodimeric interaction between LRF[POZ] and BCL-6[POZ], *in vivo* (Figure 16, lane 5). The absence of the less well conserved region between the POZ domain and zinc fingers of LRF (LRF[P-Z]) did not prevent an interaction between full-length BCL-6 (Figure 16, lane 4). Although the LRF[P-Z] might be expected to have a significantly altered tertiary structure,

A



B

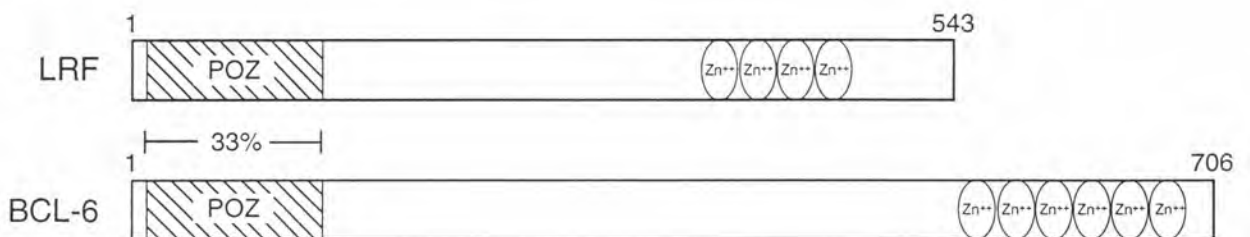


Fig. 16. A. Yeast two-hybrid assay showing homodimeric and heterodimeric interactions of wildtype and mutated LRF and BCL-6 proteins fused to the Gal4 activation domain (AD) and/or Gal4 DNA binding domain (DBD). Yeast colonies were transferred to filter paper and stained for β -galactosidase activity; a positive interaction resulting in the development of a blue colour. **B.** Schematic representation of BCL-6 and LRF proteins indicating the POZ domain (POZ) and zinc fingers (circled Zn⁺⁺ symbols). Percentage value indicates the extent of homology between the POZ domains of the two proteins.

	AD		DBD	Colour of colonies
1.	pACT II	v	pAS I	white
2.	pGAD	v	pGBT9	white
3.	pACT II	v	pGBT9	white
4.	LRF[POZ] pACT II	v	pAS I	white
5.	pACT II	v	LRF[POZ] pAS I	white
6.	BCL-6[POZ] pGAD	v	pGBT9	white
7.	pACT II	v	BCL-6[POZ] pGBT9	white
8.	LRF pACT II	v	pAS I	white
9.	pACT II	v	BCL-6 pGBT9	white
10.	LRF[P-Z] pACT II	v	pAS I	white
11.	[-POZ]LRF pACT II	v	pAS I	white
12.	[-POZ]BCL-6 pGAD	v	pAS I	white
13.	pACT II	v	[-POZ]LRF pGBT9	blue
14.	pACT II	v	[-POZ]BCL-6 pGBT9	blue

Table 2. Yeast two-hybrid controls (for construct design see section 2.3.5). Co-transformation of empty GAL4 fusion vectors together with the various BCL-6 or LRF expressing GAL4 vectors, in the combinations indicated, should not induce β -galactosidase expression due to the absence of an interaction. Both the [-POZ]LRF pGBT9 and [-POZ]BCL-6 pGBT9 GAL4 constructs when co-transformed with empty pACTII produced a blue colour (lanes 13 and 14, respectively) and could not be used in the assay system as they activate β -galactosidase expression in the absence of an interaction (self activating).

it was still able to heterodimerise with BCL-6 and the interaction was stronger than wild type LRF, as measured by the time taken to develop a blue colour (30 minutes as compared to 4 hours). In line with the results shown in lane 5 of Figure 16, LRF[P-Z] was unable to heterodimerise with just the POZ domain of BCL-6 (Figure 16, lane 6). The inability of a POZ-less LRF mutant ([-POZ]LRF) to interact with full length BCL-6 suggests the absence of a solely zinc finger mediated heterodimeric interaction (Figure 16, lane 7). This would have been best confirmed by co-transforming [-POZ]LRF with [-POZ]BCL-6 however as indicated in Table 2, rows 13 and 14, both these constructs are self activating when fused to the Gal4 DNA binding domain and produce a blue colour in the absence of an interaction, see section 2.3.1. In addition, an intriguing intramolecular interaction between the POZ domain and zinc fingers of LRF (Figure 16, lane 9) but not between the POZ domain and zinc fingers of BCL-6 (Figure 16 lane 10) was apparent. There was no heterodimeric, intermolecular, interaction between the zinc fingers of LRF and the POZ domain of BCL-6 (Figure 16, lane 8).

As indicated in the schematic representation of LRF and BCL-6 (Figure 16B) the degree of homology between the POZ domains of LRF and BCL-6 is 33 % whereas the degree of homology between the first or last four zinc fingers of the respective proteins is 42% . The results of the two hybrid analysis suggest that a heterodimeric interaction between POK proteins is mediated through both the POZ domain and zinc finger regions; the heterodimeric complex only being stable in the presence of both interacting regions. In contrast, the POZ domain has been shown in this study and by others to be sufficient for a POK protein homodimeric interaction, *in vivo*. It is important to stress that the results of this mutational analysis may not necessarily be directly applied to the mammalian system owing to the possible absence of cooperating proteins which are not conserved in yeast.

3.3.1.2 Co-immunoprecipitation

To confirm the *in vivo* interaction between full length LRF and BCL-6 in mammalian cells, a co-immunoprecipitation assay was carried out (see section 2.3.2 for procedure and section 2.3.5 for construct design). N-terminal FLAG tagged LRF (Flag-LRF) and BCL-6 proteins were transiently expressed in COS-1 cells under the control of the SV40 promoter (pSG5). COS-1 cells were transfected with 1µg of both BCL-6 and Flag-LRF (Figure 17, lanes 1 and 5), 1µg of Flag-LRF (Figure 17, lanes 2 and 6), 1µg of BCL-6 (Figure 17, lanes 3

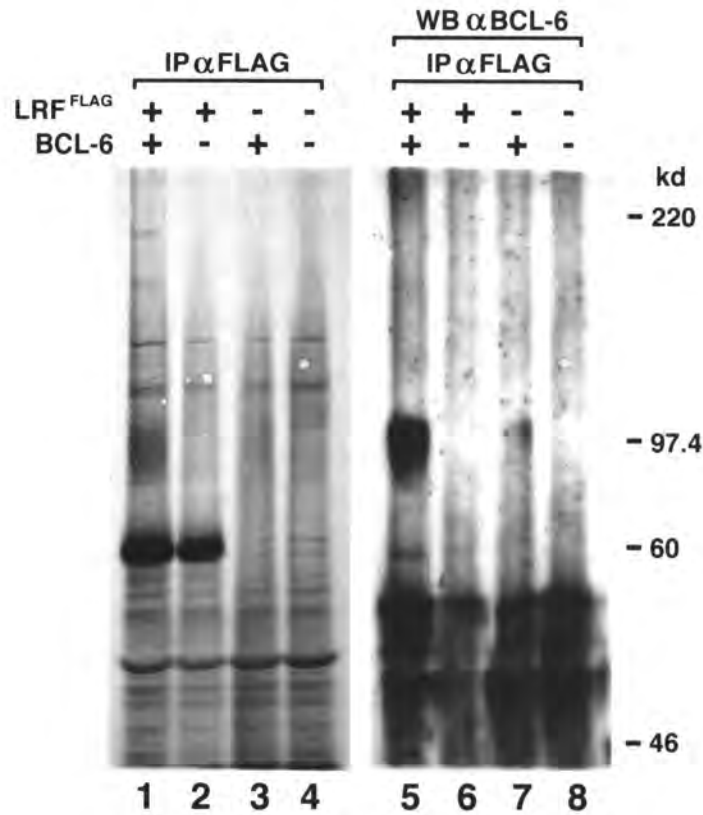


Fig. 17. Co-immunoprecipitation of BCL-6 and LRF proteins. COS-1 cells were transfected with 1 μ g of mLRF^{Flag} and/or BCL-6 expression vectors, as indicated. The ³⁵S methionine labelled transfected cell lysates were subjected to immunoprecipitation with anti-Flag M2 (Kodak) antibody (lanes 1-4). Lanes 5-6 represent a western blot of the immunoprecipitates shown in lanes 1-4 using a rabbit anti-BCL-6 antiserum.

and 7) or 2 μ g of empty pSG5 (Figure 17, lanes 4 and 8). 24 hours post transfection, cells were lysed in a high salt lysis buffer to inhibit any low specificity protein interactions and the 35 S methionine labelled lysed cells were immunoprecipitated with an anti-FLAG antibody (Figure 17, lanes 1-4). A western blot of the immunoprecipitated proteins was probed with a BCL-6 polyclonal antibody (Figure 17, lanes 5-8). The immunoprecipitation of Flag-LRF resulted in the co-immunoprecipitation of BCL-6. The interaction between LRF and BCL-6 was sufficiently strong for the co-immunoprecipitated, 35 S-labelled, BCL-6 to be visualised (Figure 17, lane 1) prior to probing with the BCL-6 specific antibody (Figure 17, lane 5).

3.3.1.3 Co-immunofluorescence

Upon transient transfection of CHO cells with 1 μ g of Flag-LRF or 1 μ g of BCL-6 pSG5 expression vectors, immunofluorescence experiments using the anti-FLAG and BCL-6 antibodies (see section 2.3.3 for procedure and section 2.3.5 for construct design) revealed that both proteins localised to the nucleus and in co-expressing cells their nuclear localisation pattern overlapped (Figure 18A). This is further supported by the observation that cells co-expressing BCL-6 and POZ-less Flag-LRF have a more mutually exclusive staining patterns with the latter being restricted to the perimeter of the nucleus (Figure 18B) and in some cases the cytoplasm of the cells. Clearly the POZ domain of LRF mediates its speckled nuclear localisation as has been confirmed for the POZ domain of BCL-6, see section 1.2.2. This co-localisation supports the *in vivo* interaction of BCL-6 and LRF.

3.3.2 The functional consequences of a BCL-6/ LRF interaction

Having shown that LRF can interact with BCL-6 *in vivo*, it was important to investigate if this heterodimerisation had an effect on BCL-6 mediated transcriptional repression. 293T cells were transiently transfected with 1 μ g of a TK luciferase reporter plasmid (pT109) in the absence (Figure 19 A and hatched bars) or presence (Figure 19 B and white bars) of three copies of the BCL-6 response element (Chang, Ye et al. 1996). The effect, on the respective luciferase reporter activities (see Figure 19 A and B), of co-transfecting BCL-6, Flag-LRF and POZ-less Flag-LRF pSG5 expression vectors in the combinations indicated in Figure 19 C-K, was examined (see section 2.3.4 for procedure and section 2.3.5 for construct design). The quantities of

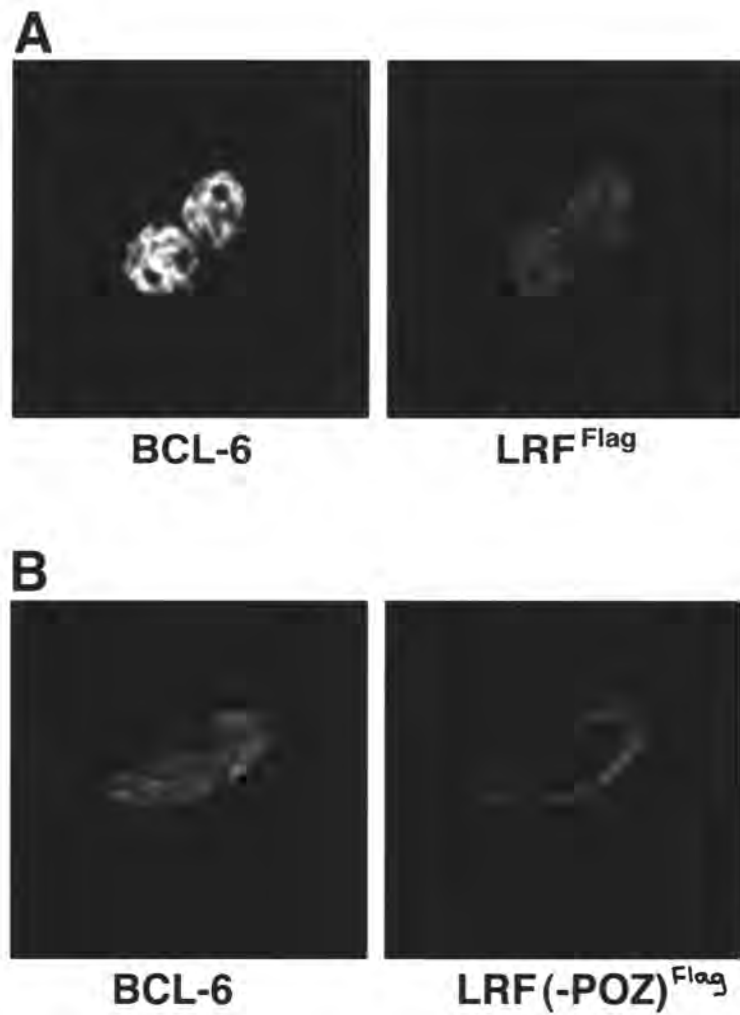


Fig. 18. Nuclear localisation patterns of co-transfected BCL-6 and LRF^{Flag} proteins (A) and co-transfected BCL-6 and LRF^{Flag}[-POZ] proteins (B). Each panel shows the immunofluorescence at a wavelength specific for the indicated individual protein. The FITC and TRITC conjugated secondary antibodies exhibited no cross-reactivity when incubated together against the appropriate primary antibody of singly transfected cells (not shown).

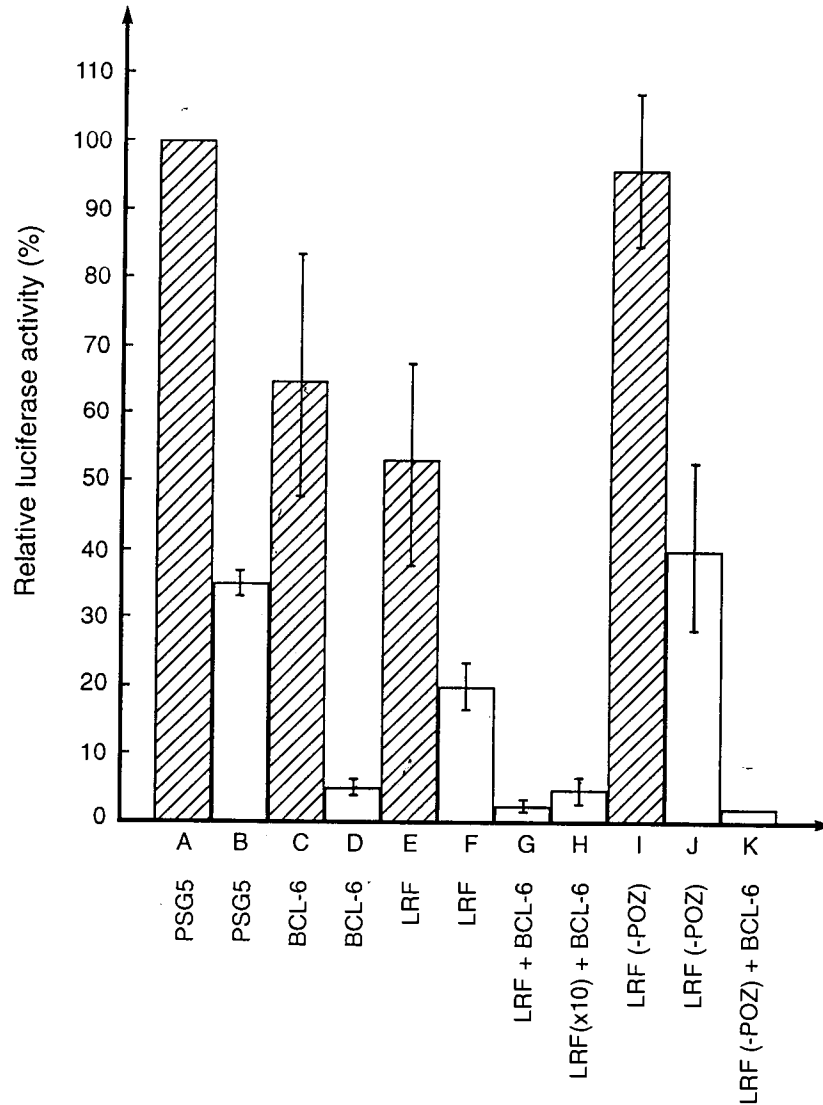


Fig. 19. The effect of mLRF on BCL-6 mediated transcriptional repression. Luciferase assays of 293T cells transiently transfected with 1µg of either the empty luciferase reporter vector (PT109, hatched bars) or the same vector with three copies of the BCL-6 response element (white bars) in the presence of 0.25µg of BCL-6, LRF or LRF[-POZ] pSG5 expression vectors, as indicated, except in lane H where 2.5µg of LRF pSG5 was transfected. The percentage luciferase activities are calculated relative to the empty vector transfection (lane A) and are the average of three separate transfection experiments.

transfected DNA are indicated in the Figure 19 legend. All luciferase activities were normalised for transfection efficiency by the co-transfection of 0.25 µg of a β-galactosidase expressing plasmid (see section 2.3.4).

As reported previously (see section 1.2.2), there was a low level endogenous repressive effect on the BCL-6 response element reporter construct (Figure 19 B) which was dramatically enhanced upon co-transfection of BCL-6 (Figure 19 D). It is unclear whether this low level repressive effect is a result of endogenous BCL-6 expressed in 293T cells or due to other endogenous proteins that recognise the BCL-6 response element (see section 1.2.2). The POZ-less LRF construct (LRF [-POZ]) was included as a negative control because in the yeast two-hybrid system this protein did not interact with BCL-6. When compared to the co-transfection of LRF[-POZ] and BCL-6 (Figure 19 K), the co-transfection of LRF with BCL-6 did not have a significant positive or negative effect on the BCL-6 mediated repression (Figure 19 G) even in 10 fold excess (Figure 19 H).

Despite this lack of effect on BCL-6 mediated repression, a comparison of the effect of LRF (Figure 19 E) and LRF [-POZ] (Figure 19 I) on the empty PT109 reporter construct indicates that LRF exerts a significant 'non-specific' general repression of basal transcription. Therefore, the apparent LRF mediated repression in the presence of the BCL-6 response element (Figure 19 F) might be a consequence of the combined non-specific effect of LRF on basal transcription (Figure 19E) and the above described endogenous repression (Figure 19B). Whilst the general repression of basal transcription by LRF appears to be dependant on the presence of the POZ domain, as highlighted in section 3.3.1.3, the altered nuclear localisation of LRF[-POZ] might prevent it from coming into contact with the transcriptional machinery in this assay.

3.4 Summary of results

In order to identify related proteins or potential protein interaction partners of PLZF, this study has led to the isolation of four POZ domain proteins. Expression patterns reveal that all four proteins might play a role in limb bud development. Two of these proteins, Bach 1 and Bach 2, have been independently published and represent a new class of POZ domain proteins which contain two protein interaction motifs. Of the two novel proteins

isolated, LRF is the most closely related to PLZF. LRF is located on human chromosome 19p13.3 and is a 60 kdalton POK protein comprising an N-terminal POZ domain and four C-terminal *Krüppel*-like zinc fingers. These two functional domains of LRF are highly conserved through evolution but are separated by a long stretch of unconserved amino acids. LRF exists as a number of isoforms and has a low level wide spread expression pattern. LRF has been shown to interact with BCL-6 in vivo and deletion analysis has revealed that this interaction is mediated through both the POZ domain and the zinc fingers of the respective partner proteins. In addition a unique intramolecular interaction between the POZ domain and zinc fingers of LRF has been demonstrated. In 293T cells, whilst expression of the LRF protein caused a general repressive effect on transcription, co-transfection of BCL-6 and LRF, did not alter significantly BCL-6 activity.

Chapter 4: Discussion

4.1 PLZF related POZ domain proteins

This study has led to the isolation of four mammalian PLZF related POZ domain proteins. POZ domain proteins have been implicated in a number of developmental processes and the functional importance of the POZ domain is highlighted by the presence of family members throughout metazoan evolution. Studies on the BAB protein suggest the existence of over 40 drosophila POZ domain proteins (Zollman, Godt et al. 1994), however neither PLZF nor LRF appear to have direct drosophila or yeast counterparts. These two POK proteins may have been derived from the same ancestral gene by chromosomal and/or gene duplication. Both genes possess a large exon (approximately 1.2Kb) encoding the N-terminus of the protein and the beginning of the zinc finger region. In contrast, the POZ domains of the two Bach proteins isolated are encoded by two separate exons suggesting a more divergent evolutionary progression which is in agreement with the presence of an alternative DNA binding motif. The existence of viral POZ domain proteins implies that cellular gene and/or exon capture by viral genome integration and subsequent transposition, might be responsible for the generation of chimeric POZ proteins with alternative C-terminal motifs.

PLZF as well as the drosophila POZ domain proteins, Abrupt and Bric-a-Brac, have been implicated in limb morphogenesis. It is clear however that it is not always possible to directly correlate the expression pattern of a gene with its actual function; PLZF is expressed in the fore and hindlimbs of the developing embryo but its knockout only affects hindlimb development (Hawe, Soares et al. 1996). Nevertheless, the higher levels of expression in limb bud of all four clones isolated in this study, suggests a role for them in limb development. In this respect, it will be interesting to see how LRF integrates within the regulatory gene network which governs limb development (Tabin 1991). For example, could the expression of LRF be influenced by the fibroblast growth factors, FGF4 and FGF8? Alternatively, or in addition, could LRF itself regulate the expression of Hox genes? The gradient of expression of LRF, across the proximo-distal axis of the limb bud, reflects the degree of proliferation of the mesenchymal cells which implies that LRF may participate in maintaining the proliferative state of these cells; a

role which has been suggested for the Hox genes (Duboule 1995). In contrast, the diffuse expression pattern for clone 2 does not shed any light on its potential role in the complex outgrowth and patterning processes that are regulated in distinct regions of the limb bud. This does not rule out the possibility that this protein exerts its effects by the formation of different regulatory heterodimeric complexes with other partner proteins of more restricted localisation. The BMP 7, which has a diffuse expression pattern at 10.5 dpc, becomes more restricted around the digit rudiments at later stages and is thought to mediate mesenchymal apoptosis (Luo, Hofmann et al. 1995). Likewise, the expression pattern of all four clones may prove to be more restricted at later stages and warrants further investigation.

4.2 Chromosomal localisation of LRF

The presence of segments of the genome on different mouse and human chromosomes is defined as linkage homology and is believed to be a consequence of multiple chromosomal rearrangements that occurred after the divergence of the lineages leading to the mouse and human genomes. A notable exception to this evolutionary rearrangement of the genome is the conservation of mouse and human gene localisation on the X chromosome (Copeland, Jenkins et al. 1993). To date there are nine published genes localised to human chromosome 19p13.3 and to the middle region of mouse chromosome 10, position 41 of the linkage map (Taylor, Burmeister et al. 1996). MatK (Megakaryocyte associated tyrosine kinase) is one of these nine genes and has also been mapped cytogenetically to band C1 of mouse chromosome 10. The symmetry of mouse chromosomes, as compared to human chromosomes, makes cytogenetic mapping less accurate than linkage recombination techniques which map genes relative to a common set of previously identified anchor loci on the chromosome (Avner, Amar et al. 1988). As is apparent from the schematic representation of mouse chromosome 10 in Figure 15A, the localisation of mLRF to band B 5.3 is in close proximity to band C1 and it is most likely that linkage analysis of mLRF will localise it to position 41 of the mouse chromosome 10 linkage map.

Numerous oncogenic deletions, amplifications and translocations involving human chromosome 19p13 have been reported as a result of the cytogenetic analysis of malignant cells (Mitelman 1994). More detailed molecular studies

have demonstrated the involvement of specific genes in oncogenic translocations. The E2A gene is localised on chromosome 19p13.3 and has been demonstrated to be present in a number of translocations associated with both acute lymphoblastic leukaemia (ALL) [t(1;19): Mellentin and Murre 1989; t(17;19): Inaba, Roberts et al. 1992] and chronic myeloid leukaemia [t(11;19): Morris, Whitman et al. 1988]. Multiple breakpoints at 19p13 have been demonstrated for another, non-E2A associated, t(11;19) translocation implicated in the pathogenesis of a number of haematopoietic malignancies including infantile pre-B ALL (Mitani, Sato et al. 1989). Whilst molecular analysis has identified the 19p13 localised gene, ENL, as being involved in this translocation, it is believed that additional, as yet unidentified, 19p13 genes will also be implicated (Yamamoto, Seto et al. 1993). It remains to be seen if LRF, like the POK protein encoding genes BCL-6 and PLZF, is involved in an oncogenic translocation.

4.3 A biologically significant POK protein heterodimerisation

The demonstrated BCL-6/LRF interaction is interesting in light of the unpublished results showing an interaction between PLZF and BCL-6. LRF is present in BCL-6 expressing cells such as DAUDI, a cell line derived from a Burkitts lymphoma patient (see section 3.2, Figure 12a) and its expression in lymphoid cell lines and tissues suggests a potential involvement in B cell development. In addition, both BCL-6 and LRF are expressed in highly proliferating cells; BCL-6 has been shown to be predominantly expressed in the rapidly proliferating germinal centre B lymphocytes whereas high levels of LRF expression are associated with the proliferating mesenchymal cells of the developing limb bud. In this respect, it will be interesting to establish the expression of BCL-6 in the developing mouse limb bud and likewise the expression of LRF in germinal centre B cells. In contrast, as highlighted in section 1.2, the expression patterns of BCL-6 and PLZF, at least during haematopoiesis, appear to be non-overlapping, making it hard to reconcile the relevance of their interaction in lymphopoiesis. LRF is therefore a more likely heterodimeric partner for BCL-6 during B-cell development. The initial premise that the clone with the highest degree of homology to the PLZF POZ domain might be an interaction partner proved to be wrong; the percentage homology between the POZ domains of BCL-6 and LRF is in fact only 33%. Perhaps the PLZF/BCL-6 interaction reflects the true biological interaction

between the PLZF related protein, LRF and BCL-6; the conserved amino acids of LRF and PLZF might prove to be those essential for the interaction with BCL-6.

Contrary to the *in vitro*, POZ domain mediated, interaction between TTK and GAGA, this study has shown that the heterodimerisation of BCL-6 and LRF, *in vivo*, requires both the POZ domain and the zinc finger region. On the other hand, the POZ domain was sufficient for *in vivo* homodimeric interactions. Although counter intuitive, this mutational study has not ruled out the possibility that the POZ domain of one partner protein might interact with the zinc fingers of the other and vice versa, rather than the interaction being through the common domains. The involvement of a conventionally recognised DNA binding region in protein interactions is not unprecedented; zinc fingers have been demonstrated to mediate the dimerisation of a number of transcriptional regulators such as the *krüppel* zinc finger proteins, Sp1 and erythroid *krüppel*-like factor (EKLF), and the GATA proteins (Crossley, Merika et al. 1995; Merika and Orkin 1995). Furthermore, the cysteine rich LIM domain, has a conformation very similar to two adjacent *krüppel*-like zinc fingers (Sanchez-Garcia and Rabbitts 1994). However, it is the involvement of this domain in protein interactions, rather than DNA binding, that has been demonstrated experimentally (Wadman, Li et al 1994; Agulnick, Taira, et al. 1996; Larson, Lavenir et al. 1996; Kuroda, Tokunaga et al. 1996; for example). Whilst this mutational analysis identified the motifs mediating a dimeric interaction, the speckled nuclear localisation of POK proteins suggests that they may exist in large, possibly heterogeneous, complexes. Such multimeric complexes might be stabilised by the zinc finger region of a particular POK protein interacting with one neighbouring protein whilst its POZ domain interacts with another protein.

The degree of conservation, through evolution, of a motif can be taken as a direct reflection of its functional importance. This study has corroborated this view by demonstrating that the POZ domain and to a greater extent, the zinc finger region of LRF are highly conserved motifs, both required for heterodimeric interactions. The breakdown in sequence conservation between the POZ domain and zinc fingers of LRF (within the P-Z region) is apparent for other POK proteins, including PLZF and BCL-6 (Cook, Gould et al. 1995; Fukuda, Miki et al. 1995), and may prove to be a common feature of this new family of transcriptional regulators. It was not surprising, therefore,

that this P-Z region proved not to be critical for protein interactions in the yeast two-hybrid system. However, the absence of the P-Z region in the LRF[P-Z] mutant did result in a conformation that was more readily available to form heterodimeric interactions. Perhaps this region acts as a hinge to mediate the intramolecular interaction of LRF resulting in a closed conformation which inhibits its heterodimeric interactions. An intramolecular interaction between the transactivation domain and zinc finger region of EKLF has recently been proposed to regulate its protein interaction and DNA binding capacity (Chen and Bieker 1996). Similarly, the Eves motif of cMyb has been implicated in mediating a phosphorylation dependant inhibitory intramolecular interaction with the protein's DNA binding motif (Dash, Orrico et al. 1996). Correspondingly, additional co-factor or phosphorylation induced modifications, may be required to overcome the proposed intrinsic, P-Z region mediated, inhibition. In this respect, it is worth noting the presence of a conserved proline dependant phosphorylation site within the P-Z region of LRF.

The demonstrated intramolecular interaction of LRF and most recently, PLZF (unpublished) but not BCL-6 might explain the discrepancies associated with POZ domain mediated inhibition of DNA binding, see section 1.1.4. Perhaps this inhibition of DNA binding is specific to those POK proteins that can form intramolecular interactions. This explanation is in conflict with the indirect *in vitro* data suggesting that the POZ domain mediated inhibition of ZID DNA binding is not a result of an interaction with the zinc fingers as it is transferable to heterologous DNA binding motifs (see section 1.1.4). However, it is clear that not all POZ proteins act in the same manner and results pertaining to the function of ZID may not apply to all POK proteins.

4.4 Functional significance of a BCL-6/ LRF interaction

This study has demonstrated that LRF is co-expressed with BCL-6 in the lymphoid compartment, physically interacts with BCL-6 *in vivo* and colocalises with BCL-6 in discrete punctuated regions of the nucleus. This strongly suggests that these two POK proteins may function in transcriptional regulation as homo and heterodimeric complexes, their formation and resultant effect being determined by the respective levels of expression. In this respect, the preliminary co-transfection experiments were unexpected.

However, there may be a number of reasons why these initial experiments failed to indicate any effect of co-expressed LRF on the transcriptional activity of BCL-6. The concentration of the interacting proteins may have a critical saturation point which needs to be characterised. In addition, in other cell types, such as BCL-6 expressing cells, essential co-factors might be present which would reflect a more in vivo setting. Alternatively, BCL-6 might be having an effect on the transcriptional activity of LRF, rather than the other way around. To investigate this further, ideally the response element of LRF must be identified. In the meantime, however, the effect of co-transfecting BCL-6 with an LRF-Gal4 fusion chimera in a GAL4 response element reporter system could be examined.

The identified variant 5' sequences of LRF (see section 3.2) may correspond to different untranslated regions, involved in translational regulation, or to different N-terminal isoforms of the protein. If they exist, these variant N-terminal isoforms might mediate the transcriptional activity of LRF in complex with BCL-6. This is in analogy to the existence of alternatively spliced N-terminal domains of both Mxi1 and Mad which interact with the repressor, mSin3, see section 1.1.2. At some stages of development LRF may be able to actively repress or activate transcription whereas alternative isoforms at other stages might be transcriptionally silent but may still play a role in multimeric complex formation. Therefore, although this study has identified the regions mediating an in vivo heterodimeric interaction between BCL-6 and LRF, the 'true' physiological interaction may comprise an alternative isoform of LRF with additional transcriptional regulatory capacity. The development of polyclonal antibodies will resolve this possibility by determining the size of endogenous LRF protein in BCL-6 expressing cells. It will then be possible to investigate the existence of alternative splice variants of LRF at different stages of development.

The presence of additional N-terminal sequences on wild type LRF might prove to enhance or inhibit the repressive effect of BCL-6. If this does prove to be the case then the lack of effect of LRF on BCL-6 mediated repression in this current study might be because the heterodimer binds to the same response element as the BCL-6 homodimer, in analogy to the Max regulatory network. If the heterodimer bound to an alternative binding site, even in the absence of the putative N-terminal regulatory region, there would have been

an appreciable effect on BCL-6 mediated repression; in complex with LRF, BCL-6 would have been prevented from binding to its response element.

The preliminary co-transfection experiments did highlight an LRF mediated repression of basal transcription. It will be of interest to see if this LRF mediated repression is apparent for other basal promoters as this will determine whether or not LRF is able to act as a general co-repressor without the need for a specific DNA binding site. Co-immunoprecipitation or yeast two hybrid assays will determine which members of the preinitiation complex (PIC) interact with LRF (see section 1.1.2). As a homodimer, or a monomer, LRF may exert its effects on basal transcription but as a heterodimer the same protein may have more cell specific transcriptional regulatory effects. In this respect, the *Krüppel* protein (Kr) has been shown to interact directly, as a monomer, with TFIIB and activate transcription whereas the Kr homodimer binds to TFIIE and represses transcription (Sauer, Fondell et al. 1995).

To contribute towards an understanding of BCL-6 associated NHL disease progression, it will be important to clarify how the BCL-6/LRF heterodimeric interaction relates to lymphopoiesis. BCL-6 knock out (-/-) mice have confirmed a specific role for BCL-6 in germinal centre formation (Dent, Shaffer et al. 1997) and it is clearly important to establish what role, if any, LRF plays at this stage of B cell development. By virtue of its wide spread expression pattern, LRF may exert its effects by interacting, in addition to BCL-6, with a number of other spatially and temporally restricted POK proteins. In light of this, results of LRF -/- mice, now in progress, may be informative with respect to its role in a number of developmental processes. In the event that LRF -/- mice are embryonic lethal, the role of LRF in haematopoiesis will need to be examined by generating recombinant LRF-/- ES cells and inducing their differentiation in vitro (Wiles 1993). Alternatively the role of LRF, specifically in lymphopoiesis, could be examined by injecting LRF-/- ES cells into RAG-/- mice. In this in vivo model, since RAG1 is required for V(D)J recombination of immunoglobulin and T cell receptor genes, mature T and B cells will only originate from the injected LRF-/- ES cells (Chen, Shinkai et al. 1994). However, in order to assess the role of LRF/BCL-6 heterodimers it may be necessary to target the inactivation of LRF as close as possible to the germinal centre stage of B cell development. This could be achieved by using the Cre-loxP knock out system (Kilby, R et al.

1993). Cre is the recombinase of bacteriophage P1 and is able to mediate the specific excision of DNA, flanked by loxP recognition sites. Intercrossing transgenic mice containing LRF sequence flanked by loxP sites with transgenic mice containing the cre gene under the control of the Ig heavy chain enhancer ($E\mu$), in theory, will generate mice in which Cre will excise the LRF gene only once cells have reached the pre-B cell stage of development (see section 1.2.2).

4.5 Transcriptional regulation by POK protein dimerisation: a model

The work of Bardwell and Triesman (1994) on the ZID POK protein suggests that as homodimers these proteins are unable to bind DNA and only as heterodimers are the DNA binding zinc fingers released from the dimeric interaction thus enabling their interaction with DNA, leading to the activation or repression of transcription. However, the demonstrated transcriptional activity of POK proteins, such as BCL-6, would suggest that as homodimers, they can bind DNA and exert an effect on transcriptional activity unless in all reporter assay systems used the partner protein is present endogenously in the transfected cells leading to the formation of DNA binding heterodimers, which seems unlikely.

Based on the deletion analysis of this study, an alternative model of transcriptional regulation by POK proteins can be proposed, see Figure 20. Since the zinc fingers are not involved in the homodimeric interaction of BCL-6 they are free to bind DNA (Figure 20A). The model implies that LRF exists as a monomer caught up in an intramolecular interaction which inhibits its DNA binding and/or dimeric interactions. The inhibitory conformation of this monomer is mediated by phosphorylation/co-factor regulation (Figure 20B). A direct phosphorylation event could enhance or inhibit the closed intramolecular conformation or a co-factor could bind and enhance the closed conformation as a consequence of its phosphorylation or dephosphorylation. This clearly provides a means of regulating LRF and possibly BCL-6, activity by extracellular stimuli mediated signal transduction. This could be tested by examining the DNA binding and dimerisation capacity of LRF in the presence of specific kinases/phosphatase inhibitors or as a consequence of mutating the potential phosphorylation site within the P-Z region of the protein.

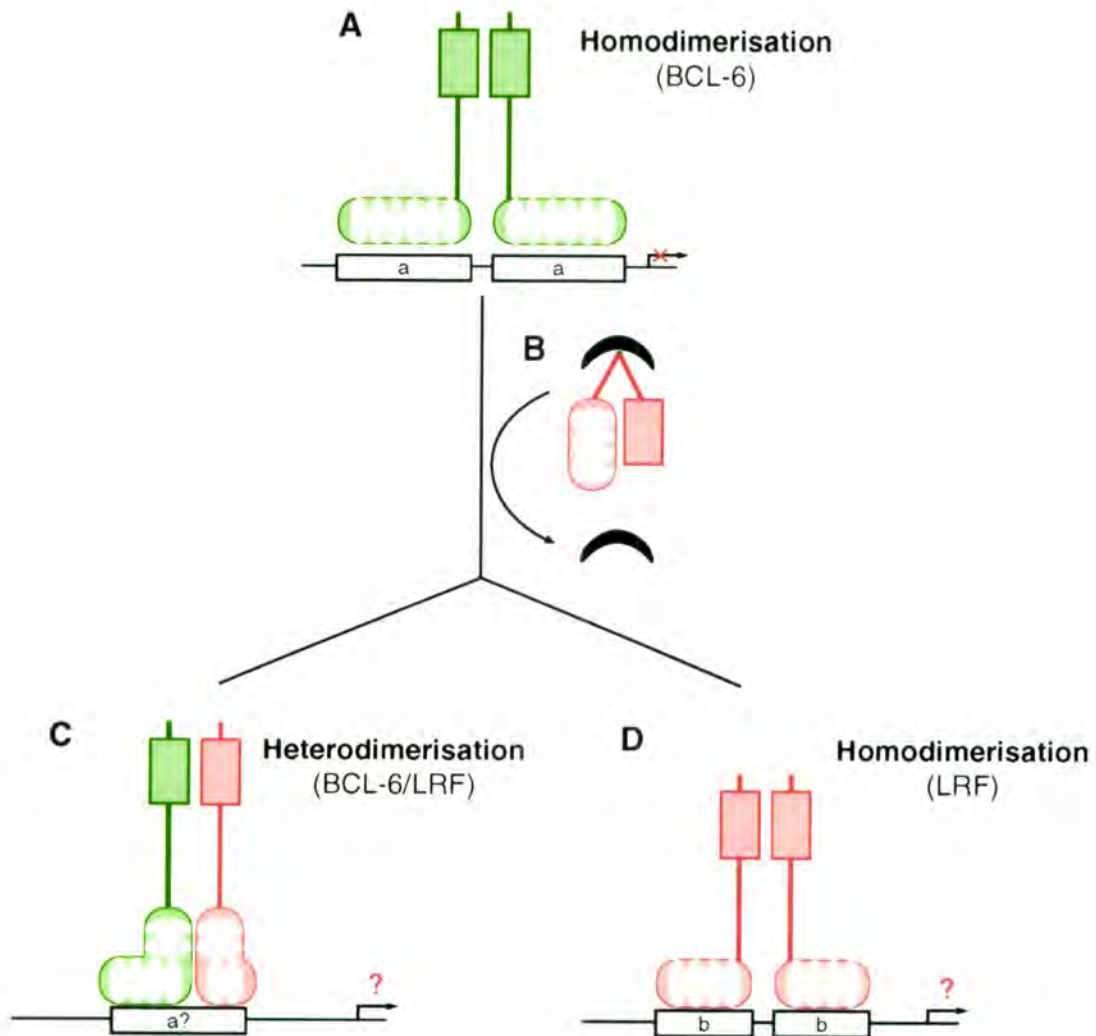


Fig. 20. A model of POK protein mediated transcriptional regulation. **A.** BCL-6 homodimeric mediated transcriptional repression. **B.** Co-factor or phosphorylation (black crescent) mediated regulation of the intramolecular interaction of the LRF monomer. **C.** Once relieved from its intramolecular interaction, LRF heterodimerises with BCL-6 and the resultant complex may bind to the same response element or an alternative response element and exert a positive or negative effect on transcriptional initiation. **D.** Alternatively, in preference to **C**, at an appropriate concentration, LRF may bind a distinct response element as a homodimer.

Depending on the relative abundance of the partner proteins, an extracellular signal could induce a conformational change leading to the generation of DNA binding heterodimers (Figure 20C) or homodimers (Figure 20D). As a heterodimer, some of the zinc fingers may be uninvolved in the interaction and will be free to bind to DNA (Figure 20C). Further clarification of this will involve a more extensive deletion analysis to determine which zinc fingers are required for the protein interaction and which are required for the DNA binding. However, based on the results of the preliminary transient transfection experiments it is anticipated that the BCL-6/LRF heterodimer will bind to the same DNA response element as the BCL-6 homodimer. However it is still possible that an alternative consensus is recognised by the heterodimer. It is also conceivable that once relieved from its intramolecular conformation, LRF binds its own distinct response element as a monomer rather than as a homodimer (Figure 20D).

4.6 In conclusion

This study has identified a previously uncharacterised POK protein, LRF, which is closely related to PLZF, at least in the first half of the POZ domain. LRF has been shown to homodimerise and heterodimerise with the BCL-6 oncoprotein. In addition, this study has demonstrated a novel, potentially regulatory, intramolecular interaction between the POZ domain and zinc fingers of LRF; a property possibly shared by some other POK proteins. The ability of POK proteins to form homo and heterodimers, which may recognise different response elements, increases the number of potential target genes that can be regulated by a limited number of proteins. The proposed regulation of dimer formation, by phosphorylation, provides a link between cell surface generated events and POK protein mediated nuclear transcriptional regulation.

References

- Agulnick, A. D., M. Taira, et al. (1996). "Interactions of the LIM-domain-binding factor Ldb1 with LIM homeodomain proteins." Nature **384**: 270-272.
- Ahuja, H., M. Bar-Eli, et al. (1986). "Alterations in the p53 gene and the clonal evolution of the blast crisis of chronic myelogenous leukemia." Proc Natl Acad Sci USA **86**: 6783-6787.
- Albagli, O., P. Dhordain, et al. (1996). "Multiple domains participate in distance independent LAZ3/BCL6 mediated transcriptional repression." Biophys Biochem Res Commun **220**: 911-915.
- Albagli, O., P. Dhordain, et al. (1995). "The BTB/POZ Domain: A new protein-protein interaction motif common to DNA and actin-binding proteins." Cell Growth Diff **6**: 1193-1198.
- Alcalay, M., D. Zangrilli, et al. (1992). "Expression pattern of the RAR α -PML fusion gene in acute promyelocytic leukemia." Proc Natl Acad Sci USA **89**: 4840-4844.
- Altingmees, M., P. Kretz, et al. (1993). "The λ Zap expression vector-factors affecting eukaryotic expression, phage particle presentation and library representation." FASEB Journal **7**(7): A1303.
- Arsura, M., A. Deshpande, et al. (1995). "Variant max protein, derived by alternative splicing, associates with c-Myc in vivo and inhibits transactivation." Molecular and Cellular Biology **15**(12): 6702-6709.
- Avery, O. T., C. M. MacLeod, et al. (1944). "Studies on the chemical nature of the substances inducing transformation of pneumococcal types. Induction of transformation by a deoxyribonucleic acid fraction isolated from Pneumococcus Type III." Journal of Exp Med **79**: 137-158.
- Aviv, H. and P. Leder (1972). "Purification of biologically active globin mRNA by chromatography on oligothymidylic acid-cellulose." Proc Natl Acad Sci USA **69**: 1408-1412.
- Avner, P., L. Amar, et al. (1988). "Genetic analysis of the mouse using interspecific crosses" Trends in Genetics **4**: 18.
- Ayer, D., L. Kretzner, et al. (1993). "Mad: A heterodimeric partner for Max that antagonises Myc transcriptional activity." Cell **72**: 211-222.
- Ayer, D., Q. A. Lawrence, et al. (1995). "Max-Max transcriptional repression is mediated by ternary complex formation with mammalian homologs of yeast repressor Sin3." Cell **80**: 767-776.
- Bardwell, V. J. and R. Treisman (1994). "The POZ domain: A conserved protein-protein interaction motif." Genes and Development **8**: 1664-1677.
- Bastard, C., C. Deweindt, et al. (1994). "LAZ3 rearrangements in Non-hodgkin's lymphoma: Correlation with histology, immunophenotype, karyotype, clinical outcome in 217 patients." Blood **83**: 2423-2427.
- Bellefroid, E. J., D. A. Poncelet, et al. (1991). "The evolutionarily conserved Krüppel-associated box domain defines a subfamily of eukaryotic multifingered proteins." Proc Natl Acad Sci USA **88**: 3608-3612.
- Benbrook, D. M. and N. C. Jones (1990). "Heterodimer formation between CREB and JUN proteins." Oncogene **5**: 295-302.

References

- Bianco, J., I.-M. Wang, et al. (1995). "Transcription factor TFIIB and the vitamin D receptor cooperatively activate ligand-dependent transcription." Proc Natl Acad Sci USA **92**: 1535-1539.
- Bjorklund, S. and Y.-J. Kim (1996). "Mediator of transcriptional regulation." Trends in Biological Science **21**: 335-337.
- Blackwood, E. M. and R. N. Eisenham (1991). "Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA binding complex with Myc." Science **251**: 1211-1217.
- Brand, N., M. Petkovich, et al. (1988). "Identification of a second human retinoic acid receptor." Nature **332**: 850-853.
- Breeden, L. and K. Nasmyth (1985). "Regulation of the yeast HO gene." Cold Spring Harbor Symp Quant Biol **50**: 643-650.
- Breitman, T. and S. Collins, BR (1981). "Induction of differentiation of the human promyelocytic cell in HL-60 by retinoic acid." Proc Natl Acad Sci USA **77**(2936-2940).
- Brown, J. L. and C. Wu (1993). "Repression of *Drosophila* pair-rule segmentation genes by ectopic expression of *tramtrack*." Development **117**: 45-58.
- Burley, S. K. and R. G. Roeder (1996). "Biochemistry and structural biology of transcription factor-II D." Annual Review Biochem **65**: 769-799.
- Castaigne, S., C. Chomienne, et al. (1990). "All-trans retinoic acid as a differentiation therapy for acute promyelocytic leukemia. I. Clinical results." Blood **76**: 1704-1709.
- Cattoretti, G., C.-C. Chang, et al. (1995). "The BCL-6 protein is expressed in germinal-centre B cells." Blood **86**: 28-37.
- Cavarec, L., S. Jenson, et al. (1997). "Molecular cloning and characterisation of a transcription factor for the copia retrotransposon with homology to the BTB-containing Lola neurogenic factor." Molecular and Cellular biology **17**(1): 482-494.
- Chambon, P. (1994). "Molecular and genetic dissection of retinoid signalling pathway." Gene **135**: 223-228.
- Chang, C.-C., B. H. Ye, et al. (1996). "BCL-6, a POZ/zinc finger protein, is a sequence-specific transcriptional repressor." Proc Natl Acad Sci USA **93**: 6947-6952.
- Chardin, P., G. Courtois, et al. (1991). "The KUP gene, located on human chromosome 14, encodes a protein with two distant zinc fingers." Nucleic Acids Research **19**(7): 1431-1436.
- Chen, J., Y. Shinkai, et al. (1994). "Probing immune functions in RAG-deficient mice." Curr. Opin. Immunol **6**: 313-319.
- Chen, S. J., A. Zelent, et al. (1993). "Rearrangements of the retinoic acid receptor alpha and promyelocytic leukemia zinc finger genes resulting from t(11;17)(q23;q21) in a patient with acute promyelocytic leukemia." J Clin Invest **91**(5): 2260-2267.
- Chen, W., S. Zollman, et al. (1995). "The BTB Domain of *bric a brac* mediates dimerization in vitro." Molecular and Cellular Biology **15**: 3424-3429.
- Chen, X. and J. J. Bieker (1996). "Erythroid kruppel-like factor (EKLF) contains a multifunctional transcriptional activation domain important for inter- and intramolecular interactions." EMBO Journal **15**(21): 5888-5896.

References

- Chen, Z., N. J. Brand, et al. (1993). "Fusion between a novel *Kruppel*-like zinc finger gene and the retinoic acid receptor- α locus due to a variant t(11;17) translocation associated with acute promyelocytic leukaemia." EMBO Journal **12**: 1161-1167.
- Chirgwin, J. M., A. E. Przybyla, et al. (1979). "Isolation of biologically active ribonucleic acid from sources enriched in ribnuclease." Biochem **18**: 5294-5299.
- Collins, S. J. (1987). "The HL-60 promyelocytic leukemia cell line: proliferation, differentiation, and cellular oncogene expression." Blood **70**: 1233-1244.
- Cook, M., A. Gould, et al. (1995). "Expression of the zinc-finger gene PLZF at rhombomere boundaries in the vertebrate hindbrain." Proc Natl Acad Sci USA **92**(6): 2249-2253.
- Cooley, L. and W. E. Theurkauf (1994). "Cytoskeleton functions during *Drosophila* oogenesis." Science **266**: 590-596.
- Cooper, M. D. (1987). "B lymphocytes, normal development and function." New Engl Journal Med **317**: 1452.
- Copeland, N., N. Jenkins, et al. (1993). "A genetic linkage map of the mouse: Current applications and future prospects." Science **262**: 57-66.
- Corey, S. J., J. Locker, et al. (1994). "A non-classical translocation involving 17q12 (retinoic acid receptor alpha) in acute promyelocytic leukemia (APML) with atypical features." Leukemia **8**(8): 1350-3.
- Crossley, M., M. Merika, et al. (1995). "Self association of the erythroid transcription factor GATA-1 mediated by its zinc finger domain." Molecular and Cellular biology **15**(5): 2448-2456.
- Das, G., C. S. Hinkley, et al. (1995). "Basal promoter elements as a selective determinant of transcriptional activator function." Nature **374**: 657-660.
- Dash, A., F. Orrico, et al. (1996). "The Eves motif mediates both intermolecular and intramolecular regulation of c-Myb." Genes and development **10**: 1858-1869.
- Dawson, M. I., E. Elstner, et al. (1994). "Myeloid differentiation mediated through retinoic acid receptor/retinoic X receptor (RXR) not RXR/RXR pathway." Blood **84**(2): 446-52.
- De La Brousse, F. C. and S. L. Mcknight (1993). "Glimpses of allostery in the control of eukaryotic gene-expression." Trends in Genetics **9**(5): 151-154.
- de The, H., C. Chomienne, et al. (1990). "The t(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor alpha gene to a novel transcribed locus." Nature **347**(6293): 558-61.
- Dent, A., A. Shaffer, et al. (1997). "Control of inflammation, cytokine expression and germinal centre formation by BCL-6." Science **276**: 589-592.
- Deweindt, C., O. Albagli, et al. (1995). "The LAZ3/BCL6 oncogene encodes a sequence-specific transcriptional inhibitor: a novel function for the BTB/POZ domain as an autonomous repressing domain." Cell growth and differentiation **6**: 1495-1503.
- Dhordain, P., O. Albagli, et al. (1995). "The BTB/POZ domain targets the LAZ3/BCL6 oncoprotein to nuclear dots and mediates homomerisation in-vivo." Oncogene **11**(12): 2689-2697.

References

- DiBello, P. R., D. A. Withers, et al. (1991). "The *Drosophila broad-complex* encodes a family of related proteins containing zinc fingers." Genetics **129**: 385-397.
- Dong, S., J. Zhu, et al. (1996). "Amino-terminal protein-protein interaction motif (POZ-domain) is responsible for activities of the promyelocytic leukemia zinc finger-retinoic acid receptor- α fusion protein." Proc Natl Acad Sci USA **93**: 3624-3629.
- Dorn, R., V. Krauss, et al. (1993). "The enhancer of position-effect variegation of *Drosophila*, *E(var)3-93D*, codes for a chromatin protein containing a conserved domain common to several transcriptional regulators." Proc Natl Acad Sci USA **90**: 11376-11380.
- Duboule, D. (1995). "Vertebrate hox genes and proliferation- an alternative pathway to homeosis." Current opinion in genes and development **5**(4): 525-528.
- DuBridge, R. B., P. Tang, et al. (1987). "Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system." Molecular and Cellular Biology **7**(1): 379-387.
- Duprez, D. M., K. Kostakopoulou, et al. (1996). "Activation of Fgf-4 and HoxD gene expression by BMP-2 expressing cells in the developing chick limb." Development **122**: 1821-1828.
- Dyck, J. A., G. G. Maul, et al. (1994). "A novel macromolecular structure is a target of the promyelocyte- retinoic acid receptor oncoprotein." Cell **76**(2): 333-43.
- Eaves, C. (1996). "Myelopoiesis" . In Leukemia (Ed ES Henderson, TS Lister, MF Greaves). W B Saunders Company: 46-64.
- Eichbaum, Q. G. (1994). "Restriction of interferon-gamma responsiveness and basal expression of the myeloid human FC-gamma-R1B gene is mediated by a functional Pu.1 site and a transcription initiator consensus." Journal of Exp Med **179**(6): 1985-1986.
- El-Baradi, T. and T. Pieler (1991). "Zinc finger proteins: what we know and what we would like to know." Mechanisms of Development **35**: 155-169
- Fagioli, M., M. Alcalay, et al. (1992). "Alternative splicing of PML transcripts predicts coexpression of several carboxy-terminally different protein isoforms." Oncogene **7**: 1083-1091.
- Farkas, G., J. Gausz, et al. (1994). "The *trithorax-like* encodes the *Drosophila* GAGA factor." Nature **371**: 806-808.
- Fenaux, P., E. Wattel, et al. (1994). "Prolonged follow-up confirms that all-trans retinoic acid followed by chemotherapy reduces the risk of relapse in newly diagnosed acute promyelocytic leukemia. The French APL Group [letter]." Blood **84**(2): 666-7.
- Fields, S. and O. Song (1989). "A novel genetic system to detect protein-protein interactions." Nature **340**: 245-247.
- Friend, C., M. C. Patuleia, et al. (1965). "Erythrocytic maturation *in vitro* of murine (friend) virus-induced leukemic cells." National Cancer Institute Monograph **22**: 505-522.
- Fukuda, T., T. Miki, et al. (1995). "The murine BCL6 gene is induced in activated lymphocytes as an immediate early gene." Oncogene **11**: 1657-1663.
- Gietz, R. and R. Woods (1994). "High efficiency transformation with lithium acetate." In molecular genetics of yeast: A practical approach. (ed. J R Johnston). Oxford university press: 121-134.

References

- Giniger, E., K. Tietje, et al. (1994). "Lola encodes a putative transcription factor required for axon growth and guidance in *Drosophila*." Development **120**: 1385-1398.
- Gluzman, Y. (1981). "SV40-transformed simian cells support the replication of early SV40 mutants." Cell **23**: 175-182.
- Godin, I. E., J. A. Garcia Porrero, et al. (1993). "Para-aortic splanchnopleura from early mouse embryos contains B1A cell progenitors." Nature **364**: 67-70.
- Goodrich, J. A. and Tjian (1994). "Transcription factors II E and II H and ATP hydrolysis direct promoter clearance by RNA-polymerase-II." Cell **77**: 145-156.
- Graham, F., J. Smiley, et al. (1977). "Characterisation of a human cell line transformed by DNA from human adenovirus type 5." J. Gen. Virol **36**: 59-74.
- Greaves, M. F. (1996). "The new biology of leukaemia". In Leukemia (ed ES Henderson, TA Lister, MF Greaves. W B Saunders Company: 34-45.
- Green, S. and I. Issemann (1988). "A versatile in vivo and in vitro eukaryotic expression vector for protein engineering." Nucleic acids research **16**(1).
- Greenberger, J. S., M. A. Sakakeeny, et al. (1983). "Demonstration of permanent factor-dependent multipotential (erythroid/neutrophil/basophil) hematopoietic progenitor cell lines." Cell Biology **60**: 2931-2935.
- Grignani, F., M. Fagioli, et al. (1994). "Acute promyelocytic leukemia: from genetics to treatment." Blood **83**(1): 10-25.
- Guarente, L. (1983). "Yeast promoters and lacZ fusions designated to study expression of clones genes in yeast. " In methods and enzymology. (eds R Wu, L Grossman, K Moldave). New York: Academic press, Inc: 181-191.
- Guidez, F., W. Huang, et al. (1994). "Poor Response to all-trans retinoic acid in a t(11;17) PLZF/RAR α AML3 patient." Leukaemia **8**: 312-317
- Hanna-Rose, W. and U. Hansen (1996). "Active repression mechanisms of eukaryotic transcription repressors." Trends in Genetics **12**(6): 229-234.
- Harper, J., G. Adami, et al. (1993). "The p21 CDK-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases." Cell **75**: 805-816.
- Hawe, N., V. Soares, et al. (1996). "Targeted disruption of the PLZF gene of acute promyelocytic leukemia results in phocomelia and altered spermatogenesis." Blood **88**(10): S1151.
- Heisterkamp (1983). "Localisation of cABL oncogene adjacent to the translocation breakpoint in Chronic Myelocytic leukemia." Nature **306**: 239-242.
- Henikoff, S. (1990). "Position-Effect Variegation After 60 Years." Trends in Genetics **6**: 422-426.
- Hershey, A. D. and M. Chase (1952). "Independent functions of viral protein and nucleic acid in growth of bacteriophage." J Gen Physiol **36**: 39-56.
- Heyworth, C. M., T. M. Dexter, et al. (1990). "The role of hemopoietic growth factors in self-renewal and differentiation of IL-3-dependent multipotential stem cells." Growth Factors **2**: 197-211.

References

- Hillstead, L. (1957). "Acute Promyelocytic Leukemia." Acta Med Scand **69**: 189.
- Horlein, A. J., A. M. Naar, et al. (1995). "Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor." Nature **377**: 397-403.
- Horowitz, H. and C. Berg (1996). "The Drosophila pipsqueak gene encodes a nuclear BTB - domain-containing protein required early in oogenesis." Development **122**: 1859-1871.
- Hou, J., U. Schindler, et al. (1994). "An interleukin-4-induced transcription factor: IL-4 STAT." Science **265**: 1701-1705.
- Hu, S., D. Fambrough, et al. (1995). "The drosophila abrupt gene encodes a BTB-zinc finger regulatory protein that controls the specificity of neuromuscular connections." Genes and Development **9**: 2936-2948.
- Huang, M., Y. Ye, et al. (1988). "Use of all-trans retinoic acid in the treatment of acute promyelocytic leukemia." Blood **72**: 567-572.
- Hunter, T. and M. Karin (1992). "The regulation of transcription by phosphorylation." Cell **70**: 375-387.
- Igarashi, K., K. Kataoka, et al. (1994). "Regulation of transcription by dimerisation of erythroid factor NF-E2 p45 with small maf proteins." Nature **367**: 568-572.
- Inaba, T., W. Roberts, et al. (1992). "Fusion of the leucine zipper gene HLF to the E2A gene in human acute B-lineage leukemia." Science **257**: 531-537.
- Jacob, F. and J. Monod (1961). "Genetic regulatory mechanisms in the synthesis of proteins." J. Molecular Biology **3**: 318-356.
- Jones-Villeneuve, E. M. V., M. W. McBurney, et al. (1982). "Retinoic acid induces embryonal carcinoma cells to differentiate into neurones and glial cells." J Cell Biology **94**: 253-262.
- Jongens, T. A., L. D. Ackerman, et al. (1994). "Germ cell-less encodes a cell type-specific nuclear pore associated protein and functions early in the germ-cell specification pathway of Drosophila." Genes and Development **8**: 2123-2136.
- Kabarowski, J. H. S., P. B. Allen, et al. (1994). "A temperature sensitive p210 BCR-ABL mutant defines the primary consequences of BCR-ABL tyrosine kinase expression in growth factor dependant cells." EMBO J **13**: 5887-5895.
- Kakizuka, A., W. H. Miller, et al. (1991). "Chromosomal translocation t(15;17) in human acute promyelocytic leukemia fuses RAR α with a novel putative transcription factor, PML." Cell **66**: 663-674.
- Kaplain, M. H., U. Schindler, et al. (1996). "STAT6 is required for mediating responses to IL-4 and for the development of TH2 cells." Immunity **4**: 313-319.
- Karim, F., G. Guild, et al. (1993). "The drosophila broad-complex plays a key role in controlling ecdysone-regulated gene expression at the onset of metamorphosis." Development **118**(3): 977-988.
- Kastner, P., A. Perez, et al. (1992). "Structure, localisation and transcriptional properties of two classes of retinoic acid receptor α fusion proteins in acute promyelocytic leukemia (APL): structural similarities with a new family of oncoproteins." EMBO J **11**: 629-642.

References

- Kataoka, K., K. Igarashi, et al. (1995). "Small Maf proteins heterodimerise with Fos and may act as competitive repressors of the NF-E2 transcription factor." Molecular and Cellular biology **15**(4): 2180-2190.
- Kawamata, N., T. Miki, et al. (1994). "The organization of the BCL6 gene." Leukemia **8**(8): 1327-1330.
- Keegan, A., K. Nelms, et al. (1994). "An IL-4 receptor region containing an insulin receptor motif is important for IL-4 mediated IRS-1 phosphorylation and cell growth." Cell **76**: 811-820.
- Keydar, I., L. Chen, et al. (1979). "Establishment and characterisation of a cell line of human breast carcinoma origin." Eur. J. Cancer **15**: 659-670.
- Kilby, N. J., S. M. R, et al. (1993). "Site-specific recombinases: tools for genome engineering." Trends in Genetics **9**: 413-421.
- Klein, G., E. Klein, et al. (1967). "Search for host defenses in Burkitt's lymphoma: membrane immunofluorescence tests on biopsies and tissue culture lines." Cancer Research **28**: 1300-1310.
- Kliwer, S. A., K. Umesono, et al. (1992). "Retinoid X receptor interacts with nuclear receptors in retinoic acid, thyroid hormone and vitamin D₃ signalling." Nature **355**: 446-449.
- Knowles, B., C. Howe, et al. (1980). "Human hepatocellular carcinoma cell lines secrete the major plasma proteins and hepatitis B surface antigen." Science **209**: 497-499.
- Koeffler, H. P. and D. W. Golde (1978). "Acute myelogenous leukaemia: a human cell line responsive to colony stimulating activity." Science **200**: 1153-1154.
- Koken, M. H., F. Puvion-Dutilleul, et al. (1994). "The t(15;17) translocation alters a nuclear body in a retinoic acid-reversible fashion." Embo J **13**(5): 1073-83.
- Koonin, E. V., T. G. Senkevich, et al. (1992). "A family of DNA virus genes that consists of fused portions of unrelated cellular genes." Trends Biochem Sci **17**: 213-214.
- Kotanides, H. and N. C. Reich (1993). "Requirement of tyrosine phosphorylation for rapid activation of a DNA binding factor IL-4." Science **262**: 1265-1267.
- Kozak, M. (1986). "Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes." Cell **44**: 283-292.
- Kretzner, L., E. M. Blackwood, et al. (1992). "Myc and Max proteins possess distinct transcriptional activities." Nature **359**: 426-429.
- Kuroda, S., C. Tokunaga, et al. (1996). "Protein-protein interactions of zinc finger LIM domains with Protein Kinase C." Journal of biological chemistry **271**(49): 31029-31032.
- Larson, R. C., I. Lavenir, et al. (1996). "Protein dimerisation between Lmo2 (Rbtn2) and Tal1 alters thymocyte development and potentiates T cell tumorigenesis in transgenic mice." EMBO Journal **15**(5): 1021-1027.
- Le, X. F., P. Yang, et al. (1996). "Analysis of the growth and transformation suppresser domains of promyelocytic leukemia gene, PML." J Biol Chem **271**(1): 130-5.
- LeBien, T. W. (1996). "Lymphopoiesis." In *Leukemia* (ed ES Henderson, TA Lister, MF Greaves). W B Saunders Company: 65-82.

- Leroy, P., A. Krust, et al. (1991). "Multiple isoforms of the mouse retinoic acid receptor- α are generated by alternative splicing and differential induction by retinoic acid." EMBO Journal 10: 59-69.
- Levine, M. and J. L. Manley (1989). "Transcriptional repression of eukaryotic promoters." Cell 59: 405-408.
- Licht, J. D., C. Chomienne, et al. (1995). "Clinical and molecular characterisation of a rare syndrome of acute promyelocytic leukemia associated with translocation (11;17)." Blood 85(4): 1083-94.
- Liu, Q., F. Shalaby, et al. (1994). "Novel zinc-finger proteins that interact with the mouse γ F-crystallin promoter and are expressed in the sclerotome during early somitogenesis." Developmental Biology 165: 165-177.
- LoCoco, F., B. H. Ye, et al. (1994). "Rearrangements of the BCL-6 gene in diffuse large-cell non-Hodgkins lymphoma." Blood 83: 1757-1759.
- Loh, E., J. Elliot, et al. (1989). "Polymerase chain reaction with single-sided specificity: analysis of T-cell receptor delta chain." Science 243: 217-220.
- Lovering, R., I. M. Hanson, et al. (1993). "Identification and preliminary characterisation of a protein motif related to the zinc finger." Proc Natl Acad Sci U S A 90(6): 2112-2116.
- Lozzio, C. G. and B. B. Lozzio (1975). "A human chronic myelogenous leukaemia cell line with positive philadelphia chromosome." Blood 45(321-334).
- Luo, G., C. Hofmann, et al. (1995). "BMP-7 is an inducer of nephrogenesis, and is also required for eye development and skeletal patterning." Genes and Development 9: 2808-2820.
- MacCumber, M. and R. L. Ornstein (1984). "Molecular model for messenger RNA splicing." Science 224: 402-405.
- Macgregor, P. F., C. Abate, et al. (1990). "Direct cloning of leucine zipper proteins- Jun binds cooperatively to the CRE with CRE-BP1." Oncogene 5(4): 451-458.
- Magrath, I. (1990). "Lymphocyte ontogeny: a conceptual basis for understanding neoplasia of the immune system" In *The Non Hodgkin's Lymphoma* (ed Magrath I). Baltimore: Williams and Wilkins: 29-48.
- Meisterernst, M., A. L. Roy, et al. (1991). "Activation of class-11 gene-transcription by regulatory factors is potentiated by a novel activity." Cell 66: 981-993.
- Mellentin, J., C. Murre, et al. (1989). "the gene enhancer binding proteins E12/E47 lies at the t(1;19) breakpoint in acute leukaemia." Science 246: 379-382
- Merika, M. and S. Orkin (1995). "Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Kruppel family proteins Sp1 and EKLF." Molecular and Cellular biology 15(5): 2437-2447.
- Migliazza, A., S. Martinotti, et al. (1995). "Frequent somatic hypermutation of the 5' noncoding region of the BCL6 gene in B-cell lymphoma." Proc Natl Acad Sci USA 92: 12520-12524.
- Minowada, J., G. Janossy, et al. (1978). "Expression of an antigen associated with acute lymphoblastic leukaemia in human leukaemia -lymphoma cell lines." J. Natl Cancer Inst. 60(6): 1269-1273.

References

- Mitani, K., Y. Sato, et al. (1989). "Heterogeneity in the breakpoints of chromosome 19 among acute leukemic patients with the t(11;19)(q23;p13) translocation." *American Journal Hematology* 31: 253-257.
- Mitelman, F. (1994). "Chromosome 19". In: A catalogue of chromosome aberrations in cancer (fourth edition) John Wiley and Sons inc 2: 1399-1451.
- Morris, C., S. Whitman, et al. (1988). "Transposition of the oncogene C-ETS-1 in a T(11-19) (q23-p13) cell-line transient during clonal evolution of blast crisis chronic myeloid-leukemia." *Leukemia* 2(2): 74-78.
- Mu, Z. M., K. V. Chin, et al. (1994). "PML, a growth suppresser disrupted in acute promyelocytic leukemia." *Molecular and Cellular Biology* 14(10): 6858-67.
- Muller, A. M., A. Medvinsky, et al. (1994). "Development of hematopoietic stem cell activity in the mouse embryo." *Immunity* 1: 291.
- Nagy, L., V. A. Thomazy, et al. (1995). "Activation of retinoid X receptors induces apoptosis in HL-60 cell lines." *Molecular and Cellular Biology* 15(7): 3540-51.
- Novina, C. and A. Roy (1996). "Core promoters and transcriptional control." *Trends in genetics* 12(9): 351-355.
- Numoto, M., O. Niwa, et al. (1993). "Transcriptional repressor ZF5 identifies a new conserved domain in zinc finger proteins." *Nucleic Acids Research* 21(16): 3767-3775.
- Onate, S., S. Tsai, et al. (1995). "Sequence and characterisation of a co-activator for the steroid hormone receptor superfamily." *Science* 270: 1354-1357.
- Onizuka, T., M. Moriyama, et al. (1995). "BCL6 gene product, a 92-98-KD nuclear phosphoprotein, is highly expressed in germinal centre B cells and neoplastic counterparts." *Blood* 86: 28-37.
- Onodera, M., T. Kunisada, et al. (1995). "Overexpression of retinoic acid receptor-alpha suppresses myeloid cell-differentiation at the promyelocyte stage." *Oncogene* 11(7): 1291-1298.
- Oyake, T., K. Itoh, et al. (1996). "Bach proteins belong to a novel family of BTB-basic leucine zipper transcription factors that interact with MafK and regulate transcription through the NF-E2 site." *Molecular and Cellular Biology* 16(11): 6083-6095.
- Palacios, R., H. Karasuyama, et al. (1987). "Ly1+ PRO-B lymphocyte clones. Phenotype, growth requirements and differentiation *in vitro* and *in vivo*." *EMBO Journal* 6(12): 3687-3693.
- Pandolfi, P. P., F. Grignani, et al. (1991). "Structure and origin of the acute promyelocytic leukemia myl/RAR α cDNA and characterisation of its retinoic binding and transactivation properties." *Oncogene* 6: 1285-1292.
- Paul, W. and J. Ohara (1987). "B cell stimulatory factor-1/interleukin 4." *Annu Rev Immunol* 5: 429-459.
- Perez, A., P. Kastner, et al. (1993). "PMLRAR homodimers: distinct DNA binding properties and heteromeric interactions with RXR." *EMBO Journal* 12: 3171-3182.
- Ptashne, M. (1988). "How eukaryotic transcriptional activators work." *Nature* 335: 683-689.
- Ptashne, M. and A. A. F. Gann (1990). "Activators and targets." *Nature* 346: 329-331.

References

- Puck, T. T., S. J. Ciciura, et al. (1958). "Genetics of somatic mammalian cells. III. Long term cultivation of euploid cells from human and animal subjects." Journal of Experimental Medicine 108(6): 945-956.
- Rabbitts, T. H. (1994). "Chromosomal translocations in human cancer." Nature 372: 143-149.
- Rabbitts, T. H. and T. Boehm (1991). "Structural and functional chimerism results from chromosomal translocation in lymphoid tumors." Adv Immunology 50: 119-146.
- Raff, J., R. Kellum, et al. (1995). "The *Drosophila* GAGA transcription factor is associated with specific regions of heterochromatin throughout the cell cycle." EMBO J 13: 5977-5983.
- Ralph, P. and I. Nakoinz (1973). "Inhibitory effects of lectins and lymphocyte mitogens on murine lymphomas and myelomas." Journal of the Nat Cancer Institute 31(3): 883-890.
- Ralph, P. and I. Nakoinz (1977). "Direct toxic effects of immunopotentiators on monocytic, myelomonocytic, and histiocytic or macrophage tumor cells in culture." Cancer Research 37: 546-550.
- Redner, R. L., E. A. Rush, et al. (1996). "The t(5;17) variant of acute promyelocytic leukemia expresses a nucleophosmin-retinoic acid receptor fusion." Blood 87(3): 882-6.
- Reid, A., A. Gould, et al. (1995). "Leukemia Translocation Gene, PLZF, Is Expressed With a Speckled Nuclear Pattern In Early Hematopoietic Progenitors." Blood 86(12): 4544-4552.
- Roeder, R. G. (1996). "The role of general initiation factors in transcription by RNA polymerase II." Trends in genetics 12(6): 332-335.
- Rowley, J. D., H. M. Golomb, et al. (1977). "15/17 Translocation: a consistent chromosomal change in acute promyelocytic leukaemia." Lancet 1: 549-550.
- Ryner, L., S. Goodwin, et al. (1996). "Control of male sexual-behaviour and sexual orientation in drosophila by the fruitless gene." Cell 87(6): 1079-1089.
- Sanchez-Garcia, I. and T. H. Rabbitts (1993). "LIM domain proteins in leukemia and development." Cancer Biology 4: 349-358.
- Sanchez-Garcia, I. and T. H. Rabbitts (1994). "The LIM domain: a new structural motif found in zinc-finger-like proteins." Trends in Genetics 10(9): 315-320.
- Sauer, F., J. D. Fondell, et al. (1995). "Control of transcription by Kruppel through interactions with TFIIB and TFIIE." Nature 375: 162-164.
- Sawyers, C. L., C. T. Denny, et al. (1991). "Leukemia and the disruption of normal hematopoiesis." Cell 64: 337-350.
- Schreiber, A. N., L. Chin, et al. (1995). "An amino-terminal domain of Mx1 mediates anti-Myc oncogenic activity and interacts with a homologue of the yeast transcriptional repressor SIN3." Cell 80: 777-786.
- Schuh, R., W. Aicher, et al. (1986). "A conserved family of nuclear proteins containing structural elements of the finger protein encoded by Kruppel, a *Drosophila* segmentation gene." Cell 47: 1025.
- Schulman, I., D. Chakravarti, et al. (1995). "Interactions between the retinoid X receptor and a conserved region of the TATA-binding protein mediate hormone-dependent transactivation." Proc Natl Acad Sci USA 92: 8288-8292.

References

- Schulz, T. C., B. Hopwood, et al. (1995). "An unusual arrangement of 13 zinc fingers in the vertebrate gene Z13." Biochem J. **311**: 219-224.
- Seyfert, V. L., D. Allman, et al. (1996). "Transcriptional repression by the proto-oncogene BCL-6." Oncogene **12**: 2331-2342.
- Shirasuna, K., M. Sato, et al. (1981). "A neoplastic epithelial duct cell line established from an irradiated human salivary gland." Cancer **48**: 745-752.
- Shivdasani, R. A. and S. H. Orkin (1996). "The transcriptional control of haematopoiesis." Blood **87**: 4025-4039.
- Sitterlin, D., P. Tiollais, et al. (1997). "The RAR α -PLZF chimera associated with acute promyelocytic leukaemia has retained a sequence-specific DNA-binding domain." Oncogene **14**: 1067-1074.
- Smith, D. B. and K. S. Johnson (1988). "Single-step purification of polypeptides expressed in Escherichia-Coli as fusion with glutathione S-transferase." Gene **67**(1): 31-40.
- Soel, W., H.-S. Choi, et al. (1996). "An orphan nuclear hormone receptor that lacks a DNA binding domain and heterodimerises with other receptors." Science **272**: 1336-1339.
- Soeller, W. C., C. Eukoh, et al. (1993). "Isolation of cDNAs encoding the *Drosophila* GAGA transcription factor." Molecular and Cellular Biology **13**(12): 7961-7970.
- Sternberg, N. L. (1992). "Cloning high-molecular-weight DNA fragments by the bacteriophage-P1 system." Trends in Genetics **8**(1).
- Stone, R. M. and R. J. Mayer (1990). "The unique aspects of acute promyelocytic leukemia." J Clinical Oncology **8**(11): 1913-21.
- Sugawara, M., T. Scholl, et al. (1994). "A factor that regulates the class II major histocompatibility complex gene DPA is a member of a subfamily of zinc finger proteins that includes a *Drosophila* developmental control protein." Molecular and Cellular Biology **14**(12): 8438-50.
- Tabin, C. (1991). "Retinoids, homeoboxes, and growth factors: toward molecular models for limb development." Cell **66**: 199-217.
- Taylor, B., M. Burmeister, et al. (1996). "Mouse chromosome 10." Mammalian genome **6**(Special): 190-200.
- Titus, J., R. Haugland, et al. (1982). "Texas red, a hydrophilic, red-emitting fluorophore for use with fluorescein in dual parameter flow micro-fluorometric and fluorescence microscopic studies." J Immunol Methods **50**: 193-204.
- Tjian, R. and T. Maniatis (1994). "Transcriptional activation: a complex puzzle with few easy pieces." Cell **77**: 5-8.
- Tsai, S. and S. J. Collins (1993). "A dominant negative retinoic acid receptor blocks neutrophil differentiation at the promyelocyte stage." Proc Natl Acad Sci USA **90**: 7153-7157.
- Tsukiyama, T., P. B. Becker, et al. (1994). "ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor." Nature **367**: 525-532.
- Valtieri, M., D. J. Tweardy, et al. (1987). "Cytokine-dependent granulocytic differentiation." Journal of Immunology **138**(11): 3829-3835.

References

- Varmus, H. E. and C. A. Lowell (1994). "Cancer genes and hematopoiesis." Blood **83**(1): 5-9.
- Verrijzer, C. P. and R. Tjian (1996). "TAFs mediate transcriptional activation and promoter selectivity." Trends in Biological Sciences **21**: 338-342.
- Vinson, C., P. Sigler, et al. (1989). "Scissors-grip model for DNA recognition by a family of leucine zipper proteins." Science **246**(4932): 911-916.
- Wadman, I., L. Jinxing, et al. (1994). "Specific in vivo association between the bHLH and LIM proteins implicated in human T cell leukemia." Embo Journal **13**(20): 4831-4839.
- Wales, M. M., M. A. Biel, et al. (1995). "p53 activates expression of HIC-1, a new candidate tumour suppresser gene on 17p13.3." Nature Medicine **1**(6): 570-576.
- Warner, N. L., M. A. S. Moore, et al. (1969). "A transplantable myelomonocytic leukemia in BALB/c mice: cytology, karyotype and muramidase content." Journal of the National Cancer Institute **43**: 963-982.
- Warrell, R. P., Jr., H. de The, et al. (1993). "Acute promyelocytic leukemia." New England Journal of Medicine **329**(3): 177-89.
- Warrell, R. P., Jr., P. Maslak, et al. (1994). "Treatment of acute promyelocytic leukemia with all-trans retinoic acid: an update of the New York experience." Leukemia **8** (Suppl 2): S33-7.
- Way, M., M. Sanders, et al. (1995). "Sequence and domain organisation of Scruin, an actin cross linking protein in the acrosomal process of Limulus sperm." J Cell Biol **128**: 51-60.
- Weber, U., V. Siegel, et al. (1995). "Pipsqueak encodes a novel nuclear protein required downstream of seven-up for the development of photoreceptors R3 and R4." EMBO Journal **14**: 6247-6257.
- Weima, S., M. van Rooijen, et al. (1988). "Differentially regulated production of platelet derived growth factor and of transforming growth factor beta by a human teratocarcinoma cell line." Differentiation **38**: 203-210.
- Weis, K., S. Rambaud, et al. (1994). "Retinoic acid regulates aberrant nuclear localisation of PML-RAR alpha in acute promyelocytic leukemia cells." Cell **76**(2): 345-56.
- Weiss, A. (1984). "The role of T3 surface molecules in the activation of human T-cells- A 2-stimulus requirement for IL-2 production reflects events occurring at a pre-translational level." J. Immunology **133**(1): 123-128.
- Wells, R. A. and S. Kamel-Reid (1996). "Numa-RARA, a new gene fusion in acute promyelocytic leukaemia." Blood **86**: 365a.
- Wiles, M. V. (1993). "Embryonic stem cell differentiation in vitro." Methods Enzymol **225**: 900-918.
- Wilkinson, D. G. (1992). "Whole mount in situ hybridisation of vertebrate embryos" in: In situ hybridisation a practical approach (ed Wilkinson D.G.) IRL press, Oxford: 75-83.
- Witzgall, R., E. O'Leary, et al. (1994). "The Krüppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression." Proc Natl Acad Sci USA **91**: 4514-4518.
- Wood, K. V. (1991). "Bioluminescence and chemiluminescence: Current status". (eds. P. Stanley and L. Kricka). : 543.

References

- Xiong, W.-C. and C. Montell (1993). "Tramtrack is a transcriptional repressor required for cell fate determination in the *Drosophila* eye." Genes & Development 7: 1085-1096.
- Xue, F. and L. Cooley (1993). "Kelch encodes a component of intercellular bridges in *Drosophila* egg chambers." Cell 72: 681-693.
- Yamamoto, K., M. Seto, et al. (1993). "Two distinct portions of LTG/ENL at 19p13 are involved in t(11;19) leukemia." Oncogene 8: 2617-2625.
- Yancopoulos, G. D., S. V. Desiderio, et al. (1984). "Preferential utilization of the most J_H-proximal V_H gene segments in pre-B-cell lines." Nature 311: 727-733.
- Ye, B. H., S. Chaganti, et al. (1995). "Chromosomal translocations cause deregulated BCL6 expression by promoter substitution in B cell lymphoma." EMBO 14(24): 6209-6217.
- Ye, B. H., F. Lista, et al. (1993). "Alterations of a zinc finger-encoding gene, BCL-6, in diffuse large-cell lymphoma." Science 262: 747-750.
- Yu, V., A. M. Naar, et al. (1992). "Transcriptional regulation by the nuclear receptor superfamily." Current Opinion in Biotechnology 3: 597-602.
- Zervos, A. S., J. Gyuris, et al. (1993). "Mxi1, a protein that specifically interacts with Max to bind Myc-Max recognition sites." Cell 72: 223-232.
- Zollman, S., D. Godt, et al. (1994). "The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in *Drosophila*." Proc Natl Acad Sci USA 91: 10717-10721.
- Zon, L. (1995). "Developmental biology of haematopoiesis." Blood 86(8): 2876-2891.

