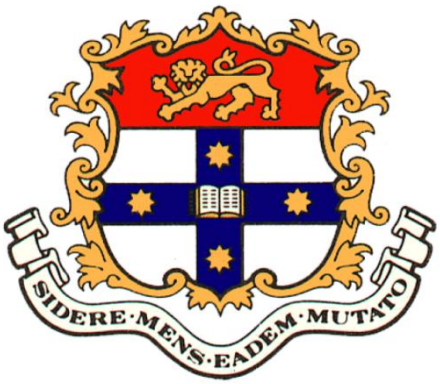
Characterisation of the zinc fingers of Erythroid Krüppel-like Factor

Samantha Hallal

A thesis submitted for the degree of Doctor of Philosophy, University of Sydney



School of Molecular and Microbial Biosciences
University of Sydney
November, 2008

Table of Contents

List of Illustrations

Statement of Originality

The work described in this thesis was conducted between March 2005 and March 2008 in the School of Molecular and Microbial Biosciences at the University of Sydney. I declare that the work presented in the thesis is, to the best of my knowledge and belief, original and my own work, except as acknowledged in the text, and that the material has not been submitted, either in whole or in part, for a degree at this or any other university.

Samantha Hallal

Acknowledgements

Firstly I would like to thank my supervisor Merlin Crossley, for his enlightening advice and persistent guidance throughout my years in the Crossley laboratory. His knowledgeable insights helped make my research an intellectually stimulating and enjoyable experience.

I express my gratitude to all members, past and present, of the Crossley and Nicholas laboratories, who helped nurture my interest in scientific research; Hannah Nicholas, José Perdomo, Alexis Verger, Richard Pearson, Jane van Vliet, Kate Quinlan, Nancy Sue, Alister Kwok, Sally Eaton, Alister Funnell, Stella Lee, Yee Mun Tan, Noelia Nunez, Anna Garsia, Briony Jack, Lit Yeen Tan, Tahni Pyke, Poh Sim Khoo, Tina Wu, Arnout Schepers, Ming Min Lee, Saul Bert and Daniel Scott. In particular, I would like to give a special mention to Kate Quinlan, for her time helping me during my Summer Scholarship in 2002 (my first research experience at the Crossley laboratory) and José Perdomo, for helping me during my Honours year. Special thanks must also be given to Robert Czolij for the considerable effort he has given to ensure the smooth organisation and management of the Crossley laboratory. In addition, I thank Sally Eaton for providing me with the Bklf and Eklf ChIP DNA material and to Alister Funnell, Briony Jack, Noelia Nunez and Richard Pearson for proofreading this thesis.

In addition, I am grateful to all the members of the undergraduate teaching laboratories, particularly Dale Hancock, Jill Johnston, Vanessa Gysbers and Gareth Denyer. I thank them sincerely for giving me the opportunity to teach undergraduate biochemistry and molecular biology because it was an extremely valuable experience for me.

I also thank members of various other laboratories for their collaboration and help. Particularly, I thank Fionna Loughlin (from the Mackay-Matthews laboratory) for her help in learning the SELEX technique and Josep Font (also from the Mackay-Matthews laboratory) for performing the fluorescence anisotropy experiments.

I express my sincerest gratitude and appreciation to my family for their tremendous support and encouragement throughout my years in education and scientific research. If it was not for their unwithering support, I would not have reached this far. I also wish to thank my husband Hassanain for his support and patience during the write up stages of this thesis.

Finally, I acknowledge the financial assistance provided by the University of Sydney Research Office through the Australian Postgraduate Award.

Publications arising from this thesis

Conference Abstract

ComBio 2007, September 2007

Poster: RNA recognition by classical zinc finger proteins

Hallal, S., Loughlin, F., Mackay, J. and Crossley, M.

Abstract

Gene expression is known to be regulated at the level of transcription. Recently, however, there has been a growing realisation of the importance of gene regulation at the post-transcriptional level, namely at the level of pre-mRNA processing (5' capping, splicing and polyadenylation), nuclear export, mRNA localisation and translation.

Erythroid krüppel-like factor (Eklf) is the founding member of the Krüppel-like factor (Klf) family of transcription factors and plays an important role in erythropoiesis. In addition to its nuclear presence, Eklf was recently found to localise to the cytoplasm and this observation prompted us to examine whether this protein has a role as an RNA-binding protein, in addition to its well-characterised DNA-binding function. In this thesis we demonstrate that Eklf displays RNA-binding activity in an *in vitro* and *in vivo* context through the use of its classical zinc finger (ZF) domains. Furthermore, using two independent *in vitro* assays, we show that Eklf has a preference for A and U RNA homoribopolymers. These results represent the first description of RNA-binding by a member of the Klf family.

We developed a dominant negative mutant of Eklf by expressing its ZF region in murine erythroleukaemia (MEL) cells. We used this to investigate the importance of this protein in haematopoietic lineage decisions by examining its effect on the multipotent K562 cell line. We provide evidence that Eklf appears to be critical not only for the promotion of erythropoiesis, but also for the inhibition of megakaryopoiesis.

Abbreviations

Ab	antibody
AD	activation domain
ARE	adenosine-uridine-rich element
Bklf	basic krüppel-like factor
bp	base pairs
CASTing	cyclic amplification and selection of targets
CFU	colony forming unit
CFU-E	colony forming unit – erythroid
CFU-GM	colony forming unit – granulocyte-macrophage (myeloid)
CFU-MK	colony forming unit – megakaryocyte
ChIP	Chromatin immunoprecipitation
CLIP	ultraviolet cross-linking and immunoprecipitation
CML	chronic myeloid leukaemia
CtBP	C-terminal binding protein
CTD	C-terminal domain
DBD	DNA-binding domain
DEPC	diethylpyrocarbonate
DMEM	Dulbecco's Modified Eagle Medium
DMSO	dimethyl sulfoxide
Dnmt3	DNA methyltransferase 3
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
dsRBP-ZFa	double-stranded RNA-binding protein ZFa
E15	embryonic day 15

Eklf	erythroid krüppel-like factor
EMSA	electrophoretic mobility shift assay
ENG	*Drosophila* engrailed protein
EPO	erythropoietin
ES	embryonic stem (cells)
FAK	focal adhesion kinase
FAR	finger associated repeats
FCS	foetal calf serum
Fli-1	friend leukaemia integration 1
FMRP	Fragile X Mental Retardation Protein
GST	glutathione *S*-transferase
GT	guanylyltransferase
HA	haemagglutinin
HAT	histone acetyltransferase
HDAC	histone deacetylase
HeBS	HEPES buffered saline
HID	Sin 3a/Histone deacetylase interaction domain
hnRNPK	heterogeneous nuclear riboprotein K
IE	intermediate element
IPTG	isopropyl-1-thio-β-D-galactopyranoside
IRES	internal ribosome entry site
JAZ	Just Another Zinc finger protein
kDa	kilodalton
KH domain	K-homology type domain
Klf	krüppel-like factor

KSRP	K homology splicing regulatory protein
LB	Luria-Bertani
Lklf	lung krüppel-like factor
MEL	murine erythroleukaemia
MEP	megakaryocyte-erythroid progenitor
MQW	Milli-Q® water
MRE	metal-response element
mRNP	messenger ribonucleic acid (mRNA)-protein complexes
MT	N7G-methyltransferase
MTF-1	MRE-binding transcription factor 1
NP-40	nonidet P-40
NUFIP	nuclear FMRP interacting protein
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
Pprint	**p**rediction of **p**rotein-**R**NA **int**eraction
PSI-BLAST	position-specific iterated-basic local alignment and search tool
PSSM	position-specific scoring matrix
PTB	Polypyrimidine Tract Binding Protein
5S rRNA	5S ribosomal RNA
RBP	RNA-binding protein
RNA	ribonucleic acid
RNP	ribonucleoprotein
rRNA	ribosomal RNA
RT	RNA 5’-triphosphatase

SAAB	selected and amplified binding site imprint assay
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SELEX	systematic evolution of ligands by exponential enrichment
SEM	standard error of the mean
SF1	Mammalian Splicing Factor 1
SUMO	small ubiquitin-related modifieer
SVM	support vector machines
TBP	TATA-binding protein
TEMED	N,N,N',N'-tetramethylethylenediamine
TF	transcription factor
TFIIIA	transcription factor IIIA
TPA	12-*O*-tetradecanoylphorbol-13-acetate
UTR	untranslated region
wig-1	wild-type p53-induced gene 1
WT1	Wilms tumour suppressor protein 1
ZFP100	zinc finger protein 100
ZF	zinc finger
ZNF	zinc finger protein
ZNF74	zinc finger protein 74

All other abbreviations and symbols used are listed in 'Instructions to Authors' (1981), *Biochemical Journal* **193,** 4-27.

Chapter 1 – General Introduction

1.1 Transcriptional regulation of gene expression

All large complex multicellular organisms originate from a single cell containing a set of genomic DNA. Proliferation of this cell and the consequent differentiation into the different cell types, as well as maintenance of the normal function within a cell is brought about by expressing the appropriate sets of genes in a temporally- and spatially-controlled manner. Control of gene expression can be carried out at several stages, the best characterised being at the level of transcription.

1.1.1 *Cis*-acting elements

Cis-acting elements are DNA sequences that influence gene activation or repression through the recruitment of various proteins to the site of transcription. They are categorised as either proximal or distal elements. Promoters that contain proximal elements are typically located up to a few hundred base pairs upstream of a gene. CAAT and CACCC sequences found upstream of the *ß-globin* gene are two examples of proximal promoter elements (Myers et al., 1986). These sequences are responsible for the direct recruitment of *trans*-acting transcription factors, which recognise and bind to these sequence elements. Transcription factors can activate gene expression through recruitment of the basal transcriptional machinery, which consists of RNA polymerase II and its associated general transcription factors TFIIA, -B, -D, -E, -F and –H (Lemon and Tjian, 2000; Muller and Tora, 2004). The general transcription factor TFIID

is composed of TATA-binding protein (TBP) and approximately ten TBP-associated factors (Woychik and Hampsey, 2002). Alternatively, transcription factors repress gene expression through the recruitment of transcriptional co-repressors (Thiel et al., 2004).

The distal elements comprise a diverse class of regulatory sequences, which are located further away from the proximal promoter and the gene. These elements include enhancers, which activate transcriptional activity at the promoter and are thought to do so by looping to a position that is physically close to the promoter (Dean, 2004); silencers, which repress transcriptional activity through the recruitment of one or more repressor proteins or a repression complex (Harju et al., 2000); and insulators, which can block spurious contact between enhancers/silencers and the promoters of other genes in close proximity (West and Fraser, 2002).

1.1.2 *Trans*-acting elements

Gene control at the level of transcription is primarily facilitated by regulatory transcription factors that bind to sequences at promoters, enhancers or silencers to activate or repress expression of their target genes. Frequently these transcription factors possess two distinct domains; a DNA-binding domain that recognises the *cis*-acting element and a regulatory domain, which allows for the recruitment of co-regulator proteins to mediate gene activation or repression (Patikoglou and Burley, 1997). Many different types of DNA-binding domains have been described in the literature, including heliex-turn-helix motifs that are found in homeodomain proteins

(Otting et al., 1990), basic leucine zipper domains found in proteins such as C/EBP (Ellenberger et al., 1992) and zinc fingers, found in proteins such as transcription factor IIIA (TFIIIA) (Lee et al., 1989; Miller et al., 1985) and the related Sp/Krüppel-like factor (Klf) family. The zinc finger proteins will be discussed in greater detail later in the chapter.

The regulatory domain of a transcription factor that activates gene expression is frequently referred to as an activation domain. Some activation domains act through the recruitment of a mediator, which in turn recruits RNA polymerase II and the general transcription factors to form the basal transcriptional machinery (Kornberg, 2005). Alternatively, activation domains function through the recruitment of histone modifying proteins such as Histone Acetyltransferases (HATs). The addition of acetyl groups to histone tails, catalysed by HATs, relieves internucleosomal contacts, thereby resulting in a transcriptionally active chromatin state (Roeder, 2005). Conversely, the regulatory domain of a transcription factor that represses gene expression is often referred to as a repression domain. This domain functions either passively through competition with transcriptional activators for DNA-binding sites (Thiel et al., 2004), or actively through the recruitment of co-repressor proteins such as DNA methyltransferase 3 (Dnmt3) and C-terminal binding protein (CtBP) (Gaston and Jayaraman, 2003). Co-repressors function to repress transcription through inhibition of the assembly of the basal transcriptional machinery or the compaction of chromatin through the recruitment of histone modifying enzymes, such as histone deacetylases (HDACs) (Gaston and Jayaraman, 2003). It must be noted that the expression of most genes is regulated in a

co-ordinated fashion by different types of activators and repressors, and that some transcription factors have the ability to recruit co-activators or co-repressors depending on the promoter or cellular context (Lemon and Tjian, 2000).

1.2 Post-transcriptional regulation of gene expression

The control of gene expression at the level of transcription has been intensively studied. However, there is now a growing appreciation of the importance of gene regulation at the post-transcriptional level (Maniatis and Reed, 2002). Post-transcriptional control of gene expression is mediated by various combinations of RNA-binding proteins (RBPs), which function to control the processing (mRNA capping, alternative splicing and polyadenylation), export, localisation, stability and translation of the mRNA transcripts within a cell (Mata et al., 2005). The structures of proteins that bind RNA are diverse, but can be classified through a few characteristic protein motifs. The most common are the RNA recognition motif (RRM) (Query et al., 1989), the KH domain (Siomi et al., 1993a), the dsRNA binding domain (St Johnston et al., 1992), the arginine-rich motif (Calnan et al., 1991), and the three classes of ZF (Laity et al., 2001). Regulation at the post-transcriptional level adds substantial complexity to gene expression. In addition, a multitude of evidence suggests that the processes involved in post-transcriptional regulation are physically and functionally coupled to each other, as well as to other steps in the gene expression pathway (Maniatis and Reed, 2002). Some of the stages where post-transcriptional gene regulation occurs will be described in this section.

1.2.1 mRNA processing and export

During transcription, nearly all pre-mRNA transcripts are post-transcriptionally modified at their 5' and 3' ends. The pre-mRNA is capped at its 5' end, introns are removed through splicing, and the 3' end is cleaved and polyadenylated (Maniatis and Reed, 2002). 5' capping has important effects on mRNA maturation, translation and stabililty and is the first of the pre-mRNA processing events, as it occurs on RNA polymerase II nascent transcripts that are 20-25 nucleotides in length (Shatkin and Manley, 2000). Phosphorylation of serine 5 of the C-terminal domain (CTD) of RNA polymerase II changes its conformation, allowing it to recruit the three capping enzymes, namely RNA 5'-triphosphatase (RT), guanylyltransferase (GT) and N7G-methyltransferase (MT), which act sequentially to modify the exposed end of the pre-mRNA transcript (Shatkin and Manley, 2000). As is the case with capping, transcription is also coupled to splicing, endonucleolytic cleavage and polyadenylation through interactions between the CTD and the respective processing complexes (Bentley, 1999; McCracken et al., 1997). Splicing requires a set of splicing-related proteins and various components of the splicing machinery to associate with the CTD. For instance, Prp40 associates with the CTD and is thought to function in bringing the 5' and 3' splice sites together during the splicing reaction (Morris and Greenleaf, 2000). As is the case with 5' capping, polyadenylation is also important for the stability of the mRNA. However, unlike capping, the protein machinery required for poly(A) synthesis is much greater in its complexity, requiring more than a dozen polypeptides (not including RNA polymerase II), while capping requires only the proteins mentioned above (Shatkin and Manley, 2000).

Release of the mature mRNA for nuclear export is another process that is tightly coupled to splicing (Luo and Reed, 1999). Evidence for this came from the observation that mRNA transcripts produced by splicing are more efficiently exported out of the nucleus than their unspliced counterparts transcribed from a complementary DNA (Luo and Reed, 1999). This increased export efficiency is thought to be due to the recruitment of the mRNA export factor ALY to the exon-exon boundaries of the transcript during the splicing reaction (Zhou et al., 2000). The ALY protein forms a complex with other proteins, including TAP, a protein that associates with the nuclear pore and is thought to be an mRNA export receptor (Le Hir et al., 2001).

1.2.2 Regulation of mRNA stability

Another level of post-transcriptional control exerted on transcripts is mRNA turnover. The decay rate of a particular mRNA transcript is specified by control elements that are usually found within its 3' untranslated region (UTR) and are recognised by various RBPs (Parker and Song, 2004; Wilusz and Wilusz, 2004). Although the majority of expression-profiling studies focus on transcriptional control, it is actually the mRNA steady-state levels that are measured; these not only reflect the production but also the stability of the transcripts. Thus, techniques were recently developed in order to globally assess mRNA stability and these are revealing important information about this level of regulation. Recently, genome-wide mRNA transcript turnover has been determined in various organisms, including bacteria (Bernstein et al., 2002; Selinger et al., 2003) and yeast (Grigull et al., 2004; Wang et al., 2002) by measuring transcript levels at various times following RNA polymerase II inactivation. Interestingly, decay rates appear to be

precisely controlled, as functionally related genes demonstrate co-ordinated changes in mRNA transcript stability, hence bringing up the notion of decay regulons, a term coined due to the similarity to bacterial operons (Keene, 2007). For example, mRNA transcripts encoding core metabolic proteins were found to have long half-lives, whereas transcripts encoding transcription factors or members of the ribosome-biogenesis machinery were frequently observed to have short half-lives (Yang et al., 2003). Short transcript half-lives enable rapid and dramatic changes in mRNA levels in response to changing conditions and this may be an advantage for transcripts encoding regulatory proteins.

1.2.3 Regulation of translation

An additional level of post-transcriptional gene control takes place during translation and involves both global and transcript-specific mechanisms to regulate protein synthesis (Gebauer and Hentze, 2004). Global regulation of translation, which affects the translation of most transcripts, usually occurs through changes in the phosphorylation state of translation initiation factors and through adjustment of the number of available ribosomes (Gebauer and Hentze, 2004). In contrast, transcript-specific regulation involves the modulation of the translation of a distinct group of mRNA species and is mediated by a number of mechanisms (Gebauer and Hentze, 2004). It involves RBPs that bind to particular secondary structures or regulatory sequences present in the UTRs of target transcripts, and is similar to the control of mRNA decay in that transcripts are modulated through functionally related regulons (described in the previous subsection) (Keene, 2007). The regulation of translation is

particularly important under conditions that require sudden and precise changes in protein levels, such as cellular responses to stress and apoptosis (Holcik and Sonenberg, 2005), regulation of cell growth and its co-ordination with mitosis (Jorgensen and Tyers, 2004), and during differentiation and development (Kuersten and Goodwin, 2003).

1.2.4 mRNA localisation

Messenger RNA transcripts can be localised to the appropriate region of the cell by four basic processes; local synthesis, local protection from degradation, diffusion and local trapping, or active transport along the cytoskeleton (St Johnston, 2005). Local synthesis is the simplest mechanism by which an mRNA can be localised in the cell, although this is rare. For instance, the transcripts for the δ- and ε- subunits of the acetylcholine receptor are transcribed in the nuclei directly below the neuromuscular junctions of mammalian myofibres, but not in the other nuclei of these syncytial cells, thereby concentrating the transcripts in close proximity to the synapses where the receptors are required to function (Brenner et al., 1990). Transcripts can also be localised through degradation of the mRNAs that are not in the correct place. For example, this mechanism has been shown for localisation of the *hsp83* mRNA to the posterior end of the *Drosophila melanogaster* oocyte (Bashirullah et al., 1999). This localisation requires two *cis*-acting elements in the 3'-UTR of the mRNA; a degradation element that targets the transcript for destruction in all regions of the oocyte and a protection element that stabilises the transcript at the posterior end (Bashirullah et al., 1999). Another mechanism through which mRNAs can be localised is by passive diffusion through the cytoplasm until they become entrapped by a localised anchor. One such example of this

mechanism occurs with localisation of the *nanos* mRNA transcript to the posterior end of the *D. melanogaster* oocyte (Forrest and Gavis, 2003). Finally, the best-characterised mechanism of mRNA localisation is through active transport along the cytoskeleton. For instance, the *ASH1* mRNA is localised to the bud tip of *Saccharomyces cerevisiae* through active transport along actin microtubules, which results in the repression of mating type switching in the daughter cell (Takizawa et al., 1997).

1.3 Classical C_2H_2 zinc fingers

The zinc finger is a common structural motif found in transcription factors and is known to function in sequence-specific DNA recognition. Classical zinc fingers (also known as C_2H_2 or CCHH zinc fingers) are a subset of this group and are the most common motif in eukaryotes, accounting for approximately 3% of genes in the human genome (Beerli and Barbas, 2002; Lu et al., 2003; Tupler et al., 2001). This motif of approximately 30 amino acid residues consists of an α-helix at its C-terminal end and a β-hairpin at its N-terminal end, which is referred to as a *ßßα* structure (Figure 1.1) (Mackay and Crossley, 1998) and has a the consensus sequence X_2-C-X_{2-4}-C-X_{12}-H-X_{2-8}-H, where X is any amino acid (Iuchi and Kuldell, 2005). Its secondary structures are stabilised by the tetrahedral co-ordination of a single zinc ion via pairs of cysteine and histidine residues (Figure 1.1) (Matthews and Sunde, 2002).

Figure 1.1 Structure of a classical C_2H_2 zinc finger.

Classical TFIIIA-type zinc finger (PDB code 1ZNF), showing the α-helix and β-hairpin (*ββα* structure) in orange. The cysteine (yellow) and histidine (green) residues co-ordinating the zinc ion (grey sphere) are shown. Adapted from Laity et al, 2001.

In addition to its well-characterised DNA-binding function, the classical zinc finger has also been reported to mediate interactions with RNA, suggesting a possible role in post-transcriptional gene regulation. Furthermore, this domain was shown to interact with proteins (Matthews and Sunde, 2002) and to act as zinc sensors (Bird et al., 2003). Thus, the classical zinc finger motif appears to be an adaptable structure for molecular recognition and for connecting the various steps in the gene expression pathway.

1.3.1 DNA-binding

Classical zinc fingers achieve sequence specificity when binding DNA through variations in key amino acid residues. The basic and hydrophobic side chains on the N-terminal surface in the α-helix of the *ββα* structure interact specifically with two to four bases in the major groove of DNA (Matthews and Sunde, 2002). Binding occurs through primary hydrogen-bond interactions from helical positions (amino acid position relative to the start of the α-helix) -1, 3 and 6 of each zinc finger to one strand of DNA, and through a secondary interaction from helical position 2 to the other strand (Beerli and Barbas, 2002). There is no strict binding code, as residues flanking those at positions -1, 2, 3 and 6 may also make contacts and contribute to specificity. In addition, there are wide variations in this canonical binding arrangement among the known zinc finger-DNA complexes. Many zinc finger proteins contain multiple zinc fingers that can make tandem contacts along the DNA, so that the α-helices of the fingers make sequence-specific contacts along a continuous stretch of the major groove. Contact is made by the first N-terminal zinc finger with the 3' end of the binding site (Pabo et al., 2001). The binding of multiple fingers generally increases the specificity of binding.

Approximately half of the known classical zinc fingers possess the highly conserved short amino acid linker region of sequence TGEKP between adjacent zinc fingers (Laity et al., 2000; Laity et al., 2001; Wolfe et al., 2000). This is thought to increase binding affinity by contacting the C-terminus of the adjacent α-helix upon binding to DNA, thereby stabilising the zinc finger-DNA complex (termed C-capping) (Laity et al.,

2000). Furthermore, the linker is thought to regulate DNA-binding, as phosphorylation of the threonine residue inhibits binding during mitosis (Dovat et al., 2002).

In a number of transcriptional regulators, tandem arrays of three or more closely spaced classical zinc fingers have been shown to mediate sequence-specific DNA-binding (Pavletich and Pabo, 1991), and these proteins have generally been shown to act as conventional transcription factors. Thus, the function of the zinc finger region in such proteins is to localise the regulator to its specific target genes. The remaining domains within the protein then act to regulate gene expression, typically through the recruitment of chromatin-modifying enzymes or the basal transcriptional machinery.

DNA-binding of the classical zinc finger has been characterised to the extent that it is possible to predict DNA-binding site of the zinc finger on the basis of its amino acid sequence composition. Alternatively, it is possible to design and construct zinc fingers with altered specificity by altering key amino acid residues. Artificial transcription factors that target a particular sequence of interest have been designed and this is referred to as zinc finger engineering (Choo and Isalan, 2000).

1.3.2 RNA-binding

Despite DNA-binding by classical zinc fingers being a well-understood interaction, many aspects of the molecular basis of RNA-binding remain elusive. In contrast to the relatively uniform helical structure of DNA, RNA exhibits more complex secondary and tertiary structures (Lu et al., 2003) and presents distinct challenges for protein recognition. Segments of intramolecularly formed double helix that exist are scattered between complex secondary structures, namely bulges and hairpin loops (Draper, 1995). In addition, the A-form conformation adopted by duplex RNA (as opposed to the B-form, which is the predominant conformation of DNA in solution) affects protein recognition (Draper, 1995). Although both A- and B-form helices are right-handed, the pitch, or the distance required to complete one helical turn, differs between the two forms. The B-form requires 10-10.6 bp or 3.4 nm to complete one turn, whereas the A-form requires 11 bp or 2.46 nm (Watson, 1983). Consequently, the A-form helix adopted by RNA has a relatively narrow and deep major groove (Figure 1.2), thereby resulting in steric restriction and limiting accessibility to proteins (Cheng et al., 2001).

Over a dozen classical zinc finger proteins are documented in the literature to bind RNA (Table 1.1), which is small in comparison to the number shown to bind DNA. Of these zinc finger proteins, the most extensively studied of these is transcription factor IIIA (TFIIIA) from *Xenopus laevis* and this will be described in greater detail in the following subsection.

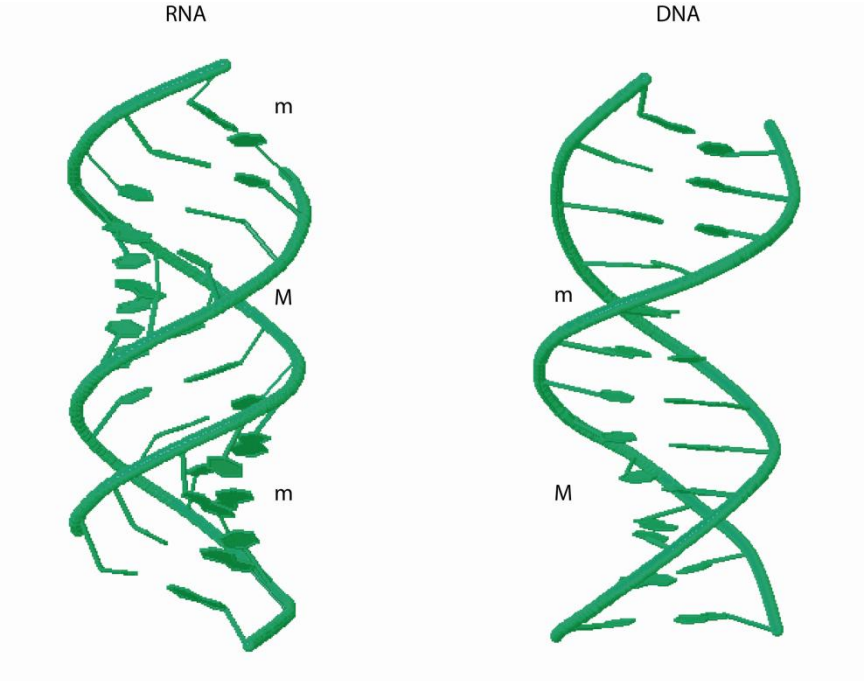


Figure 1.2 Comparison of RNA and DNA structure.

Duplex forms of RNA (PDB code 1RNA) and DNA (PDB code 1BNA) are shown. The major groove (M) of the RNA duplex is deeper and narrower than that of DNA and restricts protein accessibility. However, despite the minor groove (m) being more accessible, hydrogen bond contacts may limit sequence recognition of AU or GC base-pairs.

Table 1.1 RNA-binding classical C_2H_2 zinc finger proteins.

Listed are all the known classical zinc finger proteins reported to interact with RNA.

Protein	No. of zinc fingers	Comments	Reference
TFIIIA	9	binds 5S RNA (7S RNP) and 5S RNA gene; transcription factor	Romaniuk, 1985
p43	9	binds 5S RNA isolated from 42S RNP in *Xenopus* oocytes	Joho et al, 1990
WT1	4	+KTS form associated with splicing factors; transcription factor	Caricasole et al, 1996
Wig-1	3	p53 induced; localised to nucleoli	Méndez-Vidal et al, 2002
dsRBP-Zfa	7	binds dsRNA through 3 N-terminal fingers or 3 C-terminal zinc fingers	Finerty and Bass, 1999
JAZ	4	long linkers between zinc fingers; binds dsRNA or DNA/RNA hybrids	Yang et al, 1999
MOK2	2	long linkers between zinc fingers; binds RNA and DNA (brain and testis)	Arranz et al, 1997
NUFIP	1	associates with FMR RNA-binding protein	Bardoni et al, 1999
XFO-5, XFG-6	5-15	members of FAR family proteins	Klocke et al, 1994
ZNF74	12	deleted in DiGeorge syndrome; preferentially binds poly(U)	Grondin et al, 1996
PEP	4	co-precipitates in a complex including hnRNPK	Amero et al, 1993
ZFP100	18	binds to SLBP/SL RNA complex using fingers 2-8	Dominski et al, 2002
Xfin	31	preferentially binds poly(G)	Andreazzoli et al, 1993

1.3.2.1 TFIIIA

Of the dual DNA- and RNA-binding zinc finger transcription factors, TFIIIA is the most well-characterised and thus can be used to exemplify the role of DNA-binding zinc fingers in post-transcriptional gene regulation (Lu et al., 2003). TFIIIA consists of nine tandem classical zinc fingers (Figure 1.3A) and is an essential RNA polymerase III transcription factor for the activation of the 5S rRNA gene in *Xenopus* oocytes. In addition, TFIIIA binds directly to the 5S rRNA gene product, resulting in the formation of a 7S ribonucleoprotein (7S RNP) storage particle that stabilises the RNA until it is required for ribosome assembly (Searles et al., 2000). Furthermore, TFIIIA facilitates the export of the 5S rRNA molecule out of the nucleus (Hall, 2005).

Structural studies have revealed that TFIIIA has a DNA-binding site extending over 55 bp of the *5S rRNA* promoter (Figure 1.3B). More specifically, zinc fingers 1-3, 5 and 7-9 of TFIIIA bind to three sequence elements within the *5S rRNA* promoter, namely a 10 bp 'Box C' sequence, a 3 bp 'intermediate element' (IE) and an 11 bp 'Box A' sequence, respectively (Nolte et al., 1998) (Figure 1.3B). Fingers 4 and 6 were found not to interact with DNA and are thought to function primarily as spacer elements with respect to DNA-binding. This brought up the possibility that these zinc fingers may play an important role in RNA recognition and indeed this was found to be the case.

The recently obtained crystal structure of zinc fingers 4-6 of TFIIIA bound to a 61-nucleotide fragment of 5S rRNA demonstrates that the RNA-binding interactions of TFIIIA differ markedly from its DNA-binding interactions (Lu et al., 2003). This can be expected, due to the substantial structural differences that exist between DNA and RNA, as described earlier in this section. The zinc fingers of TFIIIA must overcome the challenges presented by the structural complexity of RNA in order to form a successful molecular interaction. In the TFIIIA-RNA structure, each zinc finger of fingers 4-6 was found to interact with a distinct element of the RNA molecule (Figure 1.3C). More specifically, zinc fingers 4 and 6 were found to recognise loops E and A, respectively, through the contact of residues that protrude from the RNA structure; while finger 4 contacts a single protruding guanosine residue, finger 6 interacts with an adenosine and a cytosine residue. The specificity of these interactions was achieved through side chain contacts from α-helices at positions -1, 1 and 2. In addition, a stacking interaction between a tryptophan residue present in finger 6 and a protruding adenosine residue

was also found to be important for TFIIIA RNA-binding. Zinc finger 5, which is a bi-functional DNA- and RNA-binding domain, demonstrates a different mode of RNA recognition. This finger interacts with a short double helix in 5S rRNA through multiple contacts between the basic amino acids in its α-helix and the sugar-phosphate backbone of the RNA molecule. These interactions have been confirmed in various independent mutagenesis studies. One example of such a study involved alanine substitutions in the RNA-binding residues of zinc fingers 4 and 6 of TFIIIA, which resulted in a substantial reduction in its RNA-binding affinity (Friesen and Darby, 1997). Such mutagenesis studies highlight the importance of these residues in the tight nanomolar affinity binding of 5S rRNA.

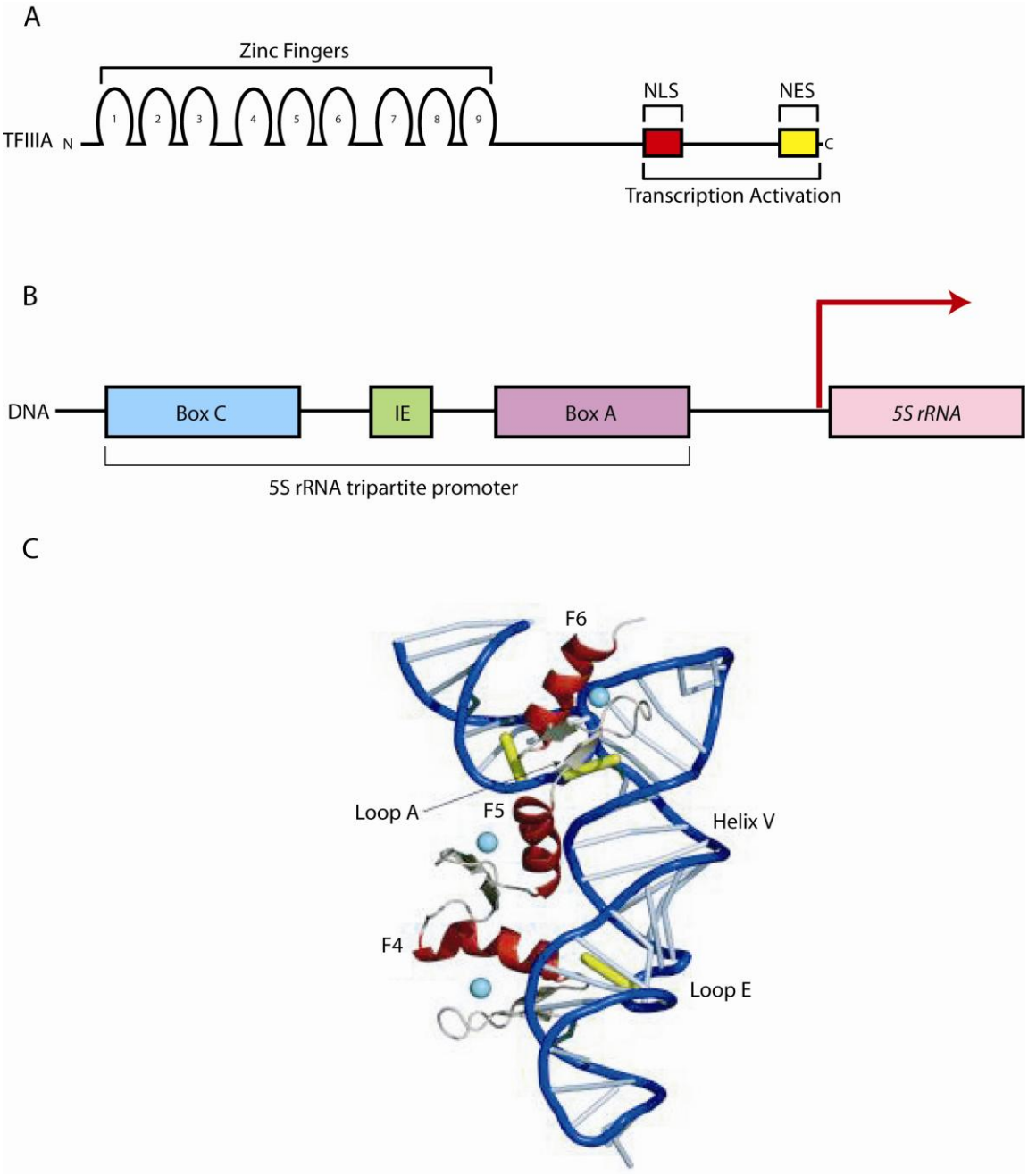


Figure 1.3 Dual DNA- and RNA-binding functions of the C_2H_2 zinc finger protein TFIIIA.

A. Schematic representation of the classical zinc finger protein TFIIIA. Relevant sequence and structural features of the protein are shown. The nine zinc fingers are numbered and their arrangement within the protein is indicated. The areas critical for nuclear import (NLS; in red) and export (NES; in yellow) are shown. Transcriptional activation is mediated by the C-terminal region of the protein, as indicated by the bracket. **B.** The DNA target of TFIIIA, the *5S rRNA* promoter. Upregulation occurs through binding of zinc

fingers 1-3, 5 and 7-9 to Box C (cyan), intermediate elements (IE; green), and Box A (purple) sequences, respectively, by wrapping around the major groove in the canonical binding arrangement of DNA-binding transcription factors. **C.** Three-dimensional structure of zinc fingers 4-6 of TFIIIA bound to 61-nucleotides of the 5S rRNA molecule (reproduced from Stefl et al, 2005). Finger 4 binds to loop E, finger 5 interacts with helix V and finger 6 binds to loop A. The protruding bases recognised by fingers 4 and 6 are shown in yellow. The zinc ions stabilising the zinc finger structures are represented as cyan spheres.

1.3.3 Protein-protein interactions

Although the majority of classical zinc fingers are known to bind nucleic acids, a number of these proteins containing these motifs have been identified to function by mediating protein-protein interactions using their zinc fingers. For instance, Roaz, GL1, SW15, Ikaros, TRPS-1 and Zac form homodimers using their classical zinc fingers (Mackay and Crossley, 1998; McCarty et al., 2003; Sun et al., 1996; Tsai and Reed, 1998; Wang et al., 2001). GL1 forms a homodimer through the hydrophobic surface of its first zinc finger that is not involved in DNA-binding. Similarly, SW15 forms homodimers using the hydrophobic surface of both the β hairpin and the C-terminal end of the α-helix of its first zinc finger. The Ikaros family of transcription factors, which play an important role in lymphoid differentiation (Georgopoulos, 2002), contain two clusters of zinc fingers. The N-terminal cluster consists of four classical zinc fingers and mediates DNA-binding activity. The C-terminal cluster consists of two zinc fingers and mediates protein dimerisation through zinc finger association. The latter cluster of zinc fingers can mediate either homodimerisation or heterodimerisation with other family members, which can result in different patterns of DNA-binding and transcriptional activity (Morgan et al., 1997). Mutagenesis studies of the two C-terminal zinc fingers of

Ikaros revealed that the amino acid residues on the α-helix are responsible for the finger-finger interactions necessary for Ikaros homodimerisation (McCarty et al., 2003).

1.3.4 Zinc sensing

Another recently discovered role for classical zinc fingers is the regulation of gene expression in response to fluctuations in cellular zinc concentrations. The response of mammalian cells to excess zinc is mediated by metallothionein, which chelates and sequesters surplus metal ions from the cellular environment. Metallothionein expression is regulated through the metal-response element (MRE) by the MRE-binding transcription factor 1 (MTF-1), which contains six classical zinc fingers (Bittel et al., 2000; Chen et al., 1999). In yeast, the cellular response to changes in zinc concentration is mediated by the transcriptional activator Zap1, which contains seven classical zinc fingers (Bird et al., 2003). Both proteins contain a core DNA-binding domain, consisting of a subset of the zinc fingers. In Zap1, the five C-terminal zinc fingers are involved in DNA-binding, while the first two zinc fingers appear to function as unique metal-sensing domains that control the activation domain of this protein in response to zinc levels (Bird et al., 2003). On the other hand, fingers two to four of MTF-1 bind DNA, while finger one appears to function as a metal sensor, which prevents MTF-1 from binding to the MRE in the absence of zinc (Bittel et al., 2000). In both cases it appears that zinc levels affect the structure of the zinc-sensing domain, which in turn either hinders or exposes the activation domain, thereby controlling interaction with downstream transcription factors.

1.4 Sp/Krüppel-Like Factor (Klf) family of transcription factors

Sp1 is the founding member of the Sp/Krüppel-like factor (Klf) family, and was one of the first transcription factors to be characterised in mammalian cells. It preferentially binds GC-rich DNA sequences using its three classical zinc fingers and is involved in the activation of gene transcription (Kaczynski et al., 2003). The DNA-binding region of Sp1 is related to a region of the *Drosophila melanogaster* segmentation gene *Krüppel*, and is present in many other developmental regulators. This DNA-binding domain defines a class of nine Sp proteins and seventeen Krüppel-like factors (Klfs); namely the Sp proteins, Sp1 to Sp9, and Klf family members, Klf1 to Klf17 (Kaczynski et al., 2003; van Vliet et al., 2006) (Figure 1.4). Beyond the DNA-binding zinc finger region, members of this family share little homology and also generally have differing expression profiles. This diversity allows them to carry out distinct physiological activities using their varied N-terminal activation or repression domains (Lomberk and Urrutia, 2005). Many are ubiquitously expressed, such as Sp1 to Sp4 (Lania et al., 1997), as well as Klf7 (also known as ubiquitous krüppel-like factor, or Uklf) (Matsumoto et al., 1998) and Klf3 (also known as basic krüppel-like factor, or Bklf) (Crossley et al., 1996). However, other members of this family have more-restricted expression profiles and are typically named according to the primary location of their expression. For instance, Klf1 (also known as Erythroid krüppel-like factor, or Eklf) (Miller and Bieker, 1993), Klf2 (also known as Lung krüppel-like factor, or Lklf) (Anderson et al., 1995), and Klf4 (also known as Gut krüppel-like factor, or Gklf) (Shields et al., 1996) are predominantly expressed in the tissues indicated by their names. Members of the Klf family of transcription factors can function as either gene activators or repressors, and

interestingly some members have the dual capacity to function as either an activator or a repressor, depending on the promoter and cellular context (Kaczynski et al., 2003). In addition, some Klf family members can regulate their own expression or the expression of other family members. For example, Eklf was found to directly activate Bklf expression in murine erythroid cells (Funnell, 2008; Funnell et al., 2007), while Klf4 autoactivates itself and is thought to be repressed by Klf5 (Dang et al., 2002). Sp/Klf transcription factors have been found to play important roles in a range of biological processes (Bieker, 2001; Dang et al., 2000), including erythropoiesis (Basu et al., 2005; Matsumoto et al., 2006; Nuez et al., 1995; Perkins et al., 1995; Van Loo et al., 2003; Wani et al., 1998), adipogenesis (Banerjee et al., 2003; Mori et al., 2005; Oishi et al., 2005) and carcinogenesis (Black et al., 2001; Ghaleb et al., 2005; Narla et al., 2001; Wang and Zhao, 2007).

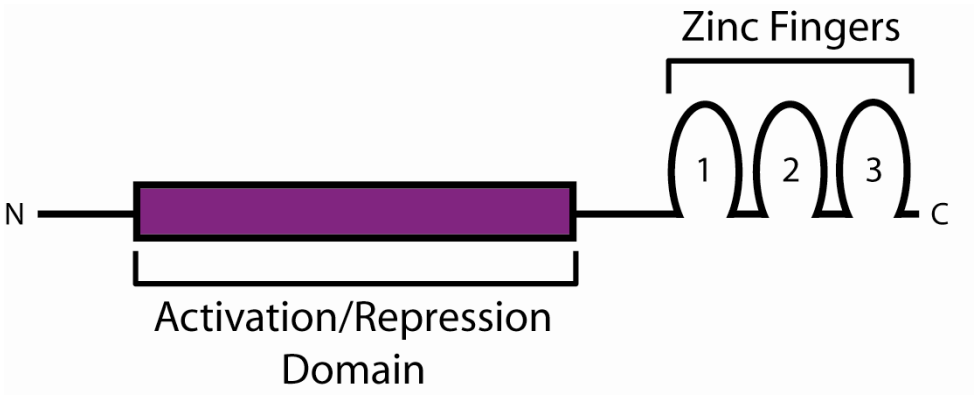


Figure 1.4 Schematic representation of the Sp/Klf family of transcription factors.

Depicted are the three C-terminal classical DNA-binding zinc fingers and the N-terminal domain that mediates either activation or repression activity (shown in purple).

1.4.1 Eklf/Klf1

Eklf is the founding member of the Klf family and was originally isolated in a subtractive cloning approach used to identify genes that are important for erythroid differentiation (Miller and Bieker, 1993). It is the best characterised member of the family (Perkins, 1999) and is a potent transcriptional activator so named because its expression is restricted mainly to erythroid cells, although expression has also been detected in mast cells (Luo et al., 2004; Miller and Bieker, 1993). This protein recognises CACCC boxes and related DNA sequence elements of the general form 5'-NCNCNCCCN-3' (Feng et al., 1994). These sequence motifs are found abundantly in the promoter regions of many erythroid genes (Feng et al., 1994; Miller and Bieker, 1993; Raich and Romeo, 1993). A particularly well-studied target of activation is the adult *ß-globin* promoter, which contains an Eklf recognition site that is critical for β-globin expression (Donze et al., 1995; Feng et al., 1994). Consequently, *Eklf* knockout mice die at approximately embryonic day 15 (E15) from severe anaemia, which reflects a β-globin deficiency (Nuez et al., 1995; Perkins et al., 1995). Eklf was found to activate gene expression through the recruitment of the transcriptional co-activators p300 and CBP (Zhang and Bieker, 1998), which acetylate both histones and Eklf itself. In addition, acetylation of Eklf is thought to interfere with its ability to interact with components of the SWI/SNF ATP-dependent chromatin remodelling complex, thereby allowing the chromatin to remain in a transcriptionally active state (Armstrong et al., 1998). Recent evidence shows that in addition to its nuclear localisation, Eklf was found to localise to the cytoplasm (Quadrini et al., 2008; Shyu et al., 2007), suggesting this protein may have an additional function that is yet to be determined.

A recent study has established that, in addition to its important role in driving erythroid differentiation, Eklf has an inhibitory effect on megakaryocyte differentiation (Frontelo et al., 2007). This finding suggests that Eklf plays a key role in haematopoietic bipotential lineage decisions in the megakaryocyte-erythroid progenitor (MEP). Thus, this unique role for Eklf distinguishes it from other transcription factors that are common to megakaryocytes and erythroid cells that play positive roles in both lineages, such as GATA1 (Pevny et al., 1991; Shivdasani et al., 1997), FOG1 (Tsang et al., 1997), SCL (Hall et al., 2003; Mikkola et al., 2003) and Gfi-1b (Saleque et al., 2002). The molecular mechanism for these lineage decisions is thought to occur because of the functional cross-antagonism that exists between Eklf and friend leukaemia integration 1 (fli-1), an Ets transcription factor critical for megakaryocyte differentiation (Starck et al., 2003).

1.4.2 Lklf/Klf2

Lklf is a transcription factor that was originally named by virtue of its high level of expression in the lung (Anderson et al., 1995), however, relatively little is known about this protein. *Lklf* null mice were found to have defects in blood vessel organisation, due to a deficiency in the recruitment of pericytes and smooth muscle cells, and died at an early stage of embryonic development from severe haemorrhaging (Kuo et al., 1997a). In order to circumvent the embryonic lethality, chimeric mice were generated through the injection of *Lklf*-deficient embryonic stem (ES) cells into recombinase-deficient blastocysts (Kuo et al., 1997b). Analysis of these chimeric animals revealed that the mature, single-positive T cells were susceptible to apoptosis, suggesting a role for Lklf in

quiescent T cells. In addition, histopathological analysis of the lungs of highly chimeric mice that died at birth revealed an abnormal pathology and demonstrated a deficiency in the late stages of development (Wani et al., 1999), suggesting an important role for Lklf in the development of the cells of the lung.

1.4.3 Bklf/Klf3

As mentioned earlier, Bklf is a ubiquitously expressed protein and is especially abundant in the yolk sac and foetal liver of developing mice (Crossley et al., 1996), suggesting a putative role in haematopoiesis. This protein has a high frequency of basic amino acid residues, hence the name basic krüppel-like factor. Bklf functions as a transcriptional repressor (Perdomo et al., 2005; Turner and Crossley, 1998; Turner et al., 2003) and within its repression domain resides a five amino acid PXDLS-like motif (PVDLT in Bklf), which is utilised to interact with the co-repressor C-terminal binding protein (CtBP), allowing recruitment of the repression protein complex. Bklf was originally isolated in a screen using a CACCC probe specific for the zinc finger domains of Eklf and Sp1 and a relaxed-stringency hybridisation against a murine erythroleukaemia (MEL) cDNA library (Crossley et al., 1996). *In vitro*, Bklf has been shown to bind to the CACCC-boxes within the promoters of various genes, including adult *ß-globin*, foetal ^{A}*γ-globin*, *Gata-1*, *carbonic anhydrase I*, *porphobilinogen deaminase* (*Pbgd*) and *pyruvate kinase* (Crossley et al., 1996; Perdomo et al., 2005; Turner and Crossley, 1998; Turner et al., 2003). A straightforward physiological role for this protein has been difficult to establish, and *Bklf* knockout mice display a mild phenotype, such as reduced size and weight (Sue et al, unpublished data). However,

recent work in our laboratory has identified the *Klf8* gene as a direct biological target of Bklf (see next subsection).

1.4.4 Klf8

The human Klf8 protein (van Vliet et al., 2000) shares a considerable degree of homology with Bklf. Similar to Bklf, Klf8 has a PXDLS-like motif (PVDLS in Klf8), which is required for its interaction with CtBP. In addition, Klf8 also binds to CACCC sequences in EMSA experiments (van Vliet et al., 2000). In addition to its repressor function, Klf8 has also been reported to function as a transcriptional activator and has been implicated in cell cycle control (Zhao et al., 2003). It is an oncogene that is upregulated in various human cancer cell lines and tumours (Wang and Zhao, 2007) and is a downstream target of focal adhesion kinase (FAK), an important mediator of various signalling pathways, including cell adhesion, cell survival and proliferation (Zhao et al., 2003). In addition, transient assays suggest that Klf8 activates the cyclin D1 promoter. A mutant lacking the PXDLS-like motif could not activate cyclin D1 expression and a truncation mutant lacking the three DNA-binding zinc fingers also did not activate expression (Zhao et al., 2003). Furthermore, chromatin immunoprecipitation (ChIP) assays demonstrated that Klf8 binds directly to the cyclin D1 promoter *in vivo* (Wei et al., 2006).

Recent work in our laboratory demonstrates that Klf8 is directly repressed by Bklf. Microarray analysis showed that *Klf8* expression is upregulated in *Bklf* knockout mice,

and ChIP assays confirm that Bklf directly associates with the *Klf8* promoter (Eaton et al, unpublished data). Our current working model is that Bklf represses *Klf8* expression and Klf8 compensates for the loss of Bklf in the *Bklf* knockout mice, as the functions of these two proteins may overlap due to their similarity. This provides a plausible explanation for the mild phenotype observed and for the inability to identify alternative Bklf target genes using microarray analysis.

1.5 Project aims

The observation that Eklf localises to the cytoplasm (Quadrini et al., 2008; Shyu et al., 2007) suggests that Eklf has a novel role that is yet to be identified. This observation led us to hypothesise that in addition to its transcription factor function in the nucleus, Eklf may function to regulate gene expression at the post-transcriptional level when present in the cytoplasm. In other words, it is hypothesised that Eklf is an RNA-binding protein. In this thesis the RNA-binding activity of the classical zinc fingers of Eklf, through the use of various *in vitro* assays, is examined. In addition, the activity of the zinc finger region of Eklf is examined in a cellular context. Thus, the recombinant expression of the zinc fingers of Eklf in two independent mammalian cell lines is described and evidence is presented that this region behaves as a dominant negative mutant.

Chapter 2 – Materials and Methods

2.1 Materials

2.1.1 Chemicals and reagents

Below is a list of important chemicals and reagents used, along with their suppliers indicated in italics. All chemicals and reagents used were of molecular biology grade unless specified otherwise.

- Acetic acid *Asia Pacific Specialty (APS) Chemicals, Seven Hills, NSW, Australia*
- Acrylamide, 30% (37.5:1 Protogel) (electrophoresis grade) *National Diagnostics, Atlanta, Georgia, USA*
- Acrylamide, 40% (19:1) (electrophoresis grade) *National Diagnostics*
- Adenosine triphosphate (ATP) *Sigma-Aldrich Company*
- Adenosine 5'-[γ-^{32}P] triphosphate (10 mCi/mL) *Geneworks, Adelaide, SA, Australia*
- Agarose (DNA grade) *Progen Industries, Darra, QLD, Australia*
- Albumin, bovine serum, fraction V powder (BSA) *Sigma-Aldrich Company*
- Ammonium acetate *Univar, APS Finechem, Seven Hills, NSW, Australia*
- Ammonium persulfate (APS) *Sigma-Aldrich Company*
- Ampicillin sodium salt *Progen Industries*

- Aprotinin *Sigma-Aldrich Company*
- Boric acid *APS Chemicals*
- 3',3'',5',5''-tetrabromophenolsulfonepthalein (bromophenol blue) *Sigma-Aldrich Company*
- Chloroform *Biolab Scientific, Clayton, VIC, Australia*
- Coomassie® Brilliant Blue-R *Sigma-Aldrich Company*
- Complete, EDTA-free protease inhibitor tablets *Roche Molecular Biochemicals, Mannheim, Germany*
- Complete Mini, EDTA-free protease inhibitor tablets *Roche Molecular Biochemicals*
- Deoxynucleotide triphosphates (dNTPs) *Roche Molecular Biochemicals*
- Diethylpyrocarbonate (DEPC) *Sigma-Aldrich Company*
- Dimethyl sulfoxide (DMSO) *Sigma-Aldrich Company*
- Dithiothreitol (DTT) *Sigma-Aldrich Company*
- Dulbecco's Modified Eagle Medium (DMEM) *Gibco-BRL Life Technologies*
- Ethidium bromide *Amresco Inc.,Solon, Ohio, USA*
- Ethylenediaminetetraacetic acid disodium dehydrate (EDTA) *Ajax Laboratory Chemicals, Auburn, NSW, Australia*

- Ficoll®-400 *Sigma-Aldrich Company*
- Foetal bovine serum, heat-inactivated *Gibco-BRL Life Technologies*
- Formaldehyde *Sigma-Aldrich Company*
- Formamide *Sigma-Aldrich Company*
- GeneRuler™ DNA ladder mix *Fermentas, Ontario, Canada*
- Glutathione-agarose beads *Sigma-Aldrich Company*
- Glutathione, reduced form *Sigma-Aldrich Company*
- Glycerol *Ajax Laboratory Chemicals*
- Glycogen *Roche Molecular Biochemicals*
- 4-(2-Hydroxyethyl)-1-piperazin-ethan-sulfonsäure (HEPES) *Sigma-Aldrich Company*
- IGEPAL CA-630 *Sigma-Aldrich Company*
- Isopropanol *Biolab Scientific, Northcote, New Zealand*
- Isopropyl-1-thio-β-D-galactopyranoside (IPTG) *Sigma-Aldrich Company*
- Leupeptin *Sigma-Aldrich Company*
- Magnesium chloride *Sigma-Aldrich Company*
- Mark-12™ protein standards *Invitrogen*
- β-mercaptoethanol *Sigma-Aldrich Company*

- 3-(N-Morpholino)-propanesulfonic acid (MOPS) *Sigma-Aldrich Company*
- Penicillin, streptomycin and glutamine solution (1%) *Gibco-BRL Life Technologies*
- Phenol:choloroform:isoamyl alcohol (25:24:1) *Sigma-Aldrich Company*
- Phenylmethylsulfonylfluoride (PMSF) *Sigma-Aldrich Company*
- Phosphate buffered saline (PBS) tablets *Sigma-Aldrich Company*
- Poly (dI·dC) *GE Healthcare Life Sciences Biotech, Little Chalfont, Buckinghamshire, UK*
- Polyoxyethylenesorbitanmonolaurate (Tween™-20) *Sigma-Aldrich Company*
- Potassium chloride *Sigma-Aldrich Company*
- Quick Spin® G-25 columns *Roche Molecular Biochemicals*
- Rainbow™ protein standards *GE Healthcare Life Sciences*
- RNasin® recombinant ribonuclease (RNase) inhibitor *Promega, Madison, WI, USA*
- Roswell Park Memorial Institute (RPMI) 1640 medium *Gibco-BRL Life Technologies*
- Sodium acetate *Ajax Laboratory Chemicals*
- Sodium borate *Ajax Laboratory Chemicals*

- Sodium chloride *Ajax Laboratory Chemicals*
- Sodium dodecyl sulphate (lauryl sulphate sodium salt) (SDS) *Sigma-Aldrich Company*
- Sodium hydroxide *APS Chemicals*
- Spin-X® centrifuge tube filters *Trace Biosciences, Castle Hill, NSW, Australia*
- N,N,N',N'-tetramethylethylenediamine (TEMED) (electrophoresis grade) *Eastern Organic Company, Rochester, NY, USA*
- Tris-hydroxymethyl-methylamine (Tris) *Ajax Laboratory Chemicals*
- t-octylphenoxypolyethoxyethanol (Triton® X-100) *Sigma-Aldrich Company*
- Vanadyl ribonucleoside complexes *New England Biolabs, Beverly, MA, USA*
- Xylene cyanole FF *Sigma-Aldrich Company*
- Zinc sulfate *Ajax Laboratory Chemicals*

2.1.2 Enzymes

- Alkaline phosphatase (from calf intestine) [phosphate-monoester phosphohydrolase (alkaline optimum), EC 3.1.3.1] *Roche Molecular Biochemicals*
- Lambda protein phosphatase (λ-PPase, EC 3.1.3.16) *New England Biolabs*

- *Pfu* polymerase [deoxynucleotide-triphosphate:DNA deoxynucleotidyltransferase (DNA-directed), EC 2.7.7.7] *Stratagene, La Jolla, CA, USA*

- Ribonuclease A (RNase, EC 3.1.27.1) *Roche Molecular Biochemicals*

- T4 DNA ligase (poly[deoxyribonucleotide]:poly[deoxyribonucleotide] ligase, EC 6.5.1.1) *Roche Molecular Biochemicals*

- T4 polynucleotide kinase (ATP: 5'-dephosphopolynucleotide 5'-phosphotransferase, EC 2.7.1.78) *New England Biolabs*

- Type II restriction endonucleases (EC 3.1.21) *New England Biolabs*

2.1.3 Bacterial strains and culture media

The bacterial strain used for all plasmid manipulations, including sub-cloning and plasmid isolation was *Escherichia coli* DH5α (genotype: *supE*44 Δ*lac*U169 [ϕ80*lac*ZΔM15] *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1) (*Bethesda Research Laboratories, Gaithersburg, MD, USA*).

The bacterial strain used for expression of GST fusion proteins was Epicurian coli® BL21 [*Escherichia coli*, genotype: B F-*dcm ompT hsdS* (r_{B-} m_{B-}) *gal*] (*Stratagene*).

Both bacterial strains were cultured in Luria-Bertani (LB) broth, or on LB agar plates:

10 g/L casein peptone *Amyl Media, Dandenong, VIC, Australia*

5 g/L yeast extract *Amyl Media*

10 g/L sodium chloride *Ajax Laboratory Chemicals*

15 g/L bacteriological agar *Amyl Media* (for plates only)

All media were made up in Milli-Q® water (MQW) and sterilised by autoclaving for 15 min at 121°C. Filter-sterilised ampicillin (50 mg/mL in MQW) was added to cooled autoclaved broth to a final concentration of 100 µg/mL. In the preparation of agar plates, this step was performed prior to pouring. All media was stored at 4°C.

2.1.4 Mammalian cell lines and culture media

Murine erythroleukaemia (MEL) cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) (*Gibco BRL Life Technologies*) that was supplemented with 10% (w/v) foetal bovine serum (heat-inactivated) (*Gibco BRL Life Technologies*) and 1% penicillin, streptomycin, glutamine (*Gibco BRL Life Technologies*). Cells were incubated at 37°C with 5% CO_2 in a CO_2 water-jacketed incubator.

Human K562 cells were maintained in Roswell Park Memorial Institute (RPMI) 1640 medium that was supplemented with 10% (w/v) foetal bovine serum (heat-inactivated)

(*Gibco BRL Life Technologies*) and 1% penicillin, streptomycin, glutamine (*Gibco BRL Life Technologies*). Cells were incubated at 37°C with 5% CO_2 in a CO_2 water-jacketed incubator.

2.1.5 Antibodies

- Anti-Bklf (α-Bklf) polyclonal antibody raised in rabbit was supplied by M. Crossley.
- Anti-Eklf (α-Eklf) polyclonal antibody raised in rabbit was supplied by M. Crossley.
- Anti-β-actin (α-β-actin) monoclonal antibody was supplied by *Sigma-Aldrich Company*.
- Horseradish peroxidase linked anti-rabbit and anti-mouse secondary antibodies were supplied by *GE Healthcare Life Sciences*.

2.2 Plasmids and oligonucleotides

2.2.1 Vectors

Vector	Provided by	Description
NpGEX2T	José Perdomo	Bacterial expression vector for expressing recombinant GST fusion proteins - contains a thrombin cleavage site following the GST tag
pEFIRES-P	Alexis Verger	Mammalian bicistronic expression vector containing an internal ribosome entry site (IRES) element between the puromycin resistance (*pac*) gene and the recombinant gene of interest
pGEX6P-1	Roland Gamsjaeger	Bacterial expression vector for expressing recombinant GST fusion proteins - contains a PreScission™ protease cleavage site following the GST tag

2.2.2 Gift plasmids

Plasmid	Provided by	Description
pc3HA.Lklf	Jane van Vliet	Mammalian expression vector encoding full length Lklf fused to a haemagglutinin (HA) tag
pEFIRES-P.GST-Bklf 241-344	Noelia Nunez	Mammalian expression vector encoding the zinc fingers (ZFs) of Bklf fused to the GST tag
pMT3.mKlf8	Alister Funnell	Mammalian expression vector encoding full length Klf8
pSG5.Eklf	James Bieker	Mammalian expression vector encoding the full-length Eklf protein

2.2.3 Constructs

Plasmid	Template	Cloning primers	Restriction sites (5', 3')
NpGEX2T.Eklf 261-376	murine spleen cDNA	A2169, A2170	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 261-376 C295A	NpGEX2T.Eklf 261-376	A3138, A3139 (one-step mutagenesis)	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 261-376 C325A	NpGEX2T.Eklf 261-376	A3140, A3141 (one-step mutagenesis)	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 261-376 C355A	NpGEX2T.Eklf 261-376	A3142, A3143 (one-step mutagenesis)	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 261-376 C295A/C325A	NpGEX2T.Eklf 261-376 C325A	A3138, A3139 (one-step mutagenesis)	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 261-376 C325A/C355A	NpGEX2T.Eklf 261-376 C355A	A3140, A3141 (one-step mutagenesis)	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 261-376 C295A/C355A	NpGEX2T.Eklf 261-376 C355A	A3138, A3139 (one-step mutagenesis)	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 261-376 C295A/C325A/C355A	NpGEX2T.Eklf 261-376 C325A/C355A	A3138, A3139 (one-step mutagenesis)	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 261-348	NpGEX2T.Eklf 261-376	A2169, A3230	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 261-318	NpGEX2T.Eklf 261-376	A2169, A3231	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 261-294	NpGEX2T.Eklf 261-376	A2169, A3232	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 289-376	NpGEX2T.Eklf 261-376	A3227, A2170	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 319-376	NpGEX2T.Eklf 261-376	A3228, A2170	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 349-376	NpGEX2T.Eklf 261-376	A3229, A2170	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 289-318	NpGEX2T.Eklf 261-376	A3227, A3231	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 319-348	NpGEX2T.Eklf 261-376	A3228, A3230	*Eco*RI, *Bam*HI
NpGEX2T.Eklf 289-348	NpGEX2T.Eklf 261-376	A3227, A3230	*Eco*RI, *Bam*HI
pEFIRES-P.GST	NpGEX2T	A2156, A2254	*Nco*I, *Nde*I
pEFIRES-P.GST	NpGEX2T	A2156, A2255	*Nco*I, *Spe*I
pEFIRES-P.GST-Eklf 261-376	NpGEX2T.Eklf 261-376	A2252, A2253	*Nde*I, *Spe*I
pEFIRES-P.GST-Eklf 261-376 C295A	NpGEX2T.Eklf 261-376 C295A	A2252, A2253, A3138, A3139	*Nde*I, *Spe*I
pEFIRES-P.GST-Eklf 261-376 C325A	NpGEX2T.Eklf 261-376 C325A	A2252, A2253, A3140, A3141	*Nde*I, *Spe*I
pEFIRES-P.GST-Eklf 261-376 C355A	NpGEX2T.Eklf 261-376 C355A	A2252, A2253, A3142, A3143	*Nde*I, *Spe*I
pEFIRES-P.GST-Lklf 253-354	pc3HA.Lklf	A3043, A3044	*Nde*I, *Spe*I
pEFIRES-P.GST-Klf8 252-355	pMT3.mKlf8	A3086, A3087	*Nde*I, *Spe*I
pGEX6P-1.GST-Eklf 289-376	NpGEX2T.Eklf 289-376	A3285, A3286	*Bam*HI, *Eco*RI

2.2.4 Oligonucleotides for cloning and mutagenesis

Database reference	Sequence
A2156	AGACCATGGC CTCCCCTATA CTAGGTTAT
A2169	CCGGAATTCG GGGCCACTGC GATCGCC
A2170	CGCGGATCCT CAGAGGTGAC GCTTCATGTG
A2252	GGTTTTCCAT ATGTTGGGGC CACTGCGATC GCC
A2253	GGACTAGTTC AGAGGTGACG CTTCATGTG
A2254	GGTTTTCCAT ATGATTTTGG AGGATGGTCG CC
A2255	GGACTAGTTC ATTTTGGAGG ATGGTCGCC
A3043	GGTTTTCATA TGTTGAGGCC AAGCCCAAAC GC
A3044	GGACTAGTCT ACATGTGTCG CTTCATGTG
A3086	GGTTTTCATA TGTTACAATG GTCCACATGC AG
A3087	GGACTAGTTC ACATGGTGTC ATGACGCC
A3138	CAGGCGGCAC ATACGGCCGG GCACGAAGGC TGC
A3139	GCAGCCTTCG TGCCCGGCCG TATGTGCCGC CTG
A3140	GAGAAGCCTT ATGCCGCCTC CTGGGACGGC TGT
A3141	ACAGCCGTCC CAGGAGGCGG CATAAGGCTT CTC
A3142	CATCGTCCCT TCTGCGCTGG CCTCTGCCCA CGT
A3143	ACGTGGGCAG AGGCCAGCGC AGAAGGGACG ATG
A3227	CCGGAATTCA GGCAGGCGGC ACATACG
A3228	CCGGAATTCG GAGAGAAGCC TTATGCC
A3229	CCGGAATTCG GACATCGTCC CTTCTGC
A3230	CGGATCCTCA AGTGTGCTTC CGGTAGTG
A3231	CGGATCCTCA CGTGTGCGTG CGCAGGTG
A3232	CGGATCCTCA CGTATGTGCC GCCTGCCTC
A3285	CGGATCCAGG CAGGCGGCAC ATACG
A3286	CCGGAATTCT CAGAGGTGAC GCTTCATGTG

2.2.5 Oligonucleotides for real-time PCR

Primer	Database reference	Sequence
18S rRNA primers	A1560	CACGGCCGGT ACAGTGAAAC
	A1561	AGAGGAGCGA GCGACCAA
Albumin ChIP primers	A2850	GGGATGAACA ACCTATGCAA TTC
	A2851	TGGGCCTTGG CATGGA
Fli-1 primers	A2896	ATCAGCCAGT GAGGGTCAAC
	A2897	GGCCATTCTT CTCGTCCATA
Fli-1 ChIP primers	A2983	CCGATCGCAA AGTGAAGTCA
	A2984	GCGGATCGAA AAAGAGACAG TT
Control fli-1 ChIP primers	A2985	AGGCCTCAAG GGCAACCT
	A2986	GCCCTGACCC CCATCTTT

2.3 Methods

2.3.1 General molecular biology techniques

Protocols for the commonly used molecular biology techniques employed in this investigation are outlined in Sambrook et al, 1989 (section references are provided below). Techniques that involved the use of a commercial kit were carried out as instructed in the manufacturers' protocols and are also listed below.

- Agarose gel DNA electrophoresis: *6.1-6.19*
- Agarose gel DNA purification: *6.22-6.23*
- Automated DNA sequencing: provided by the Automated DNA Sequencing Service, Sydney University Prince Alfred Macromolecular Analysis Centre (SUPAMAC), University of Sydney, Camperdown, Sydney
- DNA ligation: *1.53-1.69*
- DNA sub-cloning: *F.1-F.11*
- Ethanol precipitation of DNA/RNA: *E.10-E.14*
- FastPlasmid® Miniprep Kit protocol for the mini-preparation of plasmid DNA (*Eppendorf AG, Hamburg, Germany*)
- JETstar Plasmid Purification System: Plasmid Midi/Maxi Kit protocol used for the midi- or maxi- preparation of plasmid DNA (*Astral Scientific*)
- Phenol/chloroform extraction: *E.3-E.4*
- Polymerase Chain Reaction (PCR): *14.1-14.5, 14.14-14.21*

- Restriction endonuclease digestion of DNA: *5.24-5.32*
- SDS-polyacrylamide gel electrophoresis (SDS-PAGE): *6.36-6.43, 6.45, 18.47-18.55*
- SuperScript™ First Strand Synthesis System protocol for cDNA synthesis (*Invitrogen*)
- Transformation of competent bacterial cells: *1.74, 1.76, 1.86*
- Western blots: *18.60-18.61, 18.64-18.66, 18.69-18.74*
- Western Lightning™ Chemiluminescence Reagent Plus Kit: protocol used for antibody detection in Western blots (*Perkin Elmer Life Sciences*)

2.3.2 One-step PCR site-directed mutagenesis

The one-step PCR site-directed mutagenesis technique was employed to generate the NpGEX2T.Eklf 261-376 single, double and triple zinc finger (ZF) point mutants. This technique uses the same principle as the QuikChange® II Site-Directed Mutagenesis Kit (*Stratagene*) and is based on the work of Fisher and Pei (1997). It involves the use of the high-fidelity *Pwo* polymerase and does not require subcloning. A pair of long complementary mismatch primers was used to introduce the desired point mutation on both strands of the plasmid. The entire plasmid is amplified by *Pwo* polymerase, resulting in unmethylated PCR product containing the base substitution. The restriction enzyme *Dpn*I was used to preferentially digest the methylated DNA template, while leaving the PCR product intact. The method is outlined in Figure 2.2.

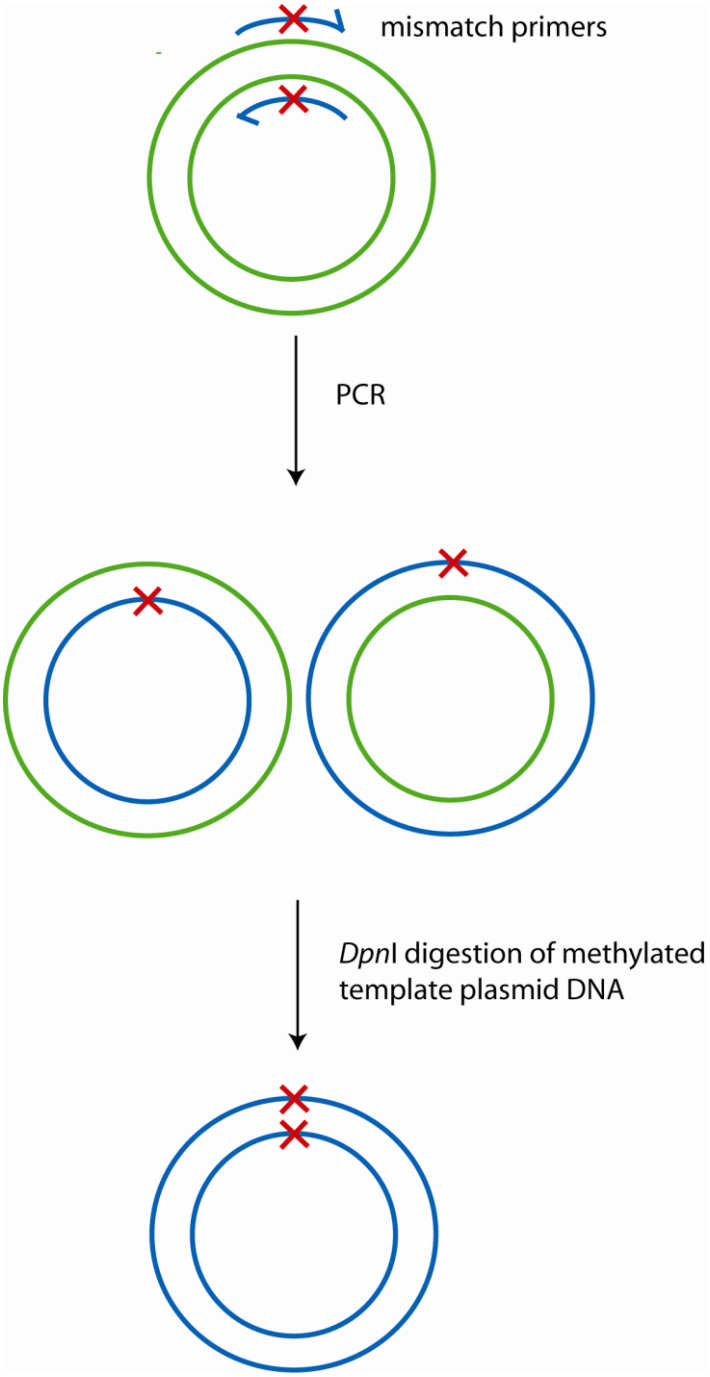


Figure 2.1 One-step PCR site directed mutagenesis technique.

A schematic representation of the one-step PCR site-directed mutagenesis technique used to generate the NpGEX2T-Eklf 261-376 ZF point mutants in this study. The methylated plasmid DNA template that is selectively digested by *Dpn*I is shown in green. The unmethylated PCR product with the incorporated point mutation is shown in blue.

2.3.3 Overlap PCR site-directed mutagenesis

Overlap PCR was used to generate the pEFIRES-P.GST-Eklf 261-376 single ZF point mutants used in this investigation, as the one step PCR site-directed mutagenesis technique was unsuccessful for this larger-sized construct. Overlap PCR site-directed mutagenesis was carried out as previously described (Bishop, 2001), and as illustrated in Figure 2.2. A pair of mismatch primers was used to introduce a substitution on both strands of the DNA, while a second set of primers bound to the sites flanking the insert. Two separate PCR reactions were performed to produce two different fragments with the same mutation; present on the 5' end of one fragment and on the 3' end of the other. These PCR products were then used as templates for a subsequent round of PCR using the flanking primers that contain restriction sites for cloning. The overlapping regions were allowed to anneal, thereby allowing extension to occur using the complementary strand as a template. The insert is regenerated in this second round of PCR to contain a nucleotide substitution within the sequence. The PCR products were digested with the appropriate restriction enzymes and sub-cloned into the vector of interest, in this case pEFIRES-P.GST.

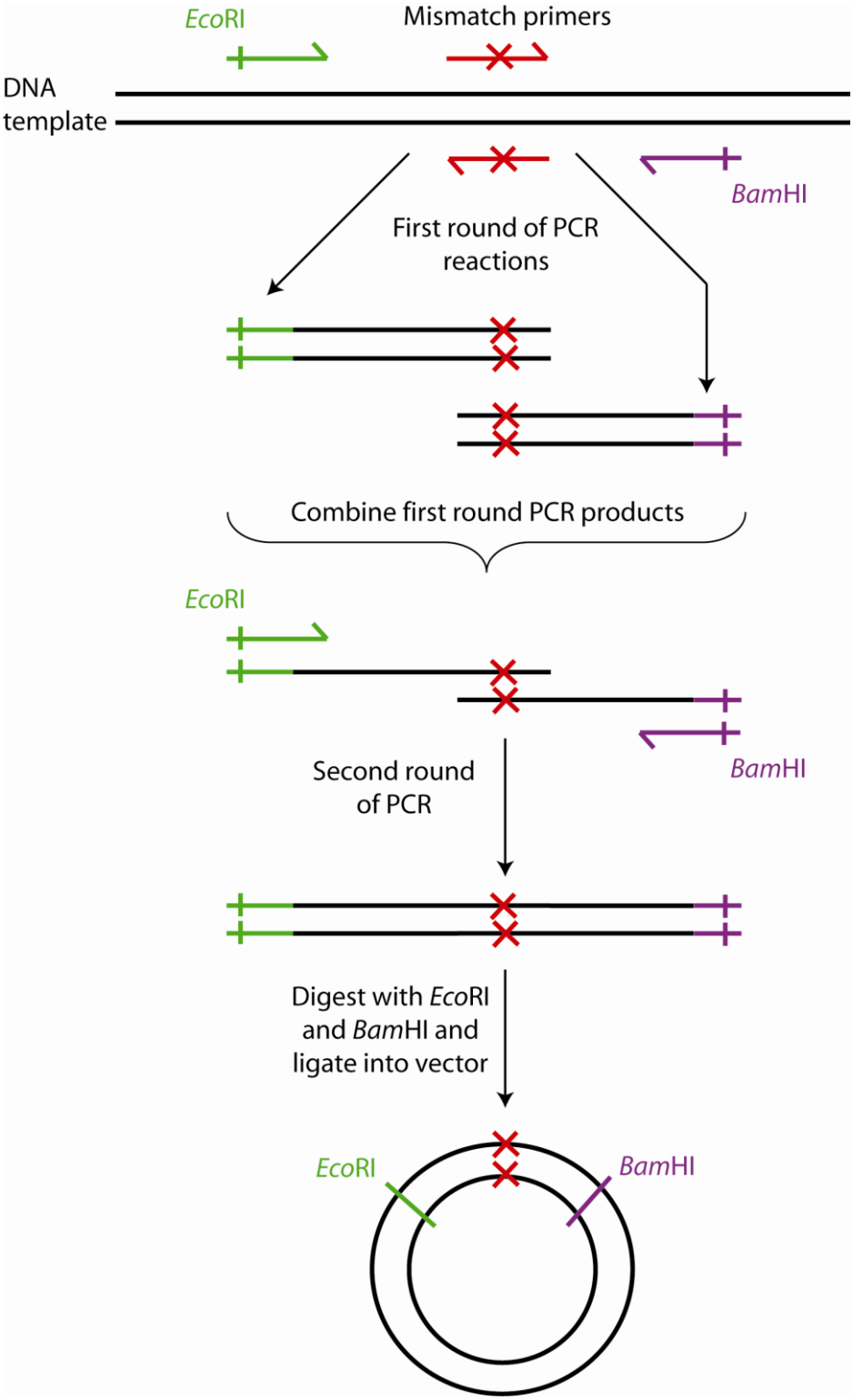


Figure 2.2 Overlap PCR site-directed mutagenesis technique.

A schematic representation of the overlap PCR site-directed mutagenesis technique used for the generation of the pEFIRES-P.GST-Eklf 261-376 point mutants used in this study.

2.3.4 Bacterial expression of recombinant GST fusion proteins

2.3.4.1 Expression of protein

Constructs were overexpressed in the bacterium Epicurian coli® BL21, which was grown in LB broth containing ampicillin (100 µg/mL) (LB+amp). A single colony containing the appropriate NpGEX2T or pGEX6P-1 plasmid was used to inoculate 10 mL LB+amp and the culture was allowed to grow through incubation for 16-18 h at 37°C with shaking at 180 rpm. An appropriate amount of culture was inoculated into 200 mL LB+amp to produce an A_{600} reading of 0.05. The culture was incubated at 37°C with shaking at 180 rpm until an A_{600} reading of approximately 0.6 was reached, and then protein overexpression was induced through the addition of IPTG (0.4 mM). After shaking at 180 rpm for 16-18 h at 20°C, cells were harvested by centrifugation (6,000 rpm, 15 min, 4°C) and the resultant cell pellets were stored at -20°C until lysis. The volume of culture used was scaled up or down as required.

2.3.4.2 Cell lysis and affinity purification

Cell pellets obtained from a 200 mL culture were resuspended in 8 mL of lysis buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.5% IGEPAL CA-630; supplemented prior to use with 1% Triton® X-100, 1 mM DTT, 10 mM $MgCl_2$, 0.2 mM PMSF, 5 µg/mL aprotinin and 5 µg/mL leupeptin) and sonicated (5× 15 s pulses at 40 W). The soluble fraction was separated from the insoluble fraction by centrifugation (15,000 rpm, 15 min, 4°C) and then incubated for 1 h at 4°C with a pre-swollen 50% (v/v) glutathione-agarose bead slurry (at a ratio of 5 mL bead slurry per L of culture used). The beads were washed 3× with 10 bead volumes of PBS-T (PBS containing 1% Triton® X-100) and

pelleted (500 *g*, 5 min, 4°C) between washes. Where elution of the protein was required, the beads were incubated at room temperature 3× 2 min with one bead volume of reduced glutathione solution (40 mM reduced glutathione, 100 mM Tris pH 7.5, 120 mM NaCl; the pH of this solution was adjusted to 9.5). Removal of the reduced glutathione was accomplished through the use of a 10 K Microsep™ centrifugal device (*Pall Gelman Sciences, Ann Arbor, MI, USA*) (7,500 *g*, 2× 60 min spins, 4°C). The eluted protein was stored at -80°C until use.

2.3.5 RNA trapping

2.3.5.1 Maintaining a ribonuclease-free environment

A ribonuclease (RNase)-free environment was maintained for all experiments requiring the handling of RNA (including this section and subsequent sections 2.3.6, 2.3.7 and 2.3.8) to minimise RNA degradation. All MQW and salt solutions were subjected to overnight treatment with 0.1% (v/v) diethylpyrocarbonate (DEPC) to deactivate RNases, prior to autoclaving. Surfaces were swabbed with a 1 M NaOH solution and allowed to dry. Cuvettes, pH-probes and stirrer bars were rinsed with 1 M NaOH, followed by thorough rinsing with DEPC-treated MQW. Glassware was baked at 160°C for 4 h or overnight, and spatulas and tweezers were flamed before use. RNase-free plasticware was used wherever possible, and all chemicals and disposable materials were obtained from supplies dedicated to RNase-free use.

2.3.5.2 Preparation of lysate from murine erythroleukaemia (MEL) cells

MEL cells were grown in low-glucose DMEM (to a cell density of approximately 2.2 $\times 10^6$ cells/mL) and harvested by centrifugation (3,200 *g*, 4 min, 4°C). Cell pellets were washed several times with ice-cold DEPC-treated PBS, resuspended in 1.5 volumes of lysis buffer (100 mM KCl, 25 mM EDTA, 5 mM $MgCl_2$, 10 mM HEPES, pH 7.0, 0.5% IGEPAL CA-630, 1% Triton® X-100, 0.1% SDS, 10% glycerol; and supplemented prior to use with 2 mM DTT, 1 tablet/10 mL Complete Mini® protease inhibitor, 100 U/mL RNasin® and 0.2% vandyl ribonucleoside complexes), and then centrifuged (14,000 *g*, 10 min, 4°C). The supernatant was collected and the pellet was resuspended in an additional 1 volume of lysis buffer and centrifuged. The resulting supernatant was pooled with the previous one and frozen in 400 µL aliquots at -80°C until use. Extracts typically had a concentration of 50 mg/mL total protein.

2.3.5.3 Expression of recombinant GST fusion proteins and affinity purification

GST-Eklf 261-376 was bacterially expressed and subjected to glutathione affinity purification, as described in Section 2.3.4. However, in this case, the proteins were retained on the glutathione-agarose beads (in other words, the elution step was omitted).

2.3.5.4 RNA trapping and isolation of bound RNA

The glutathione-agarose bead slurry, containing the recombinant GST-Eklf 261-376 protein, was equilibrated by washing (4× 500 µL) with NT2 buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.05% IGEPAL CA-630) prior to incubation with the MEL cell lysate. The bead slurry (containing approximately 0.7 mg protein) was resuspended in NT2 buffer (supplemented prior to use with 100 U/mL RNasin®, 0.2% vanadyl ribonucleoside complexes, 1 mM DTT and 20 mM EDTA) and 450 µL MEL lysate was added (all to a final volume of 10 mL). Following overnight incubation at 4°C with rocking, the samples were centrifuged (3,200 *g*, 1 min, 4°C) to pellet the beads, then washed (8× 1 mL) with supplemented NT2 buffer. The beads were resuspended in 600 µL proteinase K digestion buffer (100 mM Tris, pH 7.5, 10 mM EDTA, 50 mM NaCl and 1% SDS) and incubated with 500 µg proteinase K at 50°C for 30 min, with occasional mixing. The RNA was extracted with an equal volume of phenol:chloroform:isoamyl alcohol, vortexed for 1 min and centrifuged (22,000 *g*, 10 min, 4°C), then re-extracted with an equal volume of chloroform. Extracted RNA was precipitated with 1 volume isopropanol, 60 µL 4 M ammonium acetate, 3 µL 1 M $MgCl_2$ and 8 µL glycogen at -20°C overnight. The RNA precipitate was collected by centrifugation (14,000 *g*, 30 min, 4°C). The pellet was washed with 80% ethanol, resuspended in DEPC-treated MQW and stored at -80°C. Formaldehyde agarose gel electrophoresis was employed for the visualisation of enriched RNA species.

2.3.6 Fluorescence Anisotropy

Fluorescence Anisotropy assays can be used to monitor binding interactions between two molecules. The technique relies upon the existence of transition moments for absorption and emission which lie along specific directions within the fluorophore structure (Valeur, 2001).

When samples capable of emitting fluorescence are illuminated with polarised light at an appropriate wavelength, fluorophores with their absorption transition moments aligned parallel to the direction of polarisation have the highest probability of absorption. As a result, the excited-state population within the illuminated sample is not randomly oriented, but instead has a larger proportion of excited molecules with their transition moments aligned parallel to the polarised exciting light (Figure 2.3). A short time (1–10 ns) passes before the excited fluorophores return to the ground state by emission of light. During this time the fluorophores in solution are not static, but undergo rotational diffusion. The extent of this diffusion is dependent on a number of parameters, including the viscosity of the solvent, and importantly, the size and shape of the rotating molecules (Lakowicz, 1999).

This technique was used to investigate the binding of the ZFs of Eklf to the four RNA homoribopolymers, each with a 5' fluorescein tag ($^{Fl\text{-}}$RNAs). Short $^{Fl\text{-}}$RNA molecules rotate rapidly when free in solution due to their small size. Thus, when free $^{Fl\text{-}}$RNA molecules are illuminated with polarised light, by the time the excited state population

begins to relax to the ground state, the molecules will have undergone significant rotational diffusion, become randomised, and consequently emit depolarised light (Figure 2.3A). If however a large protein binds to the ^{Fl}RNA, the larger complex will rotate much more slowly, retaining some degree of polarisation at the time of emission, hence resulting in emission of polarised light (Figure 2.3B).

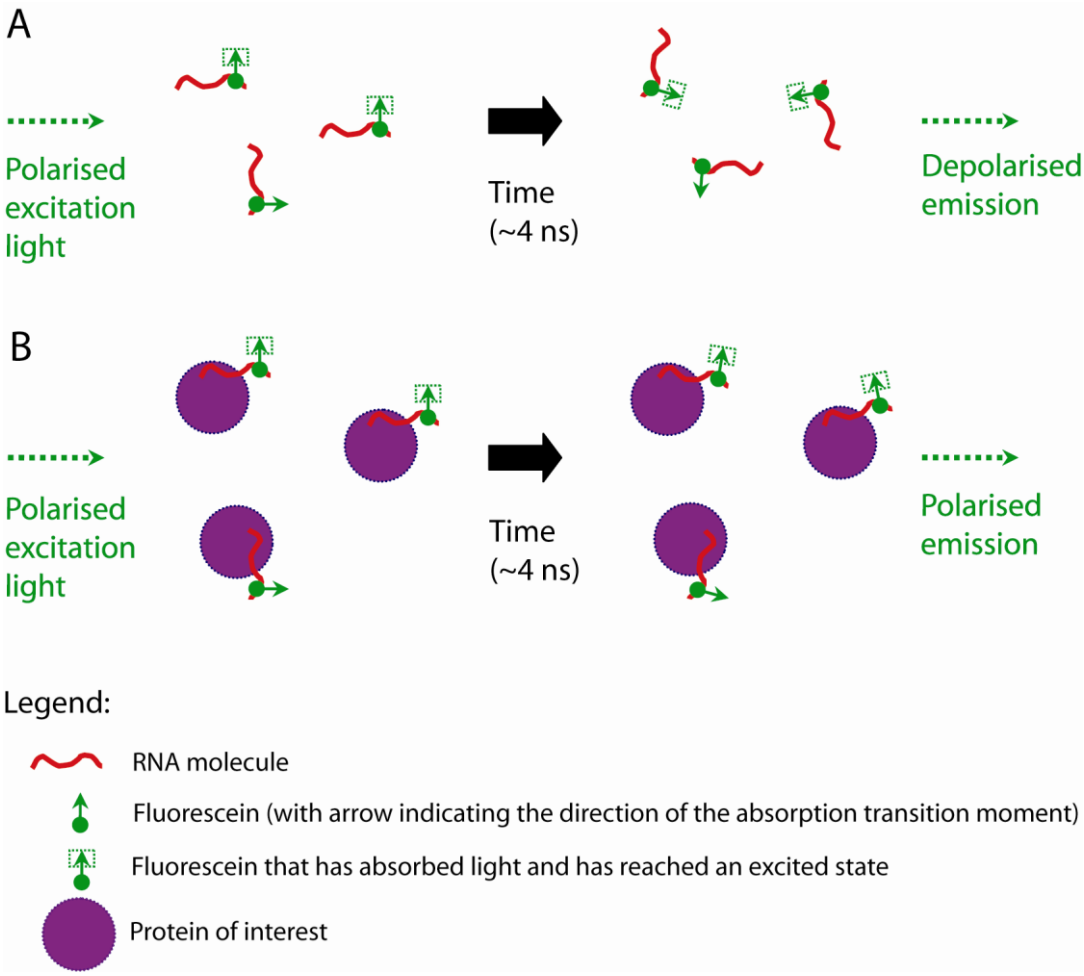


Figure 2.3 Schematic representation of fluorescence anisotropy.

A. Free ^{Fl}RNA molecules tumble rapidly in solution, becoming randomised by the time of emission, hence resulting in depolarised emitted light. **B.** Larger ^{Fl}RNA:protein complexes tumble more slowly in solution, retaining some degree of polarisation by the time of emission, hence resulting in polarised emitted light.

The degree of polarisation of a population of molecules is known as the anisotropy, r, and is given by the following equation:

$$r = \frac{I_{vv} - GI_{vv}}{I_{vv} + 2GI_{hv}}$$

where

I_{vv} = Fluorescence intensity when both the excitation and emission light are vertically polarised

I_{vh} = Fluorescence intensity when the excitation light is vertically polarised and the emission light is horizontally polarised

I_{hv} = Fluorescence intensity when the excitation light is horizontally polarised and emission light is vertically polarised

I_{hh} = Fluorescence intensity when both the excitation and emission light are vertically polarised

G = I_{hv}/I_{hh}

The G value is included to correct for any intensity bias of the fluorescence spectrophotometer towards vertically or horizontally polarised light. By calculating the anisotropy of a solution containing $^{Fl-}$RNA over a range of protein concentrations, a binding curve can be constructed, and from this a binding constant can be calculated using non-linear least squares regression.

2.3.6.1 Preparation of fluorescein-labelled RNAs

Fluorescein-labelled RNAs ($^{Fl-}$RNAs) were constructed for each of the four RNA homoribopolymers, each sequence consisting of a string of 17 nucleotides [poly(A), AAAAAAAAAAAAAAAAA; poly(G), GGGGGGGGGGGGGGGGG; poly(U), UUUUUUUUUUUUUUUUU; poly(C), CCCCCCCCCCCCCCCCC]. The $^{Fl-}$RNAs were resuspended to a concentration of 1 mM in DEPC-treated MQW. 5 µL aliquots were deprotected by diluting the $^{Fl-}$RNA 20-fold using the supplied deprotection buffer (100 mM acetic acid, adjusted to pH 3.8 with TEMED) and vortexing for 10 s before incubation at 60°C for 30 min. The $^{Fl-}$RNA was then filtered through 0.22 µm Spin-X centrifuge tube filters (10,000 rpm, 30 min, 4°C) and quantified by A_{260} and A_{493} using the extinction coefficient (ε) supplied with each $^{Fl-}$RNA and corrected for fluorescein absorbance by the equation:

$$[RNA] = \frac{A_{260} \cdot \varepsilon_{493}^{Fl} - A_{493} \cdot \varepsilon_{260}^{Fl}}{\varepsilon_{260}^{RNA} \cdot \varepsilon_{493}^{Fl}}$$

where $\varepsilon_{493}^{Fl} = 74\ 600\ M^{-1}cm^{-1}$ and $\varepsilon_{260}^{Fl} = 26\ 000\ M^{-1}cm^{-1}$ (Heyduk et al., 1996). Aliquots of $^{Fl-}$RNA were stored at -20°C until required. Exposure of the $^{Fl-}$RNAs to light was avoided wherever possible.

2.3.6.2 Preparation of proteins

Following elution of the recombinant GST-Eklf 289-376 protein off the glutathione-agarose beads (using PreScission™ buffer containing reduced glutathione; 50 mM Tris, pH 7.0, 100 mM NaCl, 2 mM DTT and 50 mM reduced glutathione), the GST tag was cleaved off the protein by the addition of 200 U PreScission™ protease/L of culture (*GE Healthcare Life Sciences*) and incubated for 1 h at 4°C. The reaction was stopped by a brief incubation at room temperature and the protein was purified using cation exchange chromatography, using the UnoS1 column with a gradient of 0-50% Buffer B (1 M NaCl) in 4 mL and a gradient of 50-100% Buffer B in 16 mL. Proteins were dialysed, against two separate volumes of fluorescence buffer, each for 4-16 h at 4°C. Aliquots of the purified protein were stored at -80°C until use.

The protein concentration was determined using the A_{280} and the extinction coefficient (ε), which was calculated from the amino acid sequence using the web-based program ProtParam on the ExPASy proteomics server (Swiss Institute of Bioinformatics, http://www.isb-sib.ch/). RNasin® was added to the protein solution to a concentration of 6.7 U/µL and incubated at 4°C for 5 min before dilution of the protein for the assays described in the next section.

2.3.6.3 Fluorescence anisotropy assays

Prior to use, all solutions used in fluorescence spectroscopy were filtered through 0.22 µm filters in order to remove light-scattering particles. Fluorescence intensity measurements were recorded at 520 nm following excitation at 475 nm, using a fluorescence spectrophotometer (*Cary Eclipse, Varian Instruments, Mulgrave VIC*) fitted with a manual polariser (*Varian Instruments, Mulgrave, VIC*) and long pass filters (*Coherent Scientific Pty. Ltd. Hilton, SA*) of 475 nm and 515 nm for excitation and emission, respectively. Temperature was maintained at 25°C using a block temperature controller (*Varian Instruments, Mulgrave, VIC*), slit widths were set to 10 nm and each recording was averaged over 15 s.

The reaction solution was prepared by the addition of 0.05 mg/mL heparin and 50 nM $^{Fl\text{-}}$RNA to fluorescence buffer in a starting volume of 300 µL. Measurements of I_{hv} (fluorescence intensity when the excitation light is horizontally polarised and emission light is vertically polarised) and I_{hh} (fluorescence intensity when both the excitation and emission light are vertically polarised) were obtained for this starting solution in order to correct for any intensity bias of the fluorescence spectrophotometer for vertically or horizontally polarised light. The stock protein solution was serially diluted into LoBind microfuge tubes, titrated into the reaction solution and mixed by inversion before the measurement of I_{vv} (fluorescence intensity when both the excitation and emission light are vertically polarised) and I_{vh} (fluorescence intensity when the excitation light is vertically polarised and the emission light is horizontally polarised) at each concentration of protein.

2.3.7 RNA homoribopolymer pulldown assay

Approximately 20 µg of recombinant GST fusion protein was incubated for 1 h at 4°C with approximately 3 µg of the homoribopolymers, poly(A), poly(G), poly(U) and poly(C), bound to Sepharose (*Sigma-Aldrich*), equilibrated in binding buffer (10 mM MOPS, pH 7.0, 50 mM KCl, 5 mM $MgCl_2$, 5% glycerol; supplemented prior to use with 0.1% Triton® X-100, 1 mM DTT, 0.1 mM PMSF and 400 U/mL RNasin®). The beads were washed 4× in binding buffer and the final pellet was drained and boiled for 5 min in 10 µL SDS-PAGE loading buffer and the bound protein separated by SDS-PAGE, and detected by Coomassie staining.

2.3.8 Systematic Evolution of Ligands by EXponential enrichment (SELEX)

The SELEX technique was first described by Tuerk and Gold (1990) and allows for the amplification and selection of the strongest binding ligands of a protein of interest. This method was used to select the best RNA ligands bound by recombinant, affinity-purified GST-Eklf 261-276 and is outlined in Figure 2.4.

A random RNA library was prepared using the synthetic oligonucleotide template 5'-CGCGGATCCTAATACGACTCACTATAGGGGCCACCAACGACATT [25N] TCTAGAATAAATAGTGCCCATGGATCCGCGGGTGTCGGG-3', where 'N' indicates random incorporation of all four nucleotides. This library was prepared for first round transcription by PCR amplification of the random N_{25} template, using the T7 (CGCGGATCCTAATACGACTCACTATAGGGGCCACCAACGACATT) and Rev

(CCCGACACCCGCGGATCCATGGGCACTATTTATTCTAGA) primers. RNA was transcribed using the RiboMax™ kit for T7 promoter (*Promega*) to yield approximately 5 nmol of full-length product (and approximately 1.1 $\times 10^{12}$ unique RNA molecules – equivalent to the coverage of a 20-mer sample space) for first round selection. Selection-amplification was performed essentially as described by Tuerk and Gold (1990). Binding reactions were carried out by incubating the RNA and GST-Eklf 261-376 bound to glutathione-agarose beads in SELEX binding buffer (10 mM MOPS, pH 7.0, 50 mM KCl, 5 mM $MgCl_2$, 5% glycerol; supplemented prior to use with 0.1% Triton® X-100, 1 mM DTT, 0.1 mM PMSF and 400 U/mL RNasin®) for 1 h at 4°C. RNA:protein molar ratios were varied from 5:1 at the beginning of selection to 20:1 at the end of selection. Following 5× 400 µL washes, the protein and bound RNA was eluted from the beads by incubating at room temperature for 15 min in 50 µL reduced glutathione (10 mM reduced glutathione, 50 mM Tris pH 8.0 – pH was adjusted to 9.5). The protein was separated from the RNA in the supernatant by three consecutive phenol-chloroform extractions, and the RNA was precipitated by overnight incubation at -80°C in 0.5 volumes of 7.5 M ammonium acetate and 2.5 volumes absolute ethanol. The precipitated RNA was collected by centrifugation (22,000 *g*, 30 min, 4°C), resuspended in 16 µL DEPC-treated MQW and reverse transcribed into cDNA using the Rev primer and the SuperScript™ III First-Strand Synthesis System for cDNA (*Invitrogen*). Half of the cDNA was used as template for a subsequent round of PCR. Following twelve rounds of binding and selection, the cDNA was digested with *Bam*HI, sub-cloned into the pBS-KS vector and sequenced using the T7 primer.

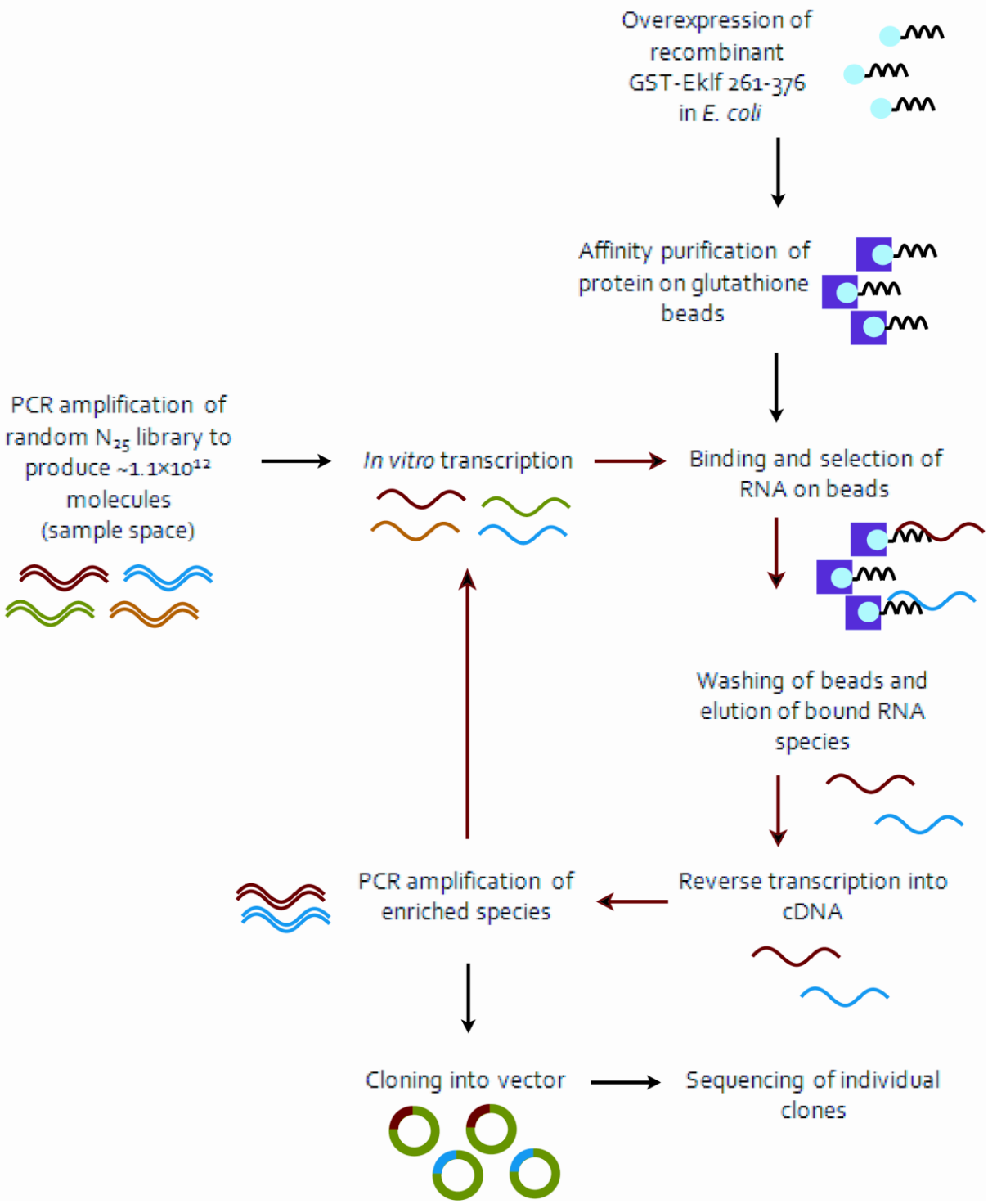


Figure 2.4 Schematic overview of the systematic evolution of ligands by exponential enrichment (SELEX) procedure.

After multiple rounds of protein binding, selection and amplification of the ligand RNAs for Eklf 261-376, isolated sequences are cloned and sequenced.

2.3.9 Electrophoretic Mobility Shift Assay (EMSA)

2.3.9.1 Probe preparation

The DNA probe used in this study is sequence from the murine *ß-major globin* promoter CACCC-box, using oligonucleotides MC132 (TAGAGCCACACCCTGGTAAG) and MC133 (CTTACCAGGGTGTGGCTCTA) as previously described (Crossley et al., 1996). Briefly, oligonucleotides were 5' end-labeled with [γ-^{32}P]-ATP using T4 polynucleotide kinase for 30 min at 37°C. To generate the double-stranded DNA probes, complementary oligonucleotides were annealed in TNE buffer (50 mM Tris, pH 8.0, 50 mM NaCl, 1 mM EDTA).

2.3.9.2 DNA-binding assays

Gel retardation assay reactions were set up in a total volume of 30 µL, comprising approximately 0.2 pmol of ^{32}P-labeled probe (final concentration 6 nM), 2 µg recombinant GST-fusion protein or 5 µL nuclear extract and gel-shift buffer (10 mM HEPES, pH 7.8, 50 mM KCl, 5 mM $MgCl_2$, 1 mM EDTA, 0.5 mM DTT, 5% glycerol and 25 µg/mL poly[dI·dC]) and 1 µL of the relevant antibody for antibody supershifting reactions. After incubation at 4°C for 20 min, samples were loaded onto a 6% native polyacrylamide gel (acrylamide:bisacrylamide ratio of 19:1 in 0.5× TBE). The gel was subjected to electrophoresis in a Sturdier™ vertical slab gel unit (*Hoefer, San Francisco, CA, USA*) in TBE (45 mM Tris, 45 mM boric acid, 1 mM EDTA) at 15 V/cm (250 V) for 1 h 50 min at 4°C. The gel was then dried under vacuum and analysed using a PhosphorImager™ (*Molecular Dynamics, Sunnyvale, CA, USA*).

2.3.10 Stable transfections

MEL and K562 cells were rinsed in HEPES buffered saline pH 7.05 (HeBS; 20 mM HEPES, 137 mM NaCl, 5 mM KCl, 0.7 mM Na_2HPO_4 and 6 mM glucose) and approximately 3 $\times 10^7$cells were used for each electroporation. One hundred micrograms of the appropriate pEFIRES-P plasmid was linearised with 200 U of the restriction enzyme *Ase*I and added to cells resuspended in 1 mL of HeBS in an electroporation cuvette (*Bio-Rad Laboratories*; 0.4 cm gap width). Cells were electroporated under the following conditions for both cell lines: 975 µF and 0.280 kV. Following electroporation, cells were incubated in 20 mL pre-warmed DMEM (for MEL cells) or RPMI (for K562 cells) for 48 h. Cells were serially diluted 10- and 100-fold and plated in 24 well plates in 1 µg/mL (MEL cells) or 0.5 µg/mL (K562 cells) puromycin, at 1 mL per well. Clones were picked and expanded once individual colonies were visible (after approximately one week) and maintained under puromycin selection.

2.3.11 Chemical induction of erythroid differentiation with dimethyl sulfoxide (DMSO)

Stable MEL cell lines were seeded at a concentration of approximately 2 $\times 10^5$ cells/mL and were chemically induced to undergo erythroid differentiation with 1.8% DMSO. Approximately 72 h post-induction, cells were harvested, washed twice in PBS to remove traces of media and enhance visualisation of the cell pellets and pelleted in a microfuge (22,000 *g*, 5 min, 4°C).

2.3.12 Nuclear extracts

Nuclear extracts were obtained from 100 mm culture dishes of MEL or K562 cells stably transfected with pEFIRES-P constructs. Cells were harvested and rinsed in 10 mL ice-cold phosphate-buffered saline (PBS). After centrifugation (10,000 *g*, 30 s), cells were resuspended in 400 µL buffer A (10 mM HEPES, pH 7.8, 1.5 mM $MgCl_2$, 10 mM KCl; supplemented with the following just before use: 1 mM DTT, 50 ng/µL PMSF, 5 µg/mL aprotinin and 5 µg/mL leupeptin), incubated at 4°C for 10 min and vortexed for 10 s. Following centrifugation (10,000 *g*, 10 s), the supernatant containing the cytosolic fraction was discarded and the pellet containing the nuclei was resuspended in 30 µL buffer C (20 mM HEPES, pH 7.8, 25% (v/v) glycerol, 420 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA; supplemented with the following just before use: 1 mM DTT, 50 ng/µL PMSF, 5 µg/mL aprotinin and 5 µg/mL leupeptin). Samples were incubated on ice for 20 min and then debris of the lysed nuclei were pelleted by centrifugation (22,000 *g*, 3 min, 4°C). The supernatants containing the nuclear extracts were collected and stored at -80°C until use.

2.3.13 SDS-PAGE

All protein preparations used in this study (mammalian and bacterial extracts) were separated by electrophoresis on polyacrylamide resolving gels (12% acrylamide from a 30% acrylamide stock [acrylamide:bisacrylamide ratio of 37.5:1], 0.375 M Tris [pH 8.8], 0.1% SDS, 0.1% APS and 0.1% TEMED). Stacking gels (4% acrylamide from a 30% acrylamide stock [acrylamide:bisacrylamide ratio of 37.5:1], 0.375 M Tris [pH 6.8], 0.1% SDS, 0.1% APS and 0.1% TEMED) were used for all SDS-PAGE runs. Prior to

electrophoresis, protein samples were heated to 95°C with an equal volume of 2× SDS-PAGE loading buffer (125 mM Tris [pH 6.8], 4% SDS, 20% glycerol, 40 µg/mL bromophenol blue and 100 mM DTT [added just prior to use]) for 5 min. Gels were run using either Mighty Small™ or Tall Mighty Small™ apparatus with 1 mm spacers (*Hoefer, San Francisco, CA, USA*) in protein running buffer (2.5 mM Tris, 0.192 M glycine, 0.1% SDS, pH 8.3) at 20 mA.

2.3.14 Western blotting

Nuclear extracts from MEL or K562 cells were separated by SDS-PAGE and transferred onto Biotrace™ nitrocellulose membranes (*Pall Gelman Sciences, Ann Arbor, MI, USA*), pre-soaked in Western transfer buffer (25 mM Tris, 0.19 M glycine, 20% (v/v) methanol). Transfer took place at 50 mA for 16-18 h at 4°C.

Following transfer, membranes were blocked twice in 50 mL skim milk solution (3 g skim milk powder in 100 mL TBST [50 mM Tris (pH 7.5), 150 mM NaCl, 0.05% Tween®-20] and washed twice in 50 mL TBST for 8 min each time. Membranes were then incubated for 1 h in primary antibody solution (diluted in 15 mL TBST). Following five washes in 50 mL TBST (8 min per wash), the membranes were incubated with the secondary antibody solution (2 µL horseradish peroxidise-linked anti-rabbit or anti-mouse secondary antibodies in 15 mL TBST). Membranes were washed an addition five times using 50 mL TBST (8 min per wash). Membranes were gently shaken throughout the washing and antibody incubation steps. Antibody detection was carried

out using the Western Lightning™ Chemiluminescence Reagent Plus Kit (*Perkin Elmer, Boston, MA, USA*) and exposure to X-ray film (*Eastman Kodak Company, Rochester, NY, USA*). The films were developed using an AGFA CP-1000 X-ray film processor using the developer and fixer supplied by the manufacturer.

Chapter 3 – Eklf binds RNA *in vitro* and *in vivo*

3.1 Introduction

Determining whether Eklf has the capacity to bind RNA will allow us to assess whether the family of Krüppel-like factors can regulate gene expression by binding directly to RNA. To date, only a small number of classical C_2H_2 zinc finger (ZF) proteins have been reported to bind RNA, including TFIIIA, JAZ, dsRBP-ZFa and wig-1 (Finerty and Bass, 1997; Honda and Roeder, 1980; Mendez-Vidal et al., 2002; Pelham and Brown, 1980; Picard and Wegnez, 1979; Yang et al., 1999). Furthermore, in addition to its nuclear function, Eklf was recently reported to localise to the cytoplasm (Quadrini et al., 2008; Shyu et al., 2007), thereby suggesting an important role for this protein other than through transcriptional regulation. Therefore, the aim of this chapter is to investigate whether Eklf has the capacity to bind RNA using its ZFs. We demonstrate that the ZFs of Eklf not only have the capacity to bind cellular RNA, but also display sequence specificity in homoribopolymer pulldown, fluorescence anisotropy and systematic evolution of ligands by exponential enrichment (SELEX) assays.

3.2 The ZFs of Eklf bind cellular RNA

To establish whether the Eklf ZFs bind RNA, a recombinant GST fusion protein containing the region spanning the ZFs of Eklf (amino acid residues 261-376; referred to as GST-Eklf 261-376) was used in the RNA pulldown experiment. Total murine

erythroleukaemia (MEL) cell extracts were used as a source of RNA. As can be seen in Figure 3.1, there was a significant enrichment in RNA levels for GST-Eklf 261-376 when compared to the GST negative control. Furthermore, the nucleic acid isolated was confirmed to be RNA by RNase digestion of the extracted nucleic acid (Figure 3.2).

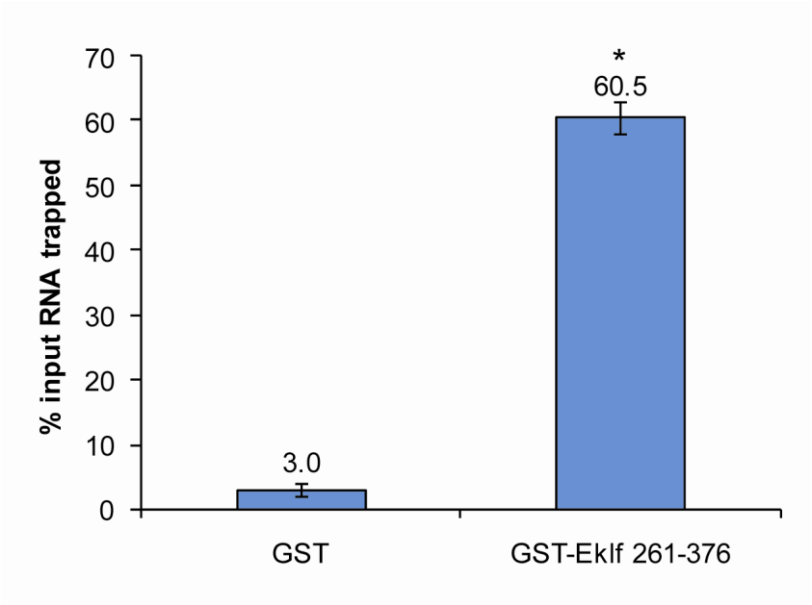


Figure 3.1 The ZFs of Eklf bind endogenous cellular RNA.

Recombinant GST and a GST-fusion protein containing the three ZFs of Eklf, spanning amino acid residues 261-376 (GST-Eklf 261-376), bound to glutathione-agarose beads (beads volumes containing approximately 0.7 mg of protein were used), were incubated with 400 µL MEL cell extract. After washing the beads, bound RNA was extracted and quantified using A_{260} readings. Results are expressed as a percentage of the input RNA used and are representative of three independent experiments. Data are presented as mean ± standard error of the mean (SEM), n=3. *$p<0.05$ (paired Student's t-test).

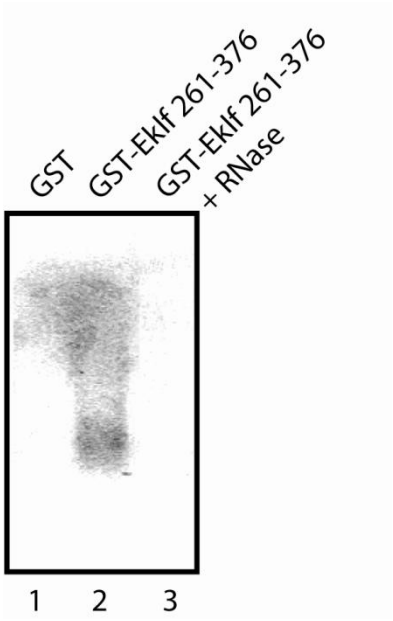


Figure 3.2 Formaldehyde agarose gel of RNA bound to GST-Eklf 261-376.

Recombinant GST and GST-Eklf 261-376 protein bound to glutathione-agarose beads, were used to pulldown RNA from MEL cell extracts (lanes 1 and 2). Approximately 5 µg of RNA bound by GST-Eklf 261-376 was treated with 30 U RNase ONE™ ribonuclease (compare lanes 2 and 3). Results shown are representative of two independent experiments.

Having established that the ZFs of Eklf are capable of binding RNA, we sought to characterise this binding in more detail using additional *in vitro* assays. This was accomplished through the use of RNA homoribopolymer pulldown assays and fluorescence anisotropy studies.

3.3 The ZFs of Eklf bind RNA homoribopolymers *in vitro*

In many instances, RNA-binding proteins have been identified by their ability to bind RNA homoribopolymers [poly(A), poly(G), poly(U) and poly(C)]. Proteins identified in such a manner include Polypyrimidine Tract Binding Protein (PTB), Mammalian Splicing Factor 1 (SF1) and Fragile X Mental Retardation Protein (FMRP) (Arning et al., 1996; Brown et al., 1998; Irwin et al., 1997; Siomi et al., 1993b).

In this study, a homoribopolymer pulldown assay was used to determine whether the ZFs of Eklf bind RNA *in vitro* and to allow us to determine whether this occurred in a sequence specific manner. The procedure involves incubating a recombinant protein with beads containing the homoribopolymers. Following multiple washes, the protein is eluted from the beads and visualised after SDS-polyacrylamide gel electrophoresis and Coomassie staining.

3.3.1 Eklf has a preference for poly(A) and poly(U) RNA homoribopolymers

Recombinant GST-Eklf 261-376 was incubated with beads containing homoribopolymers [either poly(A), poly(G), poly(U) or poly(C) (Figure 3.3; lanes 2-5, respectively)]. Following multiple washes, any bound protein was eluted from the beads by boiling in SDS-PAGE loading buffer. As shown in Figure 3.3, GST-Eklf 261-376 displays a clear preference for poly(A) and poly(U) (lanes 2 and 4, bottom panel) and weaker binding to poly(G) and poly(C) RNA sequences (lanes 3 and 5, bottom panel). These

results show that the ZFs of Eklf are capable of binding to RNA and suggest they may bind with sequence specificity, with a preference for strings of A or U. Thus, we sought to quantitate binding to the homoribopolymers and verify these results using an independent technique, as will be described in the following sub-section.

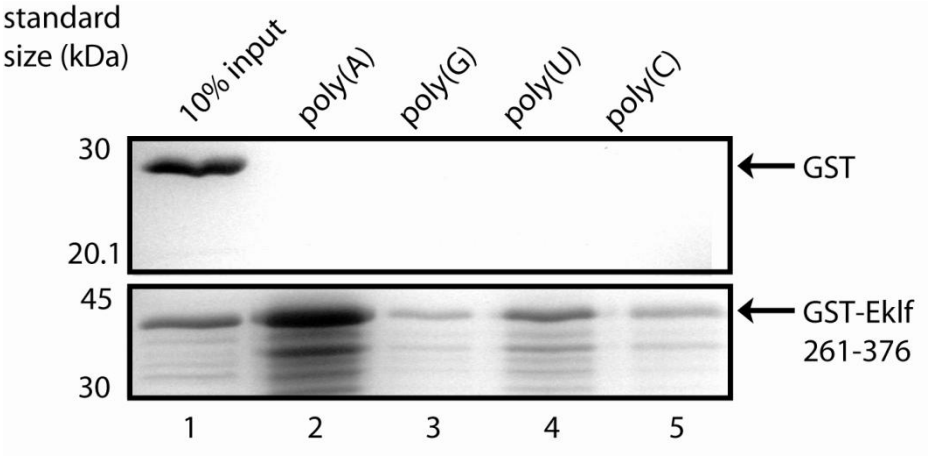


Figure 3.3 The ZFs of Eklf bind RNA homoribopolymers.

Recombinant GST and GST-Eklf 261-376 (20 µg) were incubated with ~3 µg poly(A), poly(G), poly(U) or poly(C) RNA (lanes 2-5, respectively) linked to agarose beads in binding buffer (10 mM MOPS pH 7.0, 50 mM $MgCl_2$, 5% glycerol, 0.1% Triton X-100, 1 mM DTT, 0.1 mM PMSF, 0.4 U/µL RNasin®) in a total volume of 100 µL. Beads were washed with 4× 400 µL binding buffer, loaded onto a 12% SDS-PAGE gel and stained with Coomassie Blue. Lane 1 shows 10% of the input (2 µg) used for each recombinant protein. Results shown are representative of three independent experiments.

We next sought to corroborate these results using a different technique, known as fluorescence anisotropy. In these quantitative assays, 17-nucleotide single-stranded RNA oligonucleotides were labelled with a 5' fluorescein moiety and their fluorescence anisotropy measured during the stepwise addition of purified ZFs of Eklf (this time with the GST moiety removed). As shown in Figure 3.4 and Table 3.1, relatively strong binding was observed for poly(A) and poly(U), with affinities of 3 μM and 0.3 μM, respectively (Table 3.1). In addition, no detectable level of binding was seen for poly(C), as binding was more than 100-fold weaker than that observed for poly(U) (Table 3.1). Furthermore, the poly(G) homoribopolymer displayed an uncharacteristic anisotropy binding curve (data not shown). Therefore, the binding constant could not be reliably calculated for poly(G) binding and thus the affinity for this homoribopolymer was not clearly established using this technique. One possible explanation for the unusual observation made for poly(G) is that the structure of this homoribopolymer may have interfered with the anisotropy readings, as strings of G residues are known to form planar G-quartet secondary structures (Williamson et al., 1989). Nevertheless, the remaining results obtained using this technique are consistent with the aforementioned homoribopolymer experiments and confirm that the ZFs of Eklf have a preference for strings of A or U sequences. Although poly(A) binding was observed to be stronger than poly(U) binding in the previous experiment (Figure 3.3), the former technique used was quantitative, as the amount of ribopolymer present on the beads was only approximately known. Thus, these two distinct experiments show the result that the classical ZFs of Eklf can bind to single-stranded RNA and suggest they bind with sequence specificity.

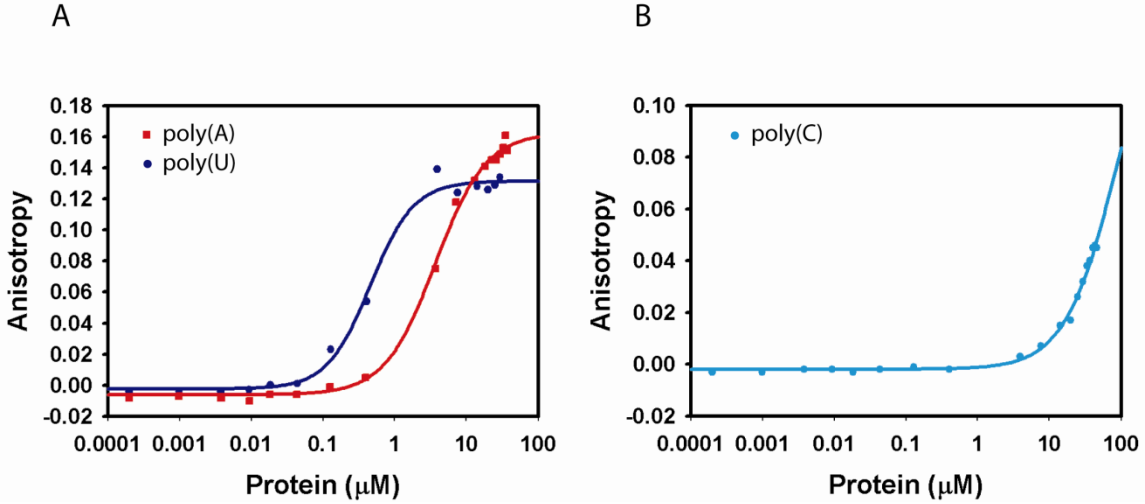


Figure 3.4 Comparison of the fluorescence anisotropy binding curves for the poly(A), poly(U) and poly(C) RNA homoribopolymers.

In each experiment, Eklf 289-376 was titrated into a 50 nM solution of fluorescein-labelled RNA and the anisotropy was calculated at each protein concentration. **A.** Fluorescence anisotropy binding curves for poly(A) and poly(U). **B.** Fluorescence anisotropy binding curve for poly(C). Curves were fitted using non-linear least squares regression. Data shown are representative of two independent experiments. The results obtained from this experiment were from a collaboration with J.Font.

Table 3.1 Relative binding affinities of Eklf 289-376 for RNA homoribopolymers

Average K_A values were calculated from two fluorescence anisotropy assays for each RNA species. Values less than 0.05 are indicative of weak or no binding. The results obtained from this experiment were from a collaboration with J. Font.

RNA substrate	**Binding constant (K_A, $\times 10^6$ M^{-1})**
poly(A)	0.26 ± 0.14
poly(U)	3.56 ± 0.63
poly(C)	<0.05

3.3.2 Point mutations in the ZFs of Eklf do not abolish RNA-binding

To investigate the mechanism of RNA-binding by Eklf 261-376, a number of point mutants were generated. A panel of mutants was generated, having the first cysteine residue of each of the three ZFs of Eklf (cysteine residues 295, 325 and 355) replaced with an alanine residue (Figure 3.5; combinations of C295A, C325A and C355A). It was anticipated that disruption of the ZF structure(s) caused by these point mutations might abolish or substantially weaken RNA-binding activity, as these mutations disrupt DNA-binding activity (data not shown). As ZFs are usually found in clusters, such mutational analysis is generally used to elucidate which of the ZFs are most important for nucleic acid binding.

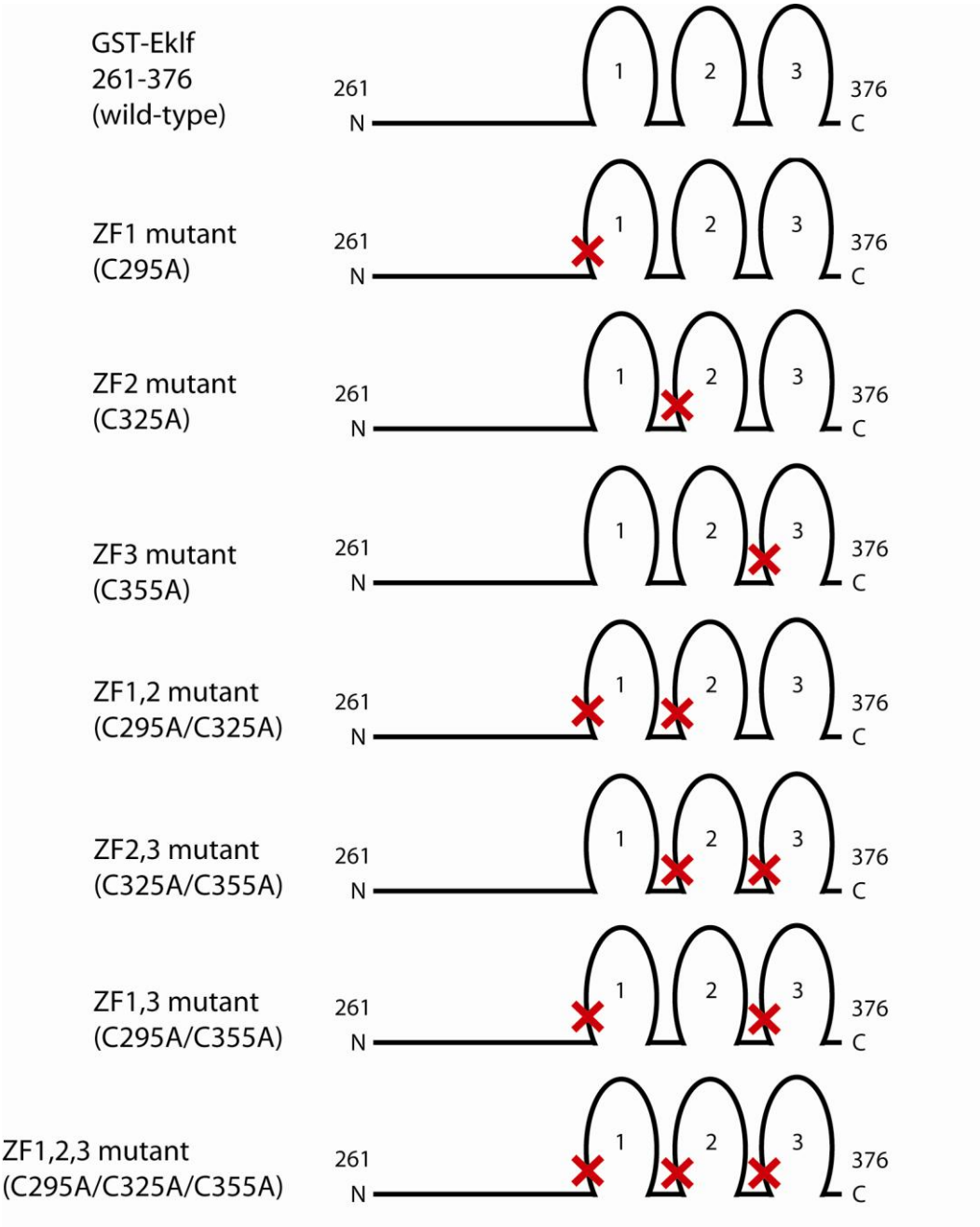


Figure 3.5 Schematic representation of GST-Eklf 261-376 point mutants.

Point mutations were generated using overlap PCR site-directed mutagenesis and then cloned into NpGEX-2T constructs, which were used to bacterially express these proteins. Each red cross depicts a mutation of the first cysteine residue within the ZF domain to an alanine residue.

Although the point mutations were predicted to disrupt the ZF structures and therefore disrupt binding, none of the point mutants generated substantially reduced binding (Figure 3.6, panels 2-9). This finding, namely that the zinc finger structures are not required for RNA-binding, is intriguing, given the importance of zinc finger binding implicated by the TFIIIA-RNA crystal structure (Figure 1.3). However, these results may suggest that the point mutations were not an effective means of disrupting RNA-binding activity. A number of possibilities exist that may explain the inability of the point mutations to abolish binding of the ZFs. The first possibility is that some or all of the point mutations partially disrupted the ZF structures and as a result, binding is not completely abolished. This is supported by the observation that binding appears to be slightly reduced with the double and triple ZF mutants (Figure 3.6, panels 6-9; C295A/C325A, C325A/C355A, C295A/C355A and C295A/C325A/C355A). The second possibility is that the ZFs are binding RNA in a non-specific manner, regardless of whether an intact structure is present or not. Furthermore, the possibility exists that the residues critical to RNA-binding lie outside the ZF domains, as 34 amino acid residues N-terminal to the first ZF are expressed in the recombinant protein (amino acid residues 261-294). However, fluorescence anisotropy experiments described in the previous subsection demonstrated high-affinity binding to poly(A) and poly(U), despite the protein lacking most of the residues preceding the ZFs (Eklf 289-376; Figure 3.4 and Table 3.1). Thus, these observations suggest that residues important for RNA-binding may exist both within the ZF domains and in the region preceding the first ZF. The above possibilities were examined through bioinformatics analysis and deletion mutant experiments, and are described in the following two subsections of this chapter.

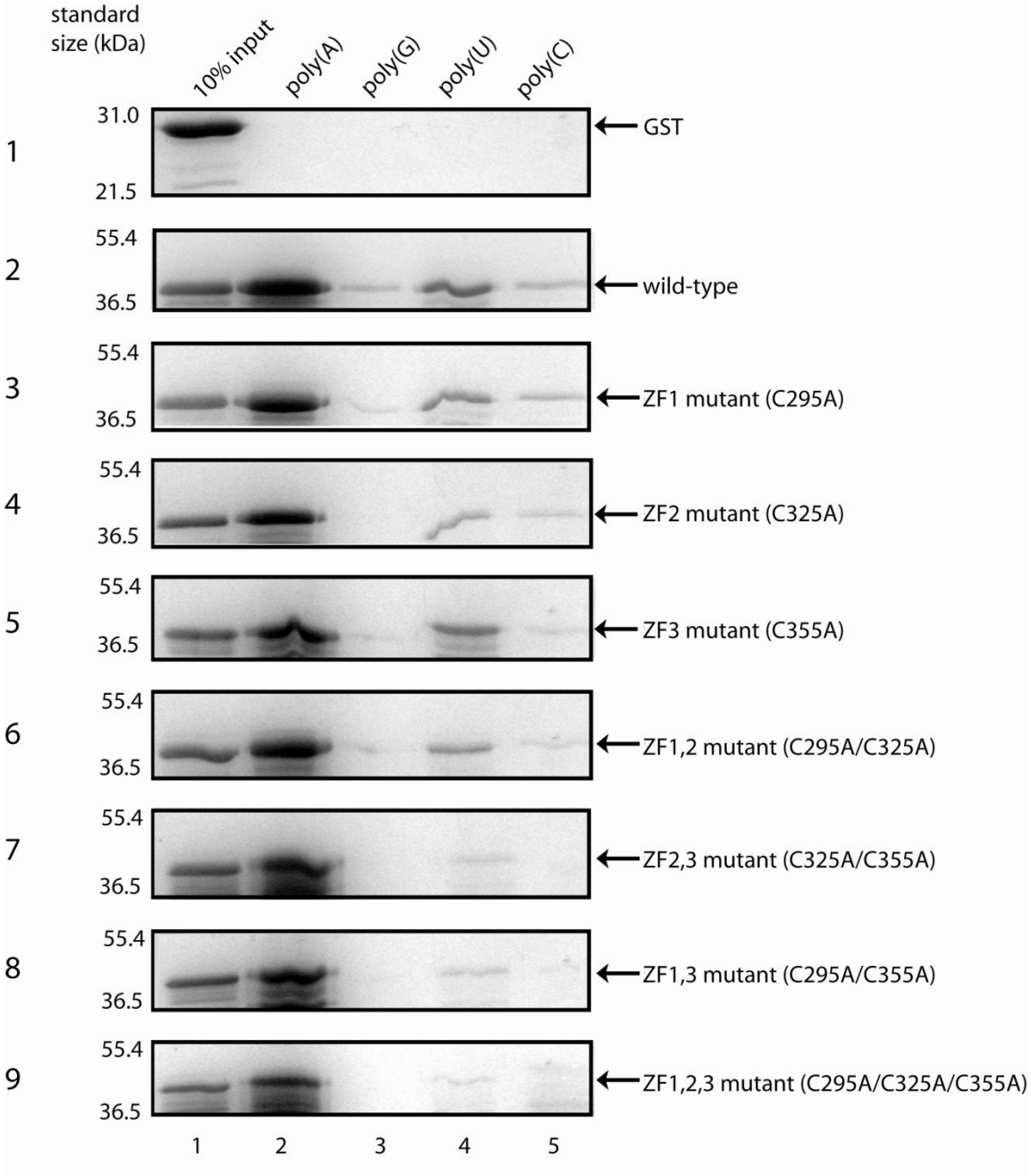

standard
size (kDa)
10% input
poly(A)
poly(G)
poly(U)
poly(C)
1
31.0
21.5
GST
2
55.4
36.5
wild-type
3
55.4
36.5
ZF1 mutant (C295A)
4
55.4
36.5
ZF2 mutant (C325A)
5
55.4
36.5
ZF3 mutant (C355A)
6
55.4
36.5
ZF1,2 mutant (C295A/C325A)
7
55.4
36.5
ZF2,3 mutant (C325A/C355A)
8
55.4
36.5
ZF1,3 mutant (C295A/C355A)
9
55.4
36.5
ZF1,2,3 mutant (C295A/C325A/C355A)
1
2
3
4
5

Figure 3.6 Point mutations in the ZFs of Eklf marginally disrupts binding to RNA homoribopolymers.

Recombinant GST (panel 1), GST-Eklf 261-376 (wild-type; panel 2) and GST-Eklf 261-376 point mutants (panels 3-9) (20 µg) were incubated with 3 µg poly(A), poly(G), poly(U) or poly(C) RNA (lanes 2-5, respectively) linked to agarose beads in binding buffer (10 mM MOPS pH 7.0, 50 mM $MgCl_2$, 5% glycerol, 0.1% Triton X-100, 1 mM DTT, 0.1 mM PMSF, 0.4 U/µL RNasin®) in a total volume of 100 µL. Beads were washed with 4× 400 µL binding buffer, loaded onto a 12% SDS-PAGE gel and stained with Coomassie. Lane 1 shows 10% of the input (2 µg) used for each recombinant protein. Results shown are representative of two independent experiments.

3.3.3 Bioinformatics analysis predicts the majority of RNA-binding residues are located within Eklf 261-376

In order to predict the residues within the Eklf protein that are important for RNA-binding, the amino acid sequence of Eklf was analysed using a number of bioinformatics programs (Kumar et al., 2007; Terribilini et al., 2006; Terribilini et al., 2007; Wang and Brown, 2006).

The RNABindR (Terribilini et al., 2006; Terribilini et al., 2007) program uses a Naïve Bayes classifier to predict the RNA-binding residues within proteins. Predictions are based on observed interactions from structures of protein-RNA interactions present in the protein data bank (PDB). It was reported that on the training set used to develop this program, 80% of the residues predicted to be RNA-binding were actually RNA-binding (Terribilini et al., 2006; Terribilini et al., 2007). The RNABindR prediction of

the RNA-binding residues within the Eklf (GenBank accession number NP_034765) protein sequence is shown in Figure 3.7. From the figure it is evident that the majority of the residues (45 out of 49 residues; approximately 92% of residues) predicted to bind RNA lie within amino acid residues 261-376. Furthermore, more than half of the residues predicted to bind RNA lie within the region 261-294 (the region immediately preceding the ZFs).

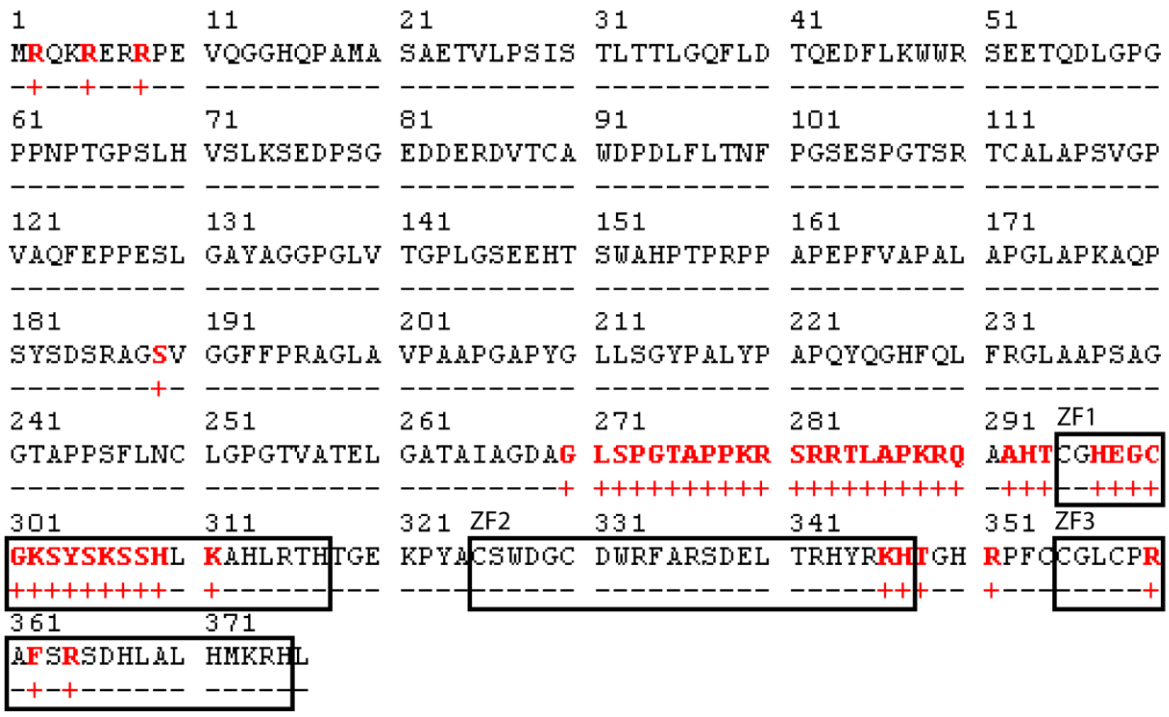


Figure 3.7 RNABindR optimal prediction of RNA-binding residues within the Eklf (GenBank accession number NP_034765) protein.

Residues predicted to bind RNA are in bold red and indicated by a '+'. The amino acid sequences of the three ZF domains of Eklf are shown in boxes. RNABindR (Terribilini et al., 2006; Terribilini et al., 2007) is a program available for use at http://bindr.gdcb.iastate.edu/RNABindR/RNABindR.aspx.

Prediction of protein-RNA interaction (Pprint) (Kumar et al., 2007) is another program that predicts RNA-binding residues within a protein sequence. It accomplishes this by using a support vector machines (SVM) model, which is trained on multiple sequence alignments. The SVM model was trained and tested on a set of 86 various RNA-binding protein chains with five-fold cross-validation.

The Pprint program assigns each amino acid residue within the protein sequence of interest an SVM value. A threshold SVM value, arbitrarily set by the user, is then used to predict the RNA-binding residues, with higher threshold values reducing the number of false positives predicted. This program is reported to have a prediction accuracy of 76%, using the default threshold SVM value of -0.2 (Kumar et al., 2007). As shown in Figure 3.8, the majority of the residues predicted to bind RNA (with an SVM value >0.2) lie at the C-terminal end of Eklf, within the region 261-376 (80 out of 91 residues have an SVM value >0.2; approximately 88% of residues). Taken together, the bioinformatics predictions made using two independent strategies suggest that the majority of RNA-binding residues exist within amino acid residues 261-376 of the Eklf protein. These predictions may explain the observations made in the point mutation experiments described in the previous subsection, as it appears that the residues preceding the ZFs, along with the ZFs, contribute to RNA-binding. In other words, introduction of the point mutations does not abolish RNA-binding, as the region immediately upstream of the ZFs may also be contributing to binding. The deletion mutant studies described in the next subsection will allow us to discern which regions are most critical for RNA-binding.

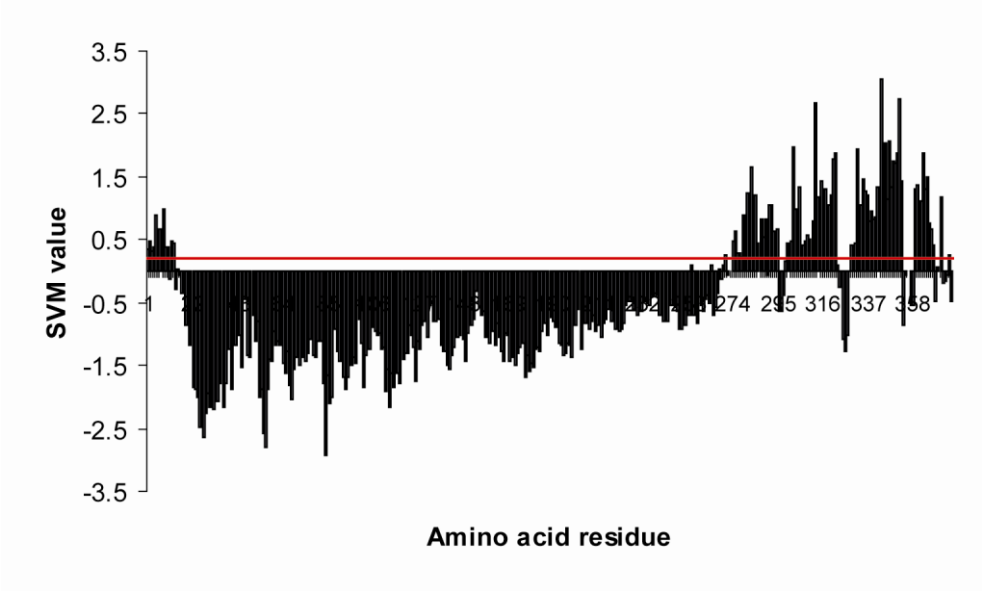


Figure 3.8 Prediction of protein-RNA interaction (Pprint) software prediction of RNA-binding residues within the Eklf protein.

Each residue is assigned a support vector machines (SVM) score. Residues predicted to interact with RNA have an SVM score greater than the threshold value of 0.2 (indicated by the solid red line). Pprint (Kumar et al., 2007) is a web program available for use at http://www.imtech.res.in/raghava/pprint/.

3.3.4 At least two ZFs are required for RNA-binding

In order to map the minimal region(s) within Eklf 261-376 that are critical for RNA-binding in finer detail, a series of deletion mutants was generated (Figure 3.9). These deletion mutants were then used in RNA homoribopolymer pulldown experiments in the same manner as described in subsections 3.3.1 and 3.3.2.

Figure 3.10 shows the results obtained from these experiments. Interestingly, the region containing all three ZFs, but lacking 28 of the preceding amino acid residues

shows binding which is similar to that of wild-type Eklf 261-376 (Figure 3.10, compare panels 2 and 6; i.e. GST-Eklf 261-376 and GST-Eklf 289-376). Furthermore, it appears that the region preceding the ZFs is not sufficient for binding, as binding is not strong when it alone is present in the absence of the ZFs (panel 5; GST-Eklf 261-294). Two tandem ZFs demonstrate similar binding to wild-type (compare panels 3, 7 and 11 with panel 2; i.e. compare GST-Eklf 261-348, GST-Eklf 319-376 and GST-Eklf 289-348 with GST-Eklf 261-376). However, in the presence of a single ZF, binding is dramatically reduced (panels 8-10; i.e. GST-Eklf 349-376, GST-Eklf 289-318 and GST-Eklf 319-348), although it is enhanced when ZF 1 is present in combination with the region preceding the ZFs (panel 4; GST-Eklf 261-318). Intriguingly, neither the region preceding the ZFs (panel 5; GST-Eklf 261-294) nor ZF 1 (panel 9; GST-Eklf 289-318) can bind poly(A) unless the two regions are present together (panel 4; GST-Eklf 261-318). Similarly, although ZF 3 appears to be the only region that is capable of binding poly(A) (panel 8; GST-Eklf 349-376), its binding is greatly enhanced when present in combination with the other two ZFs (panels 6 and 7; GST-Eklf 289-376 and GST-Eklf 319-376). Moreover, although ZF 2 cannot bind RNA in isolation (panel 10; GST-Eklf 319-348), it greatly enhances binding when present with either of the other two ZFs (panels 7 and 11; GST-Eklf 319-376 and GST-Eklf 289-348).

In order to test whether the mechanism of RNA-binding by the ZFs of Eklf is similar to DNA-binding, an electrophoretic mobility shift assay (EMSA) experiment was performed with some of the ZF deletion mutants (Figure 3.11). In this experiment, a DNA probe containing the *ß-major globin* promoter CACCC box was incubated with all three ZFs

(lane 3; GST-Eklf 289-376), ZFs 2-3 (lane 4; GST-Eklf 319-376), ZF 1 (lane 5; GST-Eklf 289-318) or ZF 2 (lane 6; GST-Eklf 319-348). In contrast to the previous experiment (Figure 3.10), only the protein containing all three ZFs was observed to bind to DNA (Figure 3.11, lane 3; GST-Eklf 289-376). Despite showing relatively strong RNA-binding (Figure 3.10, panel 7; GST-Eklf 319-376), no DNA-binding was observed for ZFs 2-3 (Figure 3.11, lane 4; GST-Eklf 319-376). These results suggest that the ZFs of Eklf bind to RNA using a mechanism that is different to DNA-binding.

Taken together, the above results demonstrate that the minimal region required for RNA-binding by Eklf is two ZF domains or one finger in combination with the 34 amino acid region preceding the fingers. In addition, the results above show that the ZFs act synergistically to enhance binding. Finally, the mechanism used for RNA-binding by the ZFs appears to differ from DNA-binding.

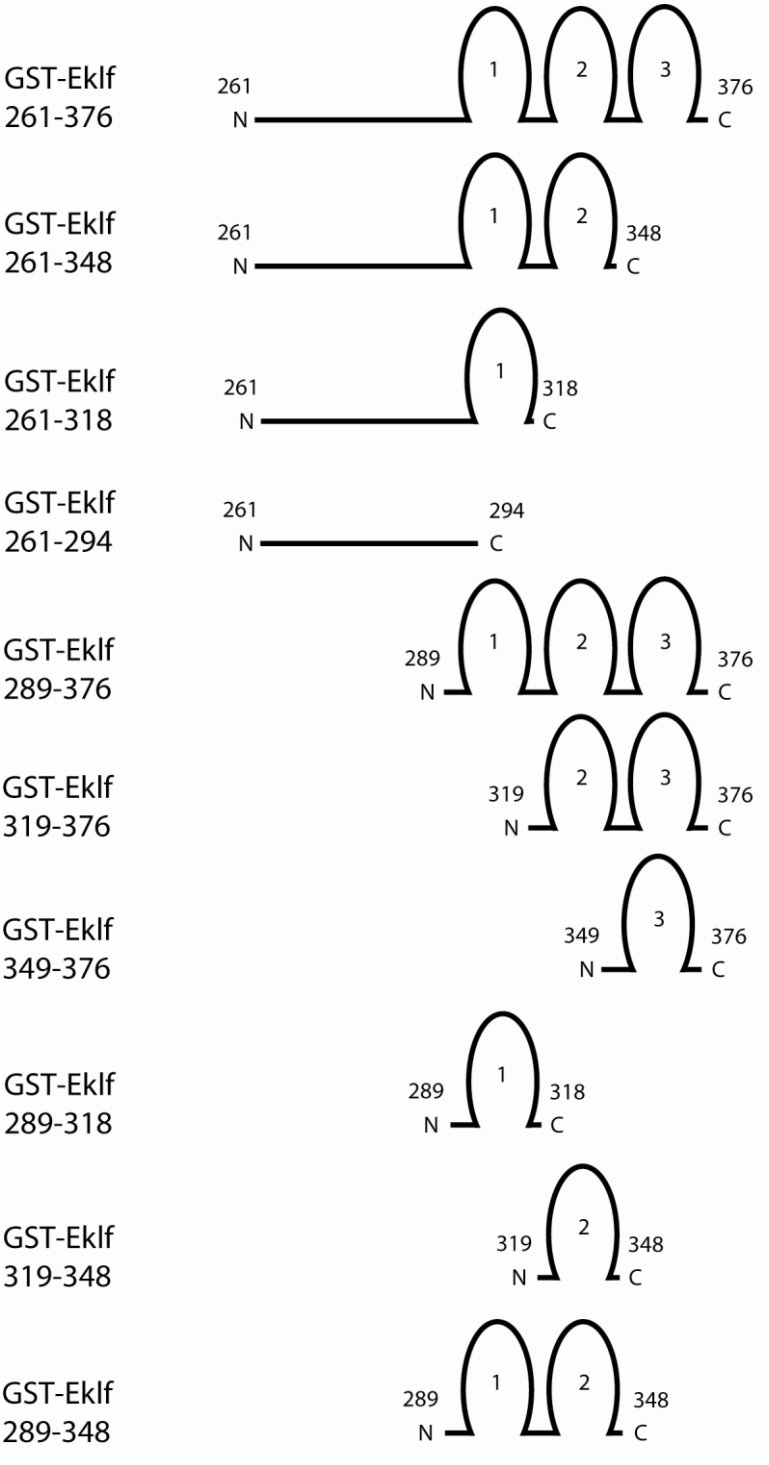


Figure 3.9 Schematic representation of GST-Eklf 261-376 deletion mutants.

N-terminal and C-terminal truncations of ZF domains of Eklf 261-376 were produced by bacterial overexpression from NpGEX-2T constructs generated to contain the desired deletions (by sub-cloning). Amino acid residues marking the N-terminal and C-terminal ends of each deletion mutant are indicated.

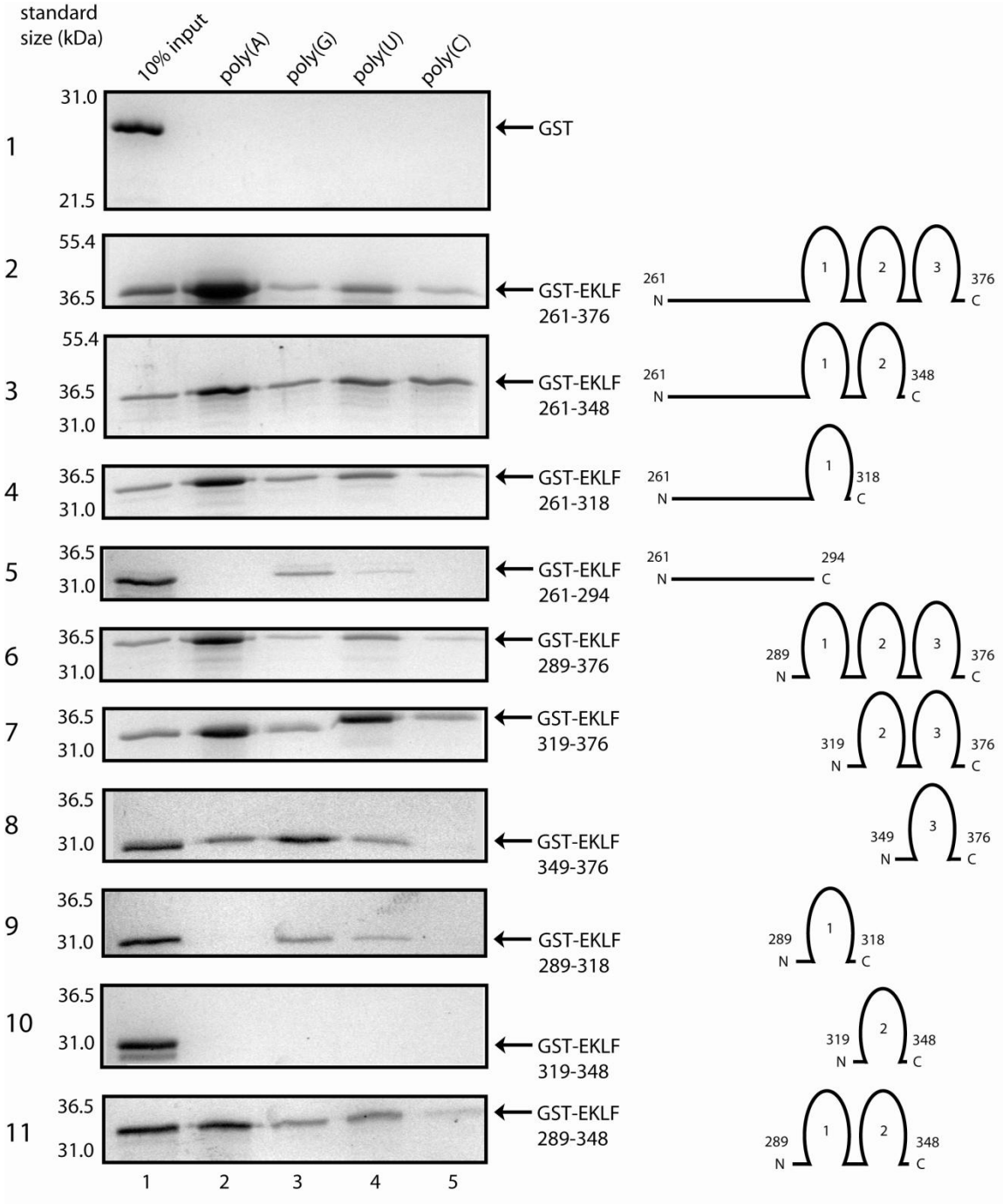

standard
size (kDa)
10% input
poly(A)
poly(G)
poly(U)
poly(C)
1
31.0
21.5
GST
2
55.4
36.5
GST-EKLF
261-376
3
55.4
36.5
31.0
GST-EKLF
261-348
4
36.5
31.0
GST-EKLF
261-318
5
36.5
31.0
GST-EKLF
261-294
6
36.5
31.0
GST-EKLF
289-376
7
36.5
31.0
GST-EKLF
319-376
8
36.5
31.0
GST-EKLF
349-376
9
36.5
31.0
GST-EKLF
289-318
10
36.5
31.0
GST-EKLF
319-348
11
36.5
31.0
GST-EKLF
289-348
1
2
3
4
5
261
N
376
C
348
318
294
289
319
349

Figure 3.10 The ZFs of Eklf display RNA sequence specificity.

Recombinant GST (panel 1), GST-Eklf 261-376 (panel 2) and GST-Eklf deletion mutants (panels 3-11) (20 μg) were incubated with 3 μg poly(A), poly(G), poly(U) or poly(C) RNA (lanes 2-5, respectively) linked to agarose beads in binding buffer (10 mM MOPS pH 7.0, 50 mM $MgCl_2$, 5% glycerol, 0.1% Triton® X-100, 1 mM DTT, 0.1 mM PMSF, 0.4 U/μL RNasin®) in a total volume of 100 μL. Beads were washed with 4× 400 μL binding buffer, loaded onto a 12% SDS-PAGE gel and stained with Coomassie Blue. Lane 1 shows 10% of the input (2 μg) used for each recombinant protein. Results shown are representative of two independent experiments.

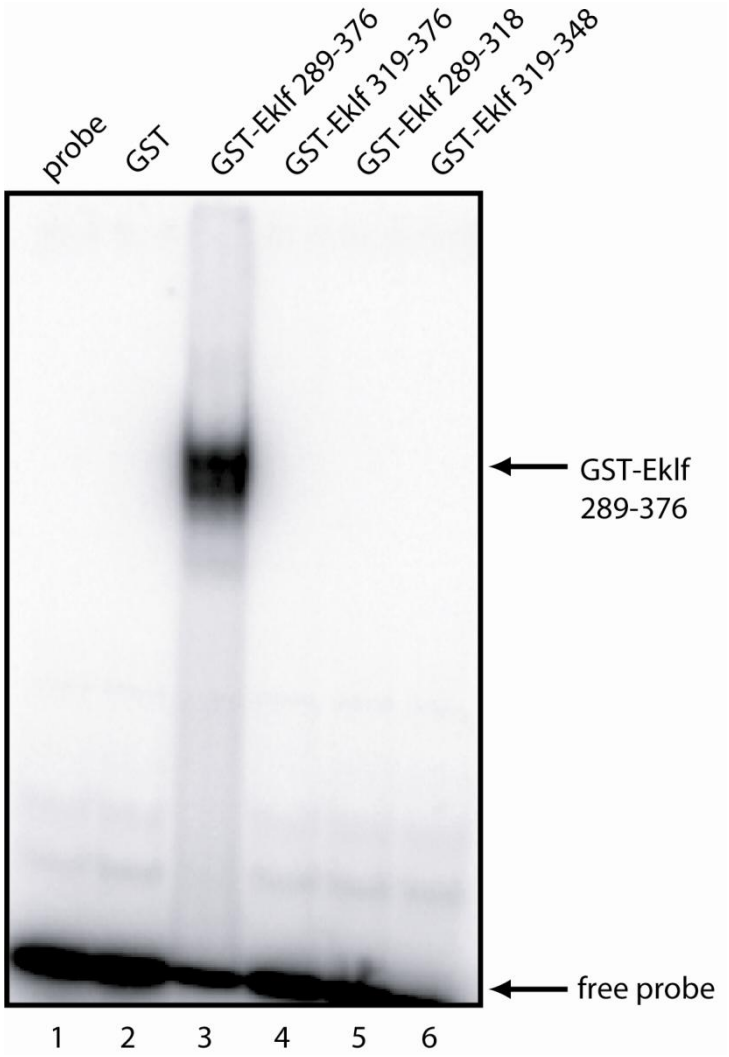


Figure 3.11 Eklf ZF deletion mutants lack the ability to bind DNA.

GST and GST-Eklf ZF deletion mutants (2 µg) were produced by bacterial overexpression (lanes 2-6). GST-Eklf 289-376 (lane 3) contains all three ZF domains; GST-Eklf 319-376 (lane 4) contains the second and third ZF domains; GST-Eklf 289-318 (lane 5) contains the first ZF domain; GST-Eklf 319-376 (lane 6) contains the second ZF domain. The recombinant proteins were incubated with a radiolabeled probe containing the *ß-major globin* promoter CACCC box. The reaction mixtures were then separated in a non-denaturing polyacrylamide gel in an electrophoretic mobility shift assay (EMSA). The interaction of GST-Eklf 289-376 with the DNA probe is shown by a shifted band.

3.4 Eklf does not appear to have a clear RNA-binding consensus sequence

As described in the previous section, the ZFs of Eklf display some degree of sequence specificity, with a preference for A or U homoribopolymers. Therefore, we carried out a selection experiment to define in detail the sequence specificity of the ZFs of Eklf and, if possible, to identify a consensus RNA-binding sequence. This involved the use of the systematic evolution of ligands by exponential enrichment (SELEX) technique (also referred to as SAAB – selected and amplified binding site imprint assay, or CASTing – cyclic amplification and selection of targets) (Blackwell and Weintraub, 1990; Wright et al., 1991). SELEX is a protocol for isolating from a pool of variant RNA sequences high-affinity ligands to a particular protein. Basically, this procedure involves cycles of affinity selection by a protein from a heterogeneous population of RNAs, replication of the bound species and *in vitro* transcription to generate an enriched pool of RNA.

A library of random 25-mer single-stranded RNA sequences, in which sequences of up to 20 bases in length were represented, was constructed. This random library was bound to the ZFs of Eklf immobilised on glutathione-agarose beads (Figure 3.12, lane 2). RNA sequences that were retained were subjected to multiple cycles of amplification and re-binding to the Eklf ZFs, until only the strongest binding sequences remained. These sequences were then identified through cloning and sequencing.

Following 12 rounds of binding and selection, sequences were obtained from 116 clones. Figure 3.13 shows the average frequency of each nucleotide within each sequence obtained. As can be seen, there is a significant enrichment in C residues present within each sequence. Furthermore, analysis of the nucleotide frequencies within the pooled sequences shows a clear enrichment of C residues (Table 3.2). Further analysis of the frequencies of dinucleotide, trinucleotide, tetranucleotide and pentanucleotide sequences was carried out (Tables 3.3-3.6) with the aim of identifying a consensus RNA-binding sequence for Eklf. The rationale for undertaking this type of analysis was the fact that RNA-binding proteins are generally found to bind short consensus sequences, since RNA-binding does not only involve recognition of specific sequences, but also the recognition of secondary structures. For instance, the Nova protein uses its K-homology (KH)-type domain to recognise a UCAY tetranucleotide element within the context of a 20-base hairpin RNA (Jensen et al., 2000). Despite there being a clear enrichment in C residues, an obvious consensus sequence could not be derived from these data (Tables 3.3-3.6). The data from this experiment appear to conflict with the earlier findings, in that the ZFs of Eklf appear to have a preference for C residues, whereas the previous experiments suggest a preference for A or U residues.

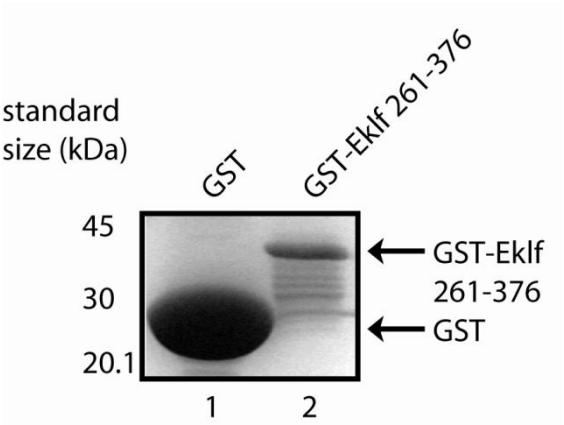


Figure 3.12 Recombinant GST-Eklf 261-376 protein used in SELEX experiments.

Coomassie stained 12% SDS-PAGE gel showing bacterially overexpressed GST (lane 1) and the GST fusion protein (lane 2), both of which were affinity-purified on glutathione agarose beads.

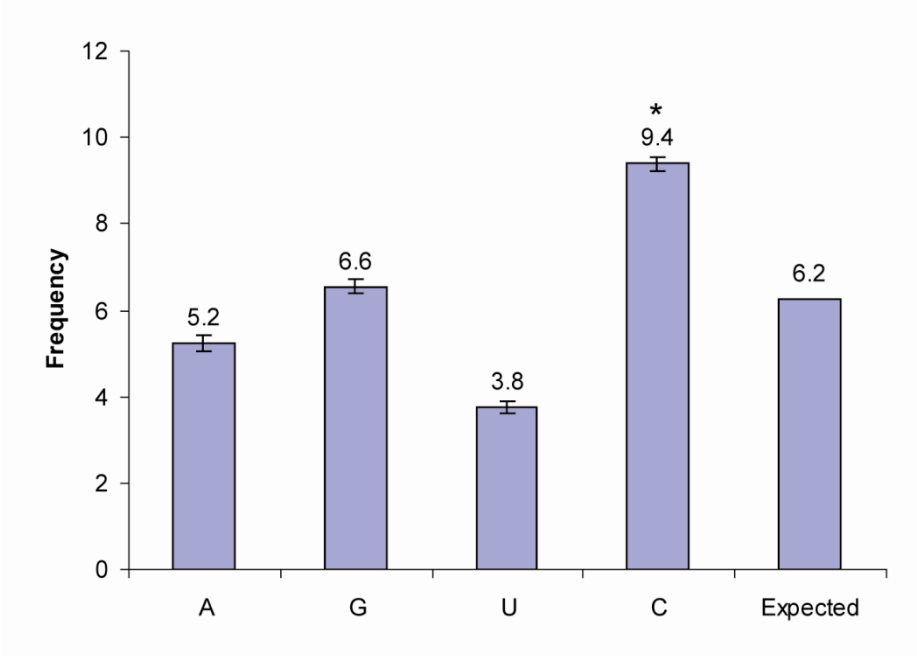


Figure 3.13 SELEX results showing significant enrichment for C residues.

Recombinant GST-Eklf 261-376 was mixed with RNA synthesised *in vitro* from oligonucleotides containing a randomised 25-mer central domain flanked by constant regions used for transcription and PCR. Bound RNAs were eluted and the binding and selection process was repeated for 12 rounds before cloning and sequencing. The average frequency of each residue from the 25-mer sequence of each isolated clone is shown. Average frequencies were obtained from the sequences of 116 isolated clones. Data are presented as mean ± SEM, n=116. *$p<0.05$ (paired Student's t-test).

Table 3.2 Nucleotide frequencies obtained from SELEX using the GST-Eklf 261-376 protein.

Recombinant GST-Eklf 261-376 was mixed with RNA synthesised *in vitro* from oligonucleotides containing a randomised 25-mer central domain flanked by constant regions used for transcription and PCR. Bound RNAs were eluted and the binding and selection process was repeated for 12 rounds before cloning and sequencing. Frequencies were obtained from a pool of 116 sequences following 12 rounds of binding and selection.

Nucleotide	Observed Frequency	Expected Frequency
C	1089	723.5
G	760	723.5
A	608	723.5
U	437	723.5

Table 3.3 Dinucleotide frequencies obtained from SELEX using the protein GST-Eklf 261-376.

Recombinant GST-Eklf 261-376 was mixed with RNA synthesised *in vitro* from oligonucleotides containing a randomised 25-mer central domain flanked by constant regions used for transcription and PCR. Bound RNAs were eluted and the binding and selection process was repeated for 12 rounds before cloning and sequencing. Frequencies were calculated from the 116 sequences obtained following 12 rounds of binding and selection. Sequences are shown in order of decreasing frequency.

Dinucleotide	Observed Frequency	Expected Frequency
CC	330	173.6
GC	308	173.6
CG	307	173.6
AC	262	173.6
CA	253	173.6
UC	164	173.6
GA	154	173.6
CU	149	173.6
AG	145	173.6
GG	141	173.6
UG	136	173.6
AA	107	173.6
GU	107	173.6
AU	82	173.6
UA	74	173.6
UU	59	173.6

Table 3.4 Trinucleotide frequencies obtained from SELEX using the protein GST-Eklf 261-376.

Recombinant GST-Eklf 261-376 was mixed with RNA synthesised *in vitro* from oligonucleotides containing a randomised 25-mer central domain flanked by constant regions used for transcription and PCR. Bound RNAs were eluted and the binding and selection process was repeated for 12 rounds before cloning and sequencing. Frequencies were calculated from the 116 sequences obtained following 12 rounds of binding and selection. Sequences are shown in order of decreasing frequency, with the trinucleotide sequences possessing the smallest frequencies in italics. A selection of the sequences with the highest and lowest frequencies is shown.

Trinucleotide sequence	**Observed Frequency**	**Expected Frequency**
CAC	107	41.6
GCC	105	41.6
CGC	105	41.6
CCG	95	41.6
ACC	88	41.6
CCC	87	41.6
GCA	80	41.6
AAA	*15*	*41.6*
AAU	*14*	*41.6*
UUA	*9*	*41.6*
UUU	*9*	*41.6*
AUU	*7*	*41.6*
UAU	*5*	*41.6*

Table 3.5 Tetranucleotide frequencies obtained from SELEX using the protein GST-Eklf 261-376.

Recombinant GST-Eklf 261-376 was mixed with RNA synthesised *in vitro* from oligonucleotides containing a randomised 25-mer central domain flanked by constant regions used for transcription and PCR. Bound RNAs were eluted and the binding and selection process was repeated for 12 rounds before cloning and sequencing. Frequencies were calculated from the 116 sequences obtained following 12 rounds of binding and selection. Sequences are shown in order of decreasing frequency, with the tetranucleotide sequences possessing the smallest frequencies in italics. A selection of the sequences with the highest and lowest frequencies is shown.

Tetranucleotide sequence	**Observed Frequency**	**Expected Frequency**
GCCC	40	9.9
CACC	38	9.9
ACGC	36	9.9
CGCA	34	9.9
GCAC	32	9.9
CCGC	30	9.9
CGCC	30	9.9
CACG	28	9.9
CCAC	28	9.9
CCUC	28	9.9
CGAC	28	9.9
AUAU	*0*	*9.9*
AUGU	*0*	*9.9*
AUUU	*0*	*9.9*
CUAU	*0*	*9.9*
UAUU	*0*	*9.9*
UUAU	*0*	*9.9*

Table 3.6 Pentanucleotide frequencies obtained from SELEX using the protein GST-Eklf 261-376.

Recombinant GST-Eklf 261-376 was mixed with RNA synthesised *in vitro* from oligonucleotides containing a randomised 25-mer central domain flanked by constant regions used for transcription and PCR. Bound RNAs were eluted and the binding and selection process was repeated for 12 rounds before cloning and sequencing. Frequencies were calculated from the 116 sequences obtained following 12 rounds of binding and selection. Sequences are shown in order of decreasing frequency, with the pentanucleotide sequences possessing the smallest frequencies in italics. A selection of the sequences with the highest and lowest frequencies is shown.

Pentanucleotide sequence	**Observed Frequency**	**Expected Frequency**
ACGCA	16	2.4
CACAC	14	2.4
CCUCG	14	2.4
ACCAC	13	2.4
GCCCG	13	2.4
CACCG	12	2.4
CACGC	12	2.4
CGCAC	12	2.4
CGGCC	12	2.4
GCACC	12	2.4
GCCGC	12	2.4
CGACG	11	2.4
GACGC	11	2.4
GCCCU	11	2.4
GCCUC	11	2.4
UGCCC	11	2.4
UUGCU	*0*	*2.4*
UUGGU	*0*	*2.4*
UUGUU	*0*	*2.4*
UUUAG	*0*	*2.4*
UUUAU	*0*	*2.4*
UUUCA	*0*	*2.4*
UUUCG	*0*	*2.4*
UUUCU	*0*	*2.4*
UUUGC	*0*	*2.4*
UUUGG	*0*	*2.4*
UUUUA	*0*	*2.4*
UUUUG	*0*	*2.4*

3.5 Discussion

The experiments described in this chapter show that the classical ZFs of Eklf can bind to single-stranded RNA and display some sequence specificity. These results represent the first description of RNA-binding by a member of the Klf family of transcription factors.

In this chapter we established that Eklf has the capacity to bind RNA in a cellular context and showed using two independent *in vitro* assays that Eklf has a preference for A or U homoribopolymers and displays high affinity binding to these sequences (Figures 3.3, 3.4 and Table 3.1). This preference for A or U may be of biological significance if it can be demonstrated that Eklf can bind to AU-rich elements (AREs). These sequence elements are typically present in the 3' untranslated region (3' UTR) of mRNA transcripts that are rapidly degraded (often with half-lives of less than one hour), including those encoding oncoproteins and signaling proteins such as cytokines, chemokines and inflammatory mediators (Barreau et al., 2005). ARE sequences are highly variable, although they usually contain one or more AUUUA pentameric motifs within or near a U-rich region (Fialcowitz-White et al., 2007). Their length is also variable, ranging between 30 and 120 nucleotides (Chen and Shyu, 1995). There are three known classes of ARE-binding proteins that can yield varying consequences when binding to their mRNA targets. For instance, AUF1, tristetraprolin and K homology splicing regulatory protein (KSRP) can direct the rapid degradation of RNA (Carballo et al., 1998; Chen et al., 2001; Gherzi et al., 2004; Lu et al., 2006; Raineri et al., 2004). Conversely, members of the Hu family of proteins inhibit the degradation of many ARE-containing mRNAs, presumably by antagonising recruitment of competing

mRNA-destabilising factors (Chen et al., 2002; Dean et al., 2001; Mobarak et al., 2000; Peng et al., 1998). Finally, members of the TIA-1/TIAR family do not appear to directly influence mRNA degradation; rather, they control the subcellular localisation and translational efficiency of bound transcripts (Lopez de Silanes et al., 2005; Piecyk et al., 2000). The RNA-binding activity demonstrated by the ZFs of Eklf may account for its unexplained presence in the cytoplasm (Quadrini et al., 2008; Shyu et al., 2007). Thus, the findings presented in this chapter could have interesting implications in uncovering a novel role for Eklf in the regulation of gene expression at the post-transcriptional level. To date, no classical ZF protein has yet been reported to bind ARE sequences within mRNA transcripts. Therefore, further work in this area may reveal a new family of proteins, the Klf family, as RNA-binding proteins.

In addition, deletion mutant studies, shown in Figure 3.10, revealed that Eklf requires at least two ZFs for RNA-binding or one ZF, when present in combination with the short amino acid peptide preceding the fingers. Moreover, EMSA studies shown in Figure 3.11 suggest that the mechanism of RNA-binding of the ZFs appears to differ from DNA-binding, since while two ZFs have the ability to bind RNA, they cannot bind DNA to an extent that can be detected in EMSA experiments. Further studies employing more quantitative approaches, such as Biacore, may allow for further elucidation of the mechanism of RNA-binding.

While the SELEX data appear to conflict with the findings described in the homoribopolymer pulldown and fluorescence anisotropy assays, there are two possibilities that may explain the seemingly contradictory data obtained using this technique. First of all, it is a well established fact that Eklf binds to C-rich sequences within the *ß -globin* promoter region of DNA, known as CACCC-boxes, with high affinity (Miller and Bieker, 1993). In addition, the deletion mutant and EMSA analyses described in this chapter provide preliminary evidence that the mode of RNA-binding of the ZFs of Eklf differs to the mode of DNA-binding, as mutants that possess the ability to bind RNA cannot bind DNA (Figures 3.9 and 3.10). Therefore, the C-rich sequences obtained using the SELEX technique may have resulted from DNA-binding and not RNA-binding. This is most likely due to the inadvertent carryover of trace amounts of DNA at the end of the *in vitro* transcription step, as a result of incomplete DNase digestion of the template DNA. This explanation could hold true especially if Eklf has a higher affinity for DNA than RNA, which means that even the slightest level of DNA contamination could lead to the unintended selective amplification of these species due to the high sensitivity of PCR. The second explanation is the possibility of sequence bias that may exist at both the PCR amplification and cloning steps, as this is one of the major limitations of the SELEX technique. Although in theory, the random library has a sample space representing all the possible 20-base combinations, in practice this is not necessarily the case. Some sequences are favoured over others as they are amplified and/or sequenced more efficiently. In light of the current findings, it appears that the former explanation is the more likely of the two. Nonetheless, further work is necessary to

elucidate the mechanism of Eklf RNA-binding in greater detail. The next two chapters will focus on examining the ZFs of Eklf using *in vivo* studies.

Chapter 4 – The ZF domain of Eklf behaves as a dominant negative mutant

4.1 Introduction

Having investigated the RNA-binding of the ZFs of Eklf using *in vitro* assays as described in the previous chapter, we sought to examine this in a more biological system. Specifically, we embarked on a procedure for identifying RNAs that might bind the ZF region of Eklf. In order to accomplish this, a cell line stably expressing the ZF region of Eklf, was produced. The major advantage of constructing such a stable cell line for these experiments is the ease of purification of overexpressed protein, particularly when an N-terminal GST purification tag is incorporated in the construct. In addition, use of overexpressed protein was preferred over endogenous Eklf protein for two reasons; firstly, due to the inability to reduce background signal under various immunoprecipitation conditions of endogenous Eklf using the existing Eklf antibody (data not shown) and secondly, in order to facilitate the isolation of RNA bound by the ZFs of Eklf and not by any other region within the protein.

In this chapter we describe the generation of a murine cell line that stably expresses a GST tag fused to the ZF region of Eklf (GST-Eklf 261-376). In addition, we provide evidence that expression of the ZFs within this stable cell line leads to an unexpected behaviour characteristic of a dominant negative mutant of Eklf. Moreover, we briefly

examine the specificity of the ZFs between members of the Klf family through the construction of various stable cells lines expressing the region containing the ZFs of some of these family members.

4.2 Generation of MEL stable cell line expressing GST-Eklf 261-376

This work was carried out in murine erythroleukaemia (MEL) cells. These cells are erythroid progenitors derived from the spleens of mice infected with the Friend virus complex (Friend, 1957). These virally transformed cells are arrested at the proerythroblast stage of development (Singer et al., 1974). Upon treatment with various chemical agents, MEL cells can be induced to undergo erythroid differentiation (between 30-100% of cells differentiate upon induction). Thus, this cell line was utilised due to its relevance to erythroid differentiation.

To construct the stable MEL cell line, the GST tag and the region encoding amino acids 261-376 of *Eklf* were sequentially sub-cloned into the bicistronic pEFIRES-P vector to produce a GST-Eklf 261-376 fusion protein. Using the pEFIRES-P vector, both the recombinant cDNA and the puromycin resistance gene (*pac*) are transcribed as a single mRNA transcript driven by the strong human polypeptide chain elongation factor 1α promoter (Hobbs et al., 1998). The presence of an internal ribosome entry site (IRES) within the message ensures that the majority of clones that are resistant to puromycin also express the recombinant protein. MEL cells were electroporated with either

pEFIRES-P.GST or pEFIRES-P.GST-Eklf 261-376 and expressing clones were selected by plating the cells in media containing puromycin. Recombinant protein obtained from the nuclear extracts of positive clones is shown in Figure 4.1.

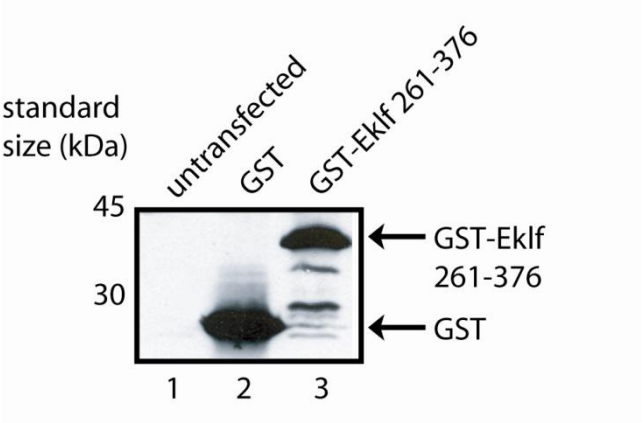


Figure 4.1 Expression of GST and GST-Eklf 261-376 in MEL nuclear extracts.

MEL cells were electroporated with either 100 μg pEFIRES-P.GST or pEFIRES-P.GST-Eklf 261-376. To generate the stable cell lines, puromycin was added to a final concentration of 1 μg/mL, 48h post-transfection. Individual puromycin resistant clones were picked, expanded and then harvested. Proteins were detected in a Western blot probed with the polyclonal Eklf antibody. Results shown are representative of four independent experiments.

4.3 The DNA-binding activity of endogenous Bklf is reduced

Using the stable cell lines described in the previous section, an EMSA experiment was conducted in order to determine whether the ZFs in the GST-Eklf fusion retained DNA-binding activity, thereby providing an indication that the recombinant protein was correctly folded. As shown in Figure 4.2, overexpression of GST-Eklf 261-376 in MEL

cells results in a strong band shift of the β-globin CACCC probe (lane 4, arrow), which is absent in the control stable cell line expressing the GST tag (lanes 1-3). The identity of the band in lane 4 is confirmed by antibody supershifting using an Eklf antibody (lane 5). Unexpectedly, the DNA-binding activity of Bklf was observed to be dramatically reduced in the GST-Eklf 261-376 stable cell line (compare the intensities of the bands corresponding to the Bklf-probe DNA complexes in lanes 1 and 4). This unusual observation led us to postulate that the decrease in Bklf protein may reflect a drop in *Bklf* mRNA transcript, since it is known that the *Bklf* gene is directly activated by Eklf (Funnell et al., 2007). In contrast, the DNA-binding activity of the Sp1/3 proteins appears to remain unaffected under these conditions (compare lanes 1 and 4). This observation could be attributed to the fact that these proteins are not known to be regulated by Eklf and therefore are unaffected. Furthermore, the reduction in endogenous Bklf DNA-binding activity was not observed in stable MEL cell lines expressing point mutations in the ZFs of Eklf (Figure 4.3A and B). Thus, these observations led us to speculate that the GST-Eklf fusion is behaving as a dominant negative mutant, thereby inhibiting the endogenous Eklf protein from activating its target genes. Therefore, these findings prompted us to further examine Bklf protein and transcript levels in this stable cell line, as described in the next section of this chapter.

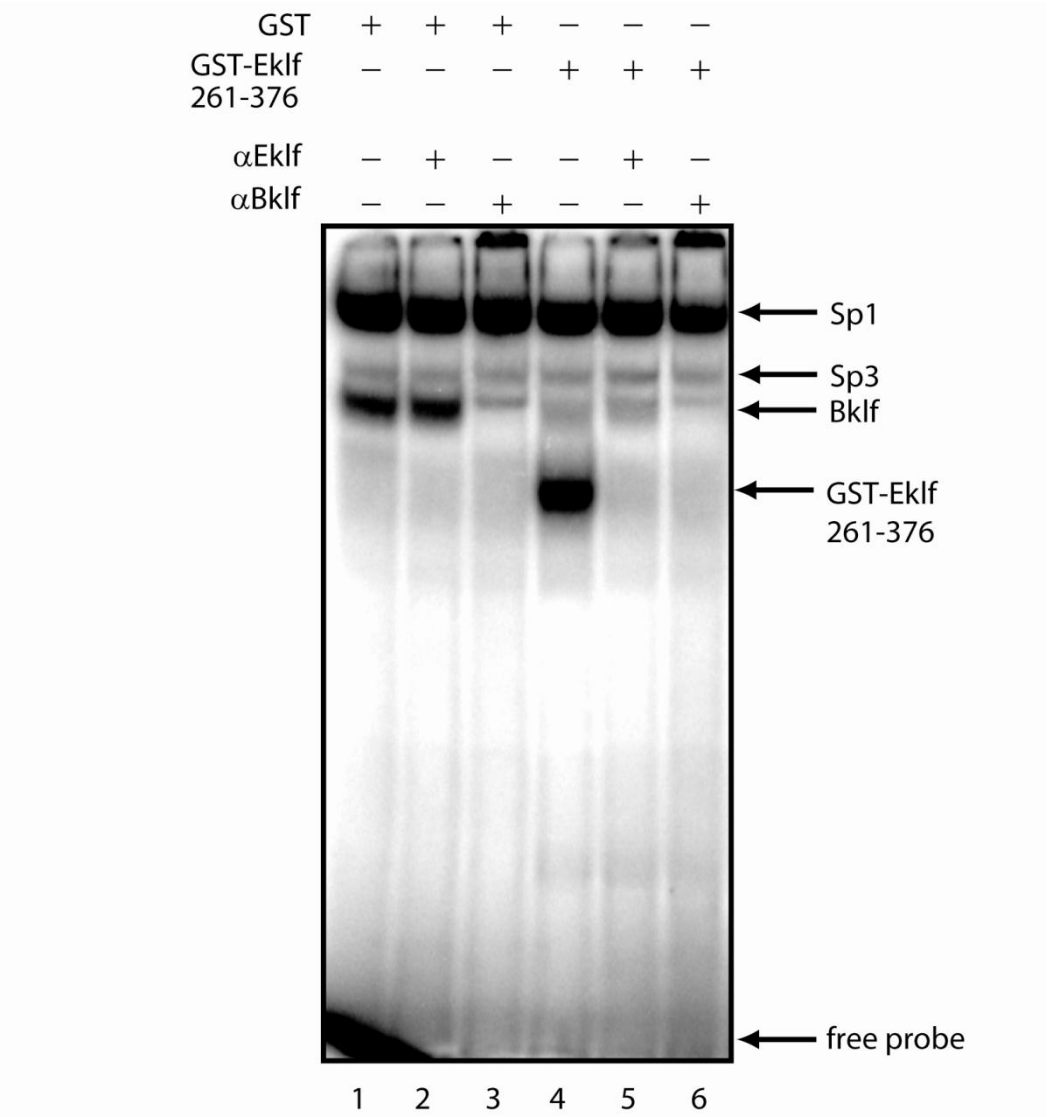


Figure 4.2 Expression of GST-Eklf 261-376 in MEL cells substantially reduces the DNA-binding activity of endogenous Bklf.

Nuclear extracts were prepared from MEL cells stably transfected with pEF-IRES-P.GST (lanes 1-3) or pEF-IRES-P.GST-Eklf 261-376 (lanes 4-6). The nuclear extracts were then incubated with a radiolabeled probe containing the *ß-major globin* promoter CACCC box, in the presence or absence of Bklf or Eklf antibody, as indicated. The reaction mixtures were then separated in a non-denaturing polyacrylamide gel in an EMSA. The interaction of GST-Eklf 261-376 with the DNA probe is shown by a shifted band, which is confirmed by antibody supershifting. Results shown are representative of three independent experiments.

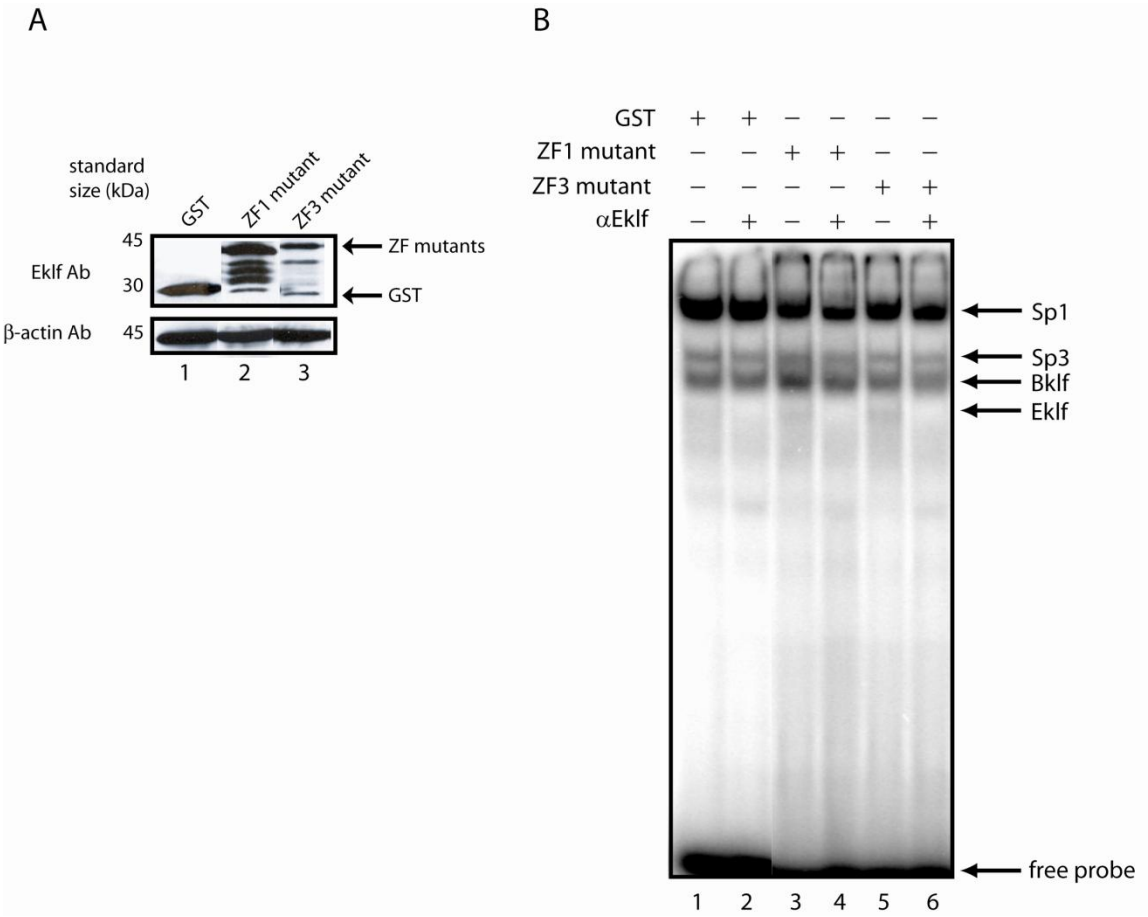


Figure 4.3 Point mutations in the ZFs of Eklf show little effect on endogenous Bklf DNA-binding activity

MEL cell lines were stably transfected with pEFIRES-P.GST, pEFIRES-P.GST-Eklf 261-376 C295A (ZF1 mutant) and pEFIRES-P.GST-Eklf 261-376 C355A (ZF3 mutant) and maintained under puromycin selection (in 1 µg/mL puromycin). **A.** Western blot showing GST (lane 1), ZF1 mutant (lane 2) and ZF3 (lane 3) expression in the nuclear extracts. **B.** EMSA experiment showing the MEL nuclear extracts incubated with a radiolabeled probe containing the *ß-major globin* promoter CACCC box, in the presence or absence of the Eklf antibody, as indicated. Results shown are representative of four independent clones.

4.4 Bklf mRNA and protein levels are substantially decreased

As described in the previous section, the MEL stable cell line expressing GST-Eklf 261-376 showed reduced endogenous Bklf DNA-binding activity. This reduction in binding could be the result of the downregulation of *Bklf* expression, caused by the GST-Eklf ZF fusion protein. In order to test this hypothesis, endogenous *Bklf* mRNA and protein levels were examined. Real-time PCR was used to measure endogenous *Bklf* transcript levels and the result is shown in Figure 4.4 (Funnell, 2008). This experiment demonstrated that *Bklf* mRNA levels are reduced approximately 5-fold in the MEL stable cell line expressing GST-Eklf 261-376 relative to untransfected MEL cells, or cells stably expressing GST (Figure 4.4). In addition, Western blot analysis of nuclear extracts showed that Bklf protein levels are reduced considerably in the cells expressing GST-Eklf 261-376 (Figure 4.5; compare lanes 1 and 2 to lane 3). Furthermore, *ß-major globin* transcripts were found to be decreased approximately 15-fold, relative to both untransfected MEL cells and cells stably expressing GST (Funnell, 2008). Taken together, these results show that the observed reduction in DNA-binding is due to the downregulation of *Bklf* gene expression. Moreover, the results suggest that the GST-Eklf ZF fusion may be directly affecting the expression levels of Eklf target genes by behaving as a dominant negative mutant, since the *Bklf* and *ß-major globin* genes are directly activated by Eklf (Bieker and Southwood, 1995; Funnell et al., 2007). Since Eklf plays a critical role in erythropoiesis, we hypothesise that overexpression of its ZFs will lead to a block in erythroid differentiation. This will be explored in the next section of this chapter.

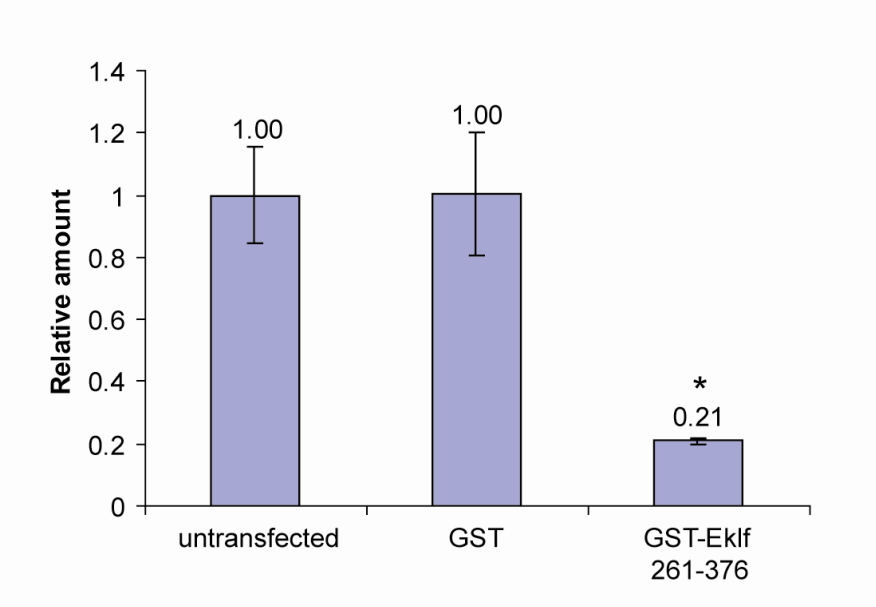


Figure 4.4 Expression of GST-Eklf 261-376 in MEL cells reduces Bklf mRNA levels.

Analysis of *Bklf* mRNA was performed by quantitative real-time RT-PCR using cDNA of untransfected and cultures stably transfected with pEF-IRES-P.GST or pEF-IRES-P.GST-Eklf 261-376. *Bklf* mRNA levels were normalised to 18S rRNA levels. Results shown are representative of four independent experiments. Data are presented as mean ± SEM, n=4. *$p<0.05$ (paired Student's t-test). Experimental data reproduced with permission from A. Funnell (Funnell, 2008).

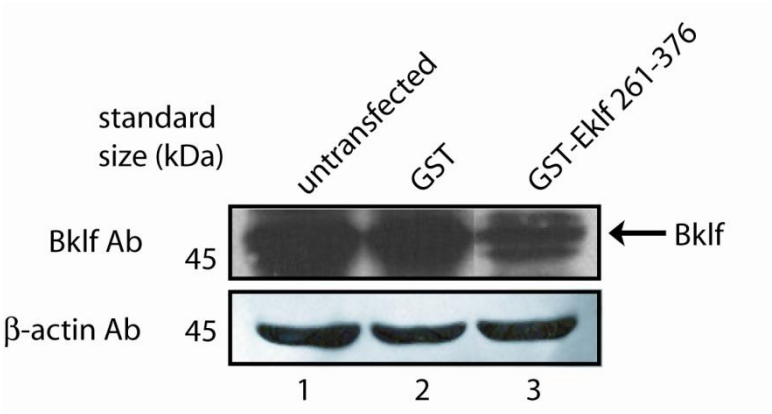


Figure 4.5 Expression of GST-Eklf 261-376 in MEL cells reduces Bklf protein levels.

Bklf protein levels were monitored in the nuclear extracts of untransfected (lane 1) and cultures stably transfected with pEF-IRES-P.GST (lane 2) or pEF-IRES-P.GST-EKLF 261-376 (lane 3). Proteins were detected in a Western blot probed with the Bklf or β-actin antibodies.

4.5 MEL: GST-Eklf 261-376 cell line fails to differentiate

In order to test the hypothesis that the GST-Eklf ZF fusion is behaving as a dominant negative mutant, the stable MEL cell line was chemically induced to undergo erythroid differentiation. Among the numerous known potent inducers of differentiation of MEL cells, dimethyl sulfoxide (DMSO), a polar-planar compound, was used as the chemical inducer in this experiment (Friend et al., 1971). The stable MEL cell line expressing GST-Eklf 261-376 was cultured in media containing DMSO. Cells that undergo erythroid differentiation turn red in colour, due to the production of haemoglobin, and this is noticeable in the colour of the cell pellet. Figure 4.6 shows that while untransfected MEL cells successfully undergo erythroid differentiation, this behaviour was absent in the GST-Eklf 261-376 cell line. Altogether, the results described in the previous sections of this chapter suggest that the GST-Eklf ZF fusion is behaving as a dominant negative

mutant (of Eklf) within MEL cells by interfering with the expression of its downstream target genes and as a result, blocking erythroid differentiation.

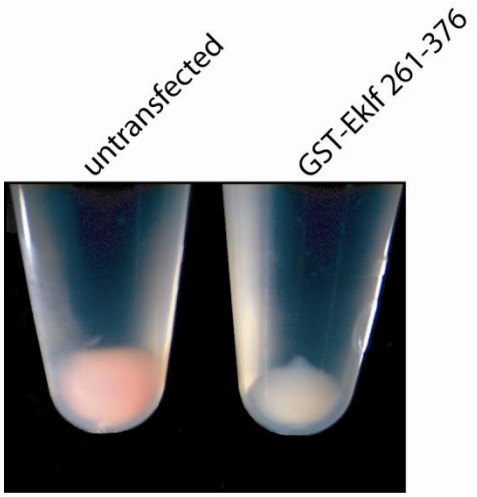


Figure 4.6 MEL cells expressing GST-Eklf 261-376 fail to differentiate into erythrocytes upon chemical induction with dimethyl sulfoxide (DMSO).

Cells, seeded at a concentration of 2 $\times 10^5$ cells/mL, were incubated for 48 h in DMEM containing 1.8% DMSO. Results shown are representative of three independent experiments.

4.6 The ZFs of Lklf also appear to behave as a dominant negative mutant

Due to the results observed in the MEL stable cell line expressing GST-Eklf 261-376 (Sections 4.3-4.5), additional stable cells lines were constructed, expressing the ZF regions of various members of the Klf family. More specifically, GST-tagged ZF regions of Lklf, Bklf and Klf8 were expressed in these cell lines. The purpose of such an experiment was to determine whether any specificity exists in the recognition of the target sequences of the ZFs. As these regions are highly conserved (67.9% identity between the ZF regions of Eklf, Lklf, Bklf and Klf8), it was predicted that each of these cell lines would behave in a similar manner to the one expressing the GST-Eklf ZF fusion. In other words, it was hypothesised that a GST fusion of the ZF region of each member of the Klf family has the capacity to function as a dominant negative mutant. However, differing effects were observed among the different family members. In the case of Klf8, no positive clones were obtained, suggesting that expression of the GST-tagged ZFs results in a lethal phenotype. In addition, although clones expressing the GST-tagged ZFs of Bklf were obtained, they were obtained with difficulty following several electroporation attempts. Furthermore, the majority of protein was found to be degraded in the positive clones and EMSA results did not show a noticeable decrease in endogenous Bklf DNA-binding activity (data not shown). However, unlike the ZFs of Klf8 and Bklf, cells expressing the GST-Lklf ZF fusion were viable. In this section, the results obtained from the stable cell line expressing the ZFs of Lklf are presented.

4.6.1 Generation of MEL cells stably expressing GST-Lklf 253-354

The stable cell line expressing the GST tag fused to the ZFs of Lklf was constructed in the similar manner as that described in Section 4.2. In brief, the GST tag and the region encoding amino acids 253-354 of *Lklf* were sequentially sub-cloned into the pEFIRES-P vector to produce a GST-Lklf 253-354 fusion protein. Following electroporation of MEL cells with the construct and selection for puromycin resistance, clones were picked and screened for recombinant protein expression. Figure 4.7 presents a Western blot showing the level of expressed protein (lane 2).

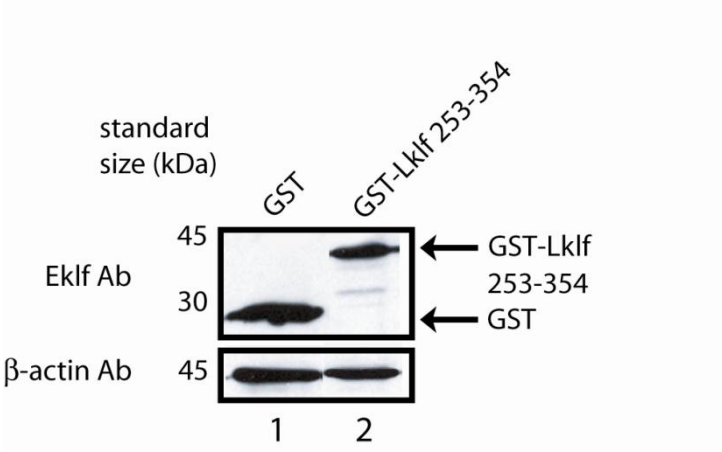


Figure 4.7 Expression of GST-Lklf 253-354 in MEL nuclear extracts.

MEL cells were electroporated with either 100 μg pEF-IRES-P.GST or pEF-IRES-P.GST-Lklf 253-354. To generate the stable cell lines, puromycin was added to a final concentration of 1 μg/mL, 48h after transfection. Individual puromycin resistant clones were picked, expanded and then harvested. Proteins were detected in a Western blot probed with the Eklf or β-actin antibodies. Results shown are representative of four independent experiments.

4.6.2 Bklf DNA-binding activity is reduced also

Using nuclear extracts obtained from the MEL stable cell line expressing GST-Lklf 253-354, an EMSA experiment was undertaken, in the same manner as outlined in Section 4.3. As can be seen in Figure 4.8, expression of the GST-Lklf ZF fusion results in a strong band shift (lane 3), similar to that observed for the GST-Eklf ZF fusion (Figure 4.8, lane 4). The identity of this band was confirmed with antibody supershifting (Figure 4.8, lane 4), as the Eklf polyclonal antibody recognises the GST tag of the expressed protein (the antibody was raised against GST fused to the region of Eklf excluding the ZFs; namely GST-Eklf 1-114). Once again, endogenous Bklf DNA-binding activity was found to be reduced (Figure 4.8; compare lanes 3 and 4 to lanes 1 and 2). This observation suggests that the ZFs of Lklf are behaving in a similar fashion to the ZFs of Eklf. In other words, this result shows, that like the GST-Eklf ZF fusion, the GST-Lklf ZF fusion appears to behave as a dominant negative mutant (of Eklf) also.

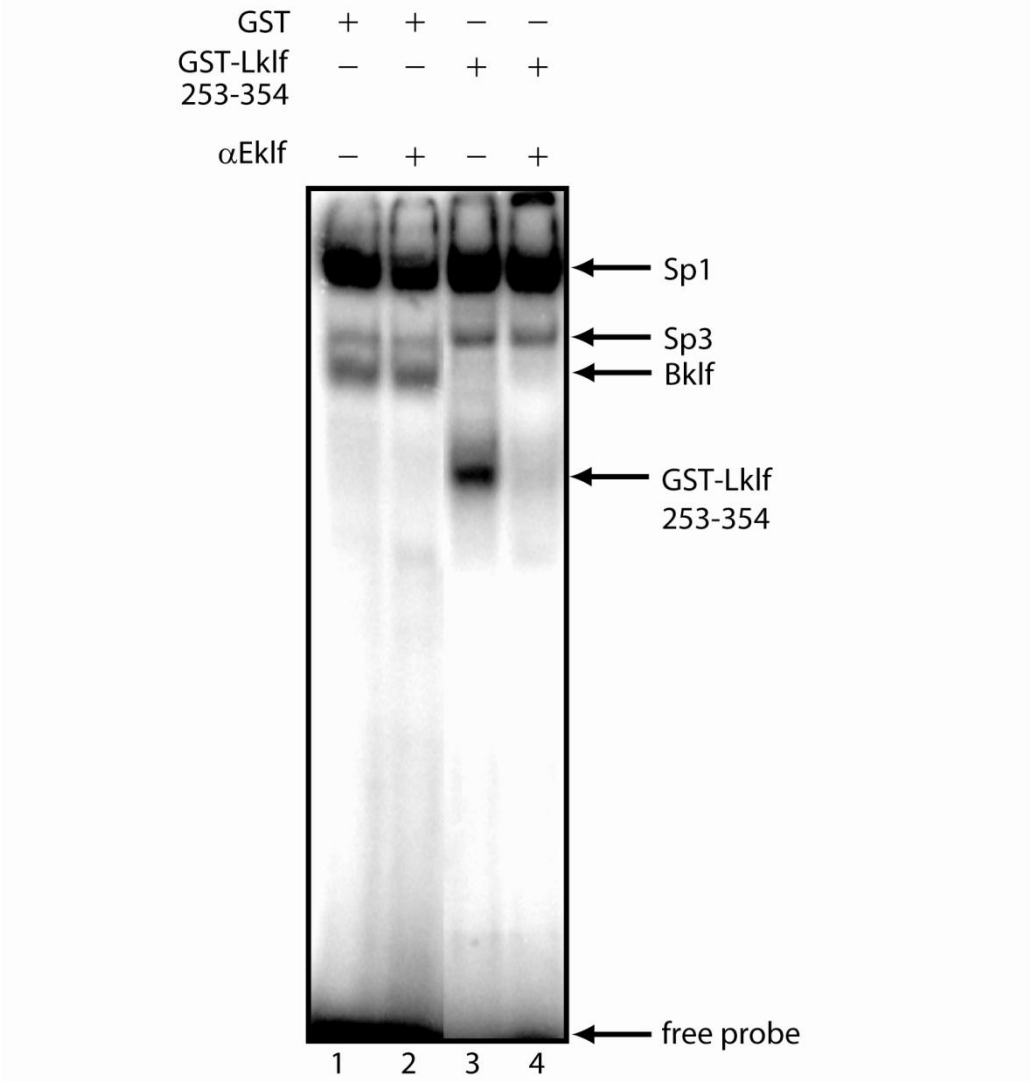


Figure 4.8 Expression of GST-Lklf 253-354 in MEL cells reduces endogenous Bklf DNA-binding activity.

Nuclear extracts were prepared from MEL cells stably transfected with pEF-IRES-P.GST (lanes 1-2) or pEF-IRES-P.GST-Lklf 253-354 (lanes 3-4). The nuclear extracts were then incubated with a radiolabeled probe containing the *ß-major globin* promoter CACCC box, in the presence or absence of Eklf antibody, as indicated. The reaction mixtures were then separated in a non-denaturing polyacrylamide gel in an EMSA. The interaction of GST-Lklf 253-354 with the DNA probe is shown by a shifted band, which is confirmed by antibody supershifting. Results shown are representative of four independent experiments.

4.7 Effects of the dominant negative mutant are diminished following multiple passages of the cell line

As demonstrated in the previous sections of this chapter, the GST-Eklf ZF fusion behaves in a manner characteristic of a dominant negative mutant. However, the effects observed appeared to diminish following multiple passages of the stable cell line, at times after as few as five passages. More specifically, after numerous passages, the DNA-binding activity observed by the recombinant GST-Eklf 261-376 protein was diminished and Bklf DNA-binding activity was restored (Figure 4.9A, lanes 1-4). Examination of the protein levels revealed that in the nuclear extracts where DNA-binding was absent, GST-Eklf 261-376 expression was still present at comparable levels (Figure 4.9B; compare lanes 1 and 2). This observation suggests that a post-translational modification of the ZFs may inhibit its DNA-binding and dominant negative capabilities. Therefore, we hypothesised that phosphorylation of the ZFs of Eklf may abolish its DNA-binding activity, as phosphorylation of the threonine residue within the conserved TGEKP linker between C_2H_2 ZFs is known to have this effect (Dovat et al., 2002). In order to test this hypothesis, a phosphatase assay was performed. It is suspected that if phosphorylation of the ZFs was responsible for the loss of the DNA-binding activity, then removal of the inhibitory phosphate groups should restore this activity. Thus, nuclear extracts obtained from cells passaged numerous times were incubated with λ protein phophatase to observe whether DNA-binding can be restored. This enzyme is a Mn^{2+}-dependent dual specificity phosphatase that releases phosphate groups from phosphorylated serine, threonine and tyrosine residues in proteins. The results from this assay are shown in Figure 4.9A.

It can be observed that phosphatase treatment of the 'old extracts' (referring to nuclear extracts obtained from cells passaged multiple times) does not restore DNA-binding activity (Figure 4.9A; lanes 2 and 4). Therefore, it appears that phosphorylation is not causing the loss of DNA-binding observed for the dominant negative mutant and that an alternative post-translational modification may be responsible.

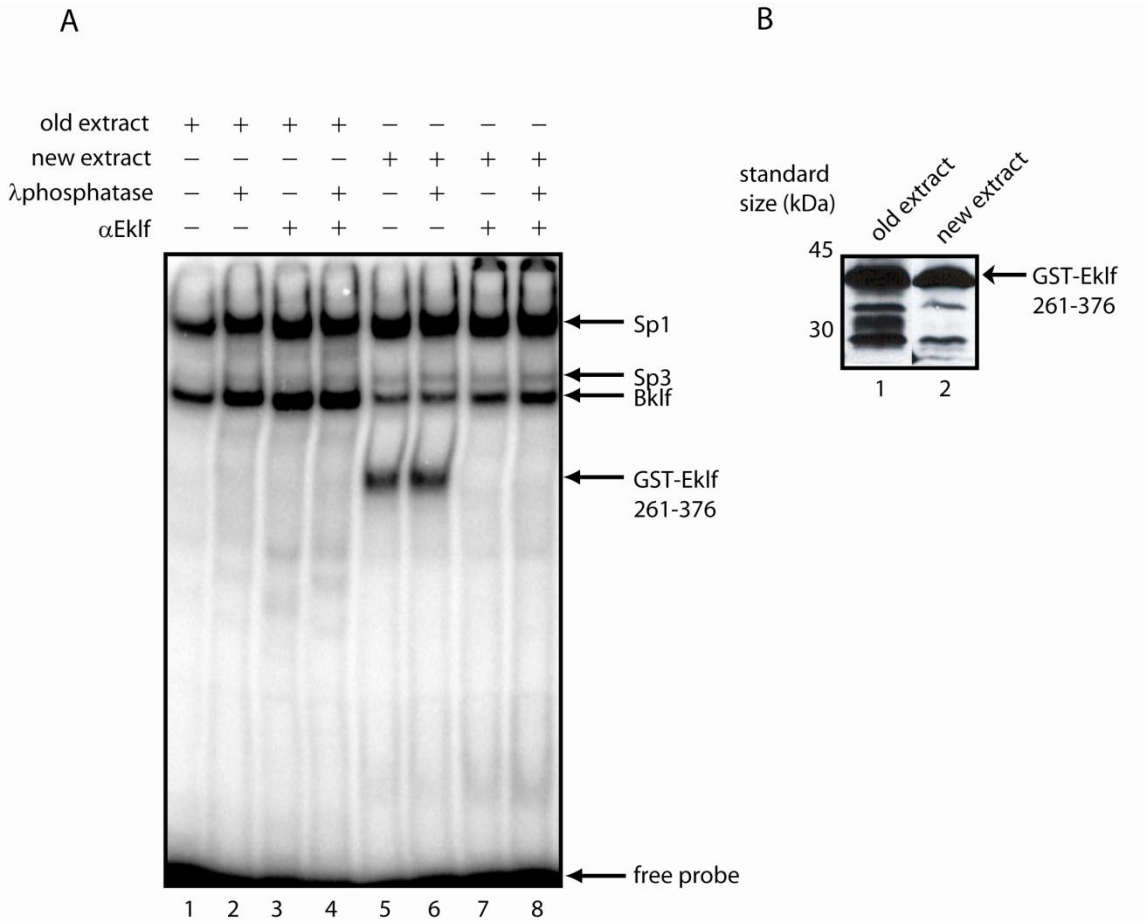


Figure 4.9 DNA-binding activity of GST-Eklf 261-376 does not appear to be affected by phosphorylation.

A. DNA-binding activities were compared by EMSA in the absence (lanes 1, 3, 5 and 7) and presence (lanes 2, 4, 6 and 8) of λ protein phosphatase. Nuclear extracts were incubated for 30 min with or without λphophatase. Following phosphatase treatment, extracts were incubated with a radiolabeled probe containing the *ß-major globin* promoter CACCC box, in the presence (lanes 3-4 and 7-8) or absence (lanes 1-2 and 5-6) of Eklf antibody. The reaction mixtures were then separated in a non-denaturing polyacrylamide gel. The interaction of GST-Eklf 261-376 with the DNA probe is shown by a shifted band, which is confirmed by antibody supershifting. **B.** Western blot showing the presence of GST-Eklf 261-376 protein in both old (lane 1) and new (lane 2) nuclear extracts. The western blot was probed with the Eklf antibody. 'Old extract' refers to the nuclear extract obtained from cells passaged multiple times and 'new extract' refers to nuclear extracts obtained from cells passaged only once.

4.8 Discussion

In this chapter we made the inadvertent discovery that the stable MEL cell line expressing the GST-Eklf ZF fusion behaves as a dominant negative mutant. In this cell line, a drastic reduction in the expression of Bklf and β-globin, two direct downstream target genes of Eklf, was observed. In addition, erythroid differentiation was found to be suppressed in this cell line, an observation supported by the fact that Eklf knockout mice die from severe anaemia due to a block in erythroid differentiation (Nuez et al., 1995; Perkins et al., 1995). Furthermore, these effects were absent in MEL cell lines expressing ZFs containing a single point mutation in one of the fingers. Moreover, these results are consistent with the observation that *ß-major globin* levels are decreased in cells expressing the ZF region of Eklf artificially fused to either the repression domain from the Drosophila engrailed protein (ENG) or the Sin 3a/Histone deacetylase interaction domain (HID) (Manwani et al., 2007). Taken together, these results suggest that expression of the GST-Eklf ZF fusion in MEL cells leads to a behaviour characteristic of a dominant negative mutant of this protein.

Since the ZF region is highly conserved between members of the Klf family (Figure 4.10; 67.9% identity between the fingers of Eklf, Lklf, Bklf and Klf8), we sought to investigate whether the aforementioned observations were due to the ZF region acting directly on Eklf in isolation, or whether the effects were more widespread, involving effects on the targets of multiple (or all) members of the Klf family. In other words, we wanted to determine whether any specificity was being exhibited by the ZFs of Eklf. In order to test this, additional stable cell lines expressing the ZFs of Lklf, Bklf and Klf8 were

produced to observe whether they behave in a similar manner to the ZF domain of Eklf. Of the four stable cell lines examined in this study, only the one expressing the GST fusion protein of the ZF region of Lklf (GST-Lklf 253-354) behaved in a similar fashion to the dominant negative mutant of Eklf. This result is not surprising, as phylogenetic analysis demonstrates that, of the Klfs examined, the ZFs of Lklf are the most closely related to those of Eklf (Figure 4.11). Altogether these results suggest that the effects observed by the ZF domain are not as widespread as anticipated and that some degree of specificity is involved, despite the highly conserved nature of the ZFs among the members of the Klf family.

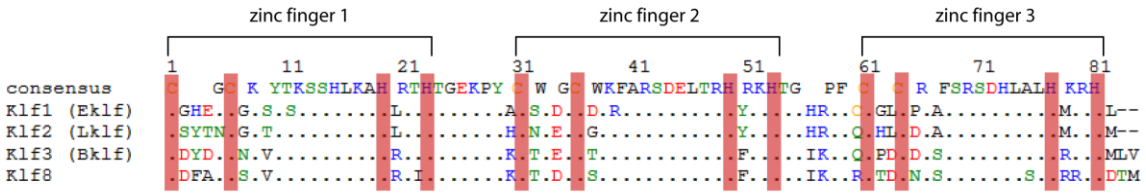


Figure 4.10 Multiple alignment of the Krüppel-like factors used in this investigation.

The ClustalW program (Thompson et al., 1994), used under the BioManager (by ANGIS) interface (http://www.angis.org.au), was used to align the ZF regions of the Klf proteins used in this study. Each protein contains three ZFs at its C-terminus. The cysteine and histidine residues involved in the co-ordination of zinc are shown in red. Amino acids are colour coded according to the chemical properties of their side chains: C, orange (sulphur); D and E, red (acidic); K, R and H, blue (basic); A, P, F, W, G, L, M, V and I, black (aliphatic); S, T, Y, N and Q, green (amide and hydroxyl). The consensus sequence is shown above the alignment. Numbers indicate amino acid residue positions within the ZF region. Shown are amino acid residues 295-376 of Eklf, residues 273-354 of Lklf, residues 261-344 of Bklf and residues 272-355 of Klf8.

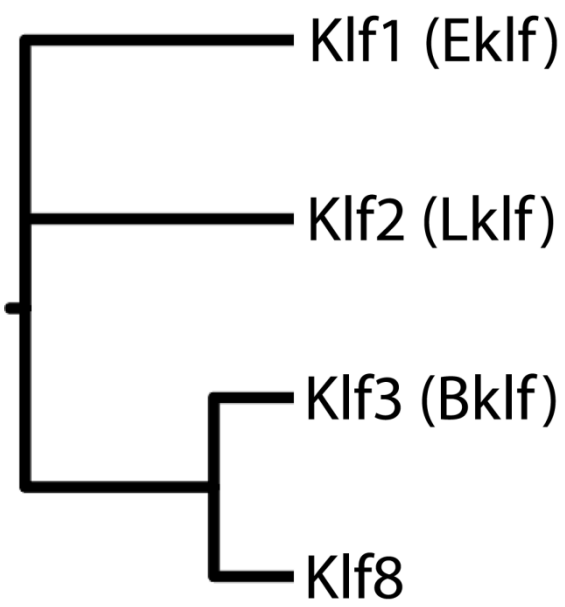


Figure 4.11 Phylogenetic relationship of the ZF domains among members of the Klf family examined in this study.

The ClustalW/Protml/DrawGram programs (Adachi and Hasegawa, 1996; Felsenstein, 1989; Thompson et al., 1994), used under the BioManager (by ANGIS) interface (http://www.angis.org.au), were used to align the ZF domains of murine Klf proteins and determine their relatedness. Also indicated are alternative names used in the literature for the same proteins. GenBank™ accession numbers for protein sequences used for this analysis are as follows: Klf1 (Eklf), NP_034765; Klf2 (Lklf), NP_032478; Klf3 (Bklf), NP_032479; Klf8, NP_776141.

Following multiple passages of this stable cell line, the effect of the dominant negative was noticeably diminished, despite expression of the GST-Eklf 261-376 protein being sustained at similar levels. The most likely explanation for this observation is that a post-translational modification was responsible for the loss of DNA-binding and dominant negative activity. We suspected that phosphorylation may be responsible, as

phosphorylation of the conserved TGEKP amino acid linkers between the ZFs of some C_2H_2 proteins, such as Sp1 and Ikaros, abolishes their DNA-binding activity (Dovat et al., 2002). The results presented in this chapter suggest that a novel modification other than phosphorylation may be responsible for the loss of DNA-binding activity observed, such as sumoylation or acetylation. Eklf has been reported to possess these modifications (Siatecka et al., 2007; Zhang and Bieker, 1998). For instance, two acetylation sites exist within the Eklf 261-376 region, namely K288 and K302 (Ouyang et al., 1998; Zhang et al., 2001). Nevertheless, further work is necessary in order to confirm these findings and to elucidate the mechanism behind the observed loss in DNA-binding. Use of mass spectrometry may allow us to determine how the recombinant protein is being modified and to identify the modified amino acid residues.

Finally, due to the profound effects of this recombinant protein on the physiological conditions of the cells, its appropriateness for use in *in vivo* RNA-binding studies, as initially intended, is questionable. In addition, loss of the dominant negative effect following multiple passages of the cell line in question may prove problematic if applications involving its use require the expansion or passaging of the cells for experiments. Nonetheless, the GST-Eklf ZF fusion protein may serve as a powerful tool with many potential applications; for instance, in studies that require the use of a cell line with a knockdown/knockout of Eklf, or for the identification of proteins that may interact with the ZFs of Eklf through the use of mass spectrometry.

Chapter 5 – The ZF domain of Eklf induces a megakaryocyte morphology in K562 cells

5.1 Introduction

In the previous chapter we developed a dominant negative mutant of Eklf (GST-Eklf 261-376) in MEL cells. In order to determine whether Eklf plays an important role in lineage decisions, the effects of this dominant negative mutant on a multipotent haematopoietic cell line were examined. Therefore, a stable cell line expressing GST-Eklf 261-376 was generated using human K562 cells. This pluripotent cell line was derived from a chronic myeloid leukaemia (CML) patient in blast crisis (Klein et al., 1976) and can be chemically induced to undergo either erythroid or megakaryocytic differentiation (Tsiftsoglou et al., 2003), making it another model for the study of erythropoiesis.

In this chapter we describe another result, where stable expression of the GST-Eklf ZF fusion in the K562 cell line results in a noticeable change in the phenotype of these cells. Therefore, we set out to characterise this observation in more detail and the results are described herein.

5.2 GST-Eklf 261-376 overexpression induces a megakaryocyte morphology

K562 cells stably expressing GST-Eklf 261-376 were generated in the same manner as outlined in Section 4.2, with the exception that expressing clones were selected in media containing 0.5 µg/mL puromycin, approximately 48 hours post-electroporation. Surprisingly, it was observed that puromycin-resistant clones positive for GST-Eklf 261-376 expression displayed an enlarged morphology, characteristic of megakaryocytes, with at least 60% of the cells adhering to the plate in a monolayer (Figure 5.1; panel 3). In contrast, this phenomenon was absent in both untransfected cells and cells stably expressing the GST tag; the cells were found to be relatively smaller in size and were non-adherent, floating in suspension (Figure 5.1; panels 1 and 2). This result demonstrates that expression of the ZFs of Eklf in K562 cells has a more profound effect than in MEL cells by driving megakaryocyte differentiation, without the requirement for chemical induction. In the next section we aim to confirm this finding at the molecular level by measuring the expression levels of *friend leukaemia integration 1* (*fli-1*), a member of the Ets family of transcription factors and a key driver of megakaryopoiesis (Athanasiou et al., 1996; Bastian et al., 1999; Jackers et al., 2004).

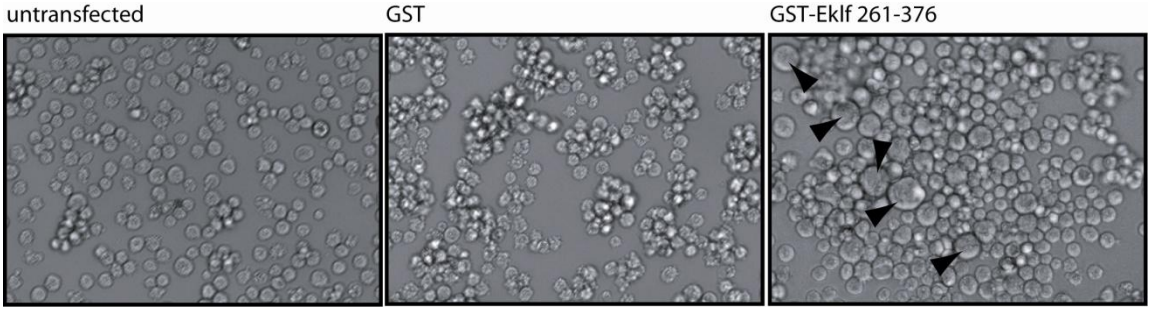


Figure 5.1 Expression of GST-Eklf 261-376 in K562 cells results in an increase in the number of cells with megakaryocyte morphology.

Phase contrast microscopy images show a noticeable increase in the number of cells with megakaryocyte morphology (some indicated by arrowheads) in cells stably transfected with pEF-IRES-P.GST-Eklf 261-376 (panel 3), but not in untransfected cells (panel 1), or cells transfected with pEF-IRES-P.GST (panel 2). Results shown are representative of three independent experiments. Original magnification, ×10.

5.3 GST-Eklf 261-376 expression increases *fli-1* transcript levels

In order to confirm at the molecular level that the stable K562 cell line expressing GST-Eklf 261-376 is undergoing megakaryopoiesis, *fli-1* expression levels were measured in these cells. To achieve this, real-time PCR was employed to quantitate *fli-1* mRNA transcript levels. As shown in Figure 5.2, *fli-1* levels are significantly increased in K562 cells stably expressing GST-Eklf 261-376, approximately 4-fold higher than untransfected cells and approximately 20-fold higher than cells stably expressing the GST tag. This marked increase was evident for at least four independent clones. Therefore, this result supports the morphological observations described in the previous section and suggests that Eklf may be responsible for repressing *fli-1* expression, either directly or indirectly via Bklf, as Bklf is a potent transcriptional repressor that is directly activated by Eklf (Funnell et al., 2007). Thus, we measured the level of *fli-1* transcripts in Bklf knockout mice to determine whether Bklf regulates *fli-1* expression.

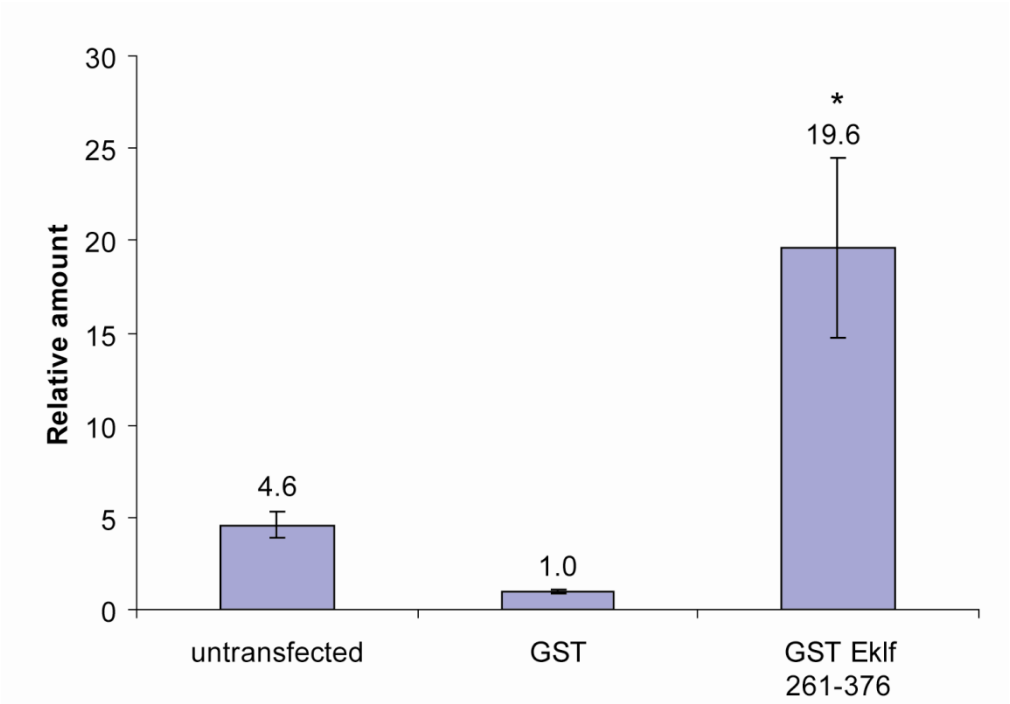


Figure 5.2 Fli-1 mRNA levels are increased in K562 cells stably expressing GST-Eklf 261-376.

Analysis of *fli-1* transcripts was performed by quantitative real-time RT-PCR using cDNA of untransfected and cultures stably transfected with pEF-IRES-P.GST or pEF-IRES-P.GST-Eklf 261-376. *Fli-1* mRNA levels were normalised to 18S rRNA levels. Results shown are representative of four independent experiments. Data are presented as mean ± SEM, n=4. *$p<0.05$ (paired Student's t-test).

5.4 Bklf knockout mice have modestly elevated *fli-1* transcript levels

Having established that *fli-1* expression is increased in K562 cells expressing the dominant negative mutant of Eklf, we wanted to determine whether this increase was attributed to a decrease in Bklf levels. To achieve this, we again used real-time PCR to measure *fli-1* mRNA transcript levels in the bone marrow cells of *Bklf* knockout mice and the results are shown in Figure 5.3. As anticipated, *fli-1* levels are significantly increased in *Bklf* knockout mice, with transcript levels 2.2-fold greater than that observed in wild-type mice. Furthermore, an increase was also observed in *Bklf* heterozygous mice. Therefore, these results support the hypothesis that Bklf may be involved in directly regulating *fli-1* gene expression.

Furthermore, an analysis of the *fli-1* promoter was undertaken, where approximately 600 bp upstream of the translational start site was scanned for potential Klf DNA-binding sites, pertaining to the consensus sequence NCNCNCCCN. Several such sites were found in this region and are indicated in Figure 5.4. Taken together, real-time PCR data of *fli-1* levels in the bone marrow of *Bklf* knockout mice and analysis of the promoter region of *fli-1* suggest that Bklf may be directly responsible for the regulation of *fli-1* expression. Thus, we sought to determine whether Bklf binds directly to the *fli-1* promoter using Chromatin Immunoprecipitation (ChIP), which will be described in the subsequent section of this chapter.

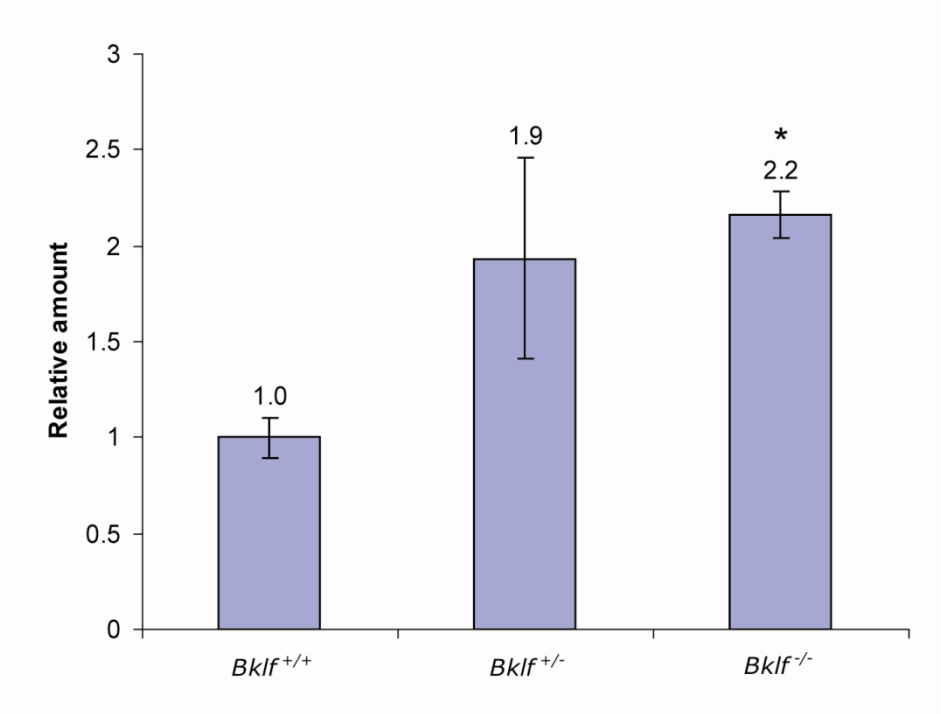


Figure 5.3 Fli-1 mRNA levels are increased in the bone marrow of Bklf knockout mice.

Analysis of *fli-1* mRNA was performed by quantitative real-time RT-PCR using cDNA synthesised from RNA extracted from the bone marrow cells of *Bklf*$^{+/+}$, *Bklf*$^{+/-}$ and *Bklf*$^{-/-}$ mice. *Fli-1* mRNA levels were normalised to 18S rRNA levels. Results shown are representative of three independent experiments. Data are presented as mean ± SEM, n=3. *p<0.05 (paired Student's t-test).

Figure 5.4 The murine fli-1 promoter showing putative Klf binding sites.

The two major transcriptional initiation sites (reported by Starck et al, 1999) are marked with arrowheads. Putative Klf binding sites, pertaining to the consensus Klf binding sequence 5'-NCNCNCCCN-3' (or its reverse complement) are underlined.

5.5 Eklf and Bklf were not detected at the *fli-1* promoter

In the previous section it was demonstrated that *fli-1* expression is modestly increased in *Bklf* knockout mice and analysis of the *fli-1* promoter showed the presence of putative Klf DNA-binding sites (Figure 5.4). Thus, these observations led us to hypothesise that Bklf directly regulates *fli-1* gene expression by binding to its promoter and recruiting a repression complex, as this protein is known to function as a transcriptional repressor (Turner and Crossley, 1998). In order to test our hypothesis, preliminary ChIP experiments were performed using DNA extracted from the foetal livers of E14.5 mice and a Bklf antibody. Following elution and extraction, immunoprecipitated DNA was analysed by real-time PCR. The primers used for amplification were designed to be in close proximity to the putative Klf DNA-binding sites, with the amplicon spanning the region -252 to -174 (Figure 5.4). In contrast to our prediction, Bklf was not detected at the *fli-1* promoter (data not shown).

This unexpected result prompted us to examine whether Eklf directly represses *fli-1* expression, as Eklf has been reported to function as a repressor (Chen and Bieker, 2004; Siatecka et al., 2007), in addition to its well-documented activator function. Thus, preliminary ChIP experiments were conducted with an Eklf antibody and real-time PCR was used to analyse the immunoprecipitated DNA with the aforementioned primer set. Once again, no enrichment was detected for Eklf at the *fli-1* promoter (data not shown). Therefore, the results from these preliminary experiments did not provide evidence that either Bklf and/or Eklf directly bind to the *fli-1* promoter. This raises the possibility that Eklf and Bklf may not responsible for regulating *fli-1*

expression and that its regulation may occur via some other indirect mechanism that is yet to be determined. Another possibility is the lack of success of ChIP in this system. In any case, further work is necessary to completely exclude the possibility that neither of these two Klf proteins directly regulate *fli-1* expression, for example through the use of primer sets at multiple locations across the *fli-1* promoter and the inclusion of robust positive controls.

5.7 Discussion

In this chapter we have constructed a human stable cell line that expresses the dominant negative mutant of Eklf, which results in the transformation of these cells to possess megakaryocyte morphology, without the need for phorbol esters to chemically induce differentiation. In addition, we demonstrated that the expression of *fli-1*, a key driver of megakaryopoiesis, is upregulated in these cells, which is consistent with the observed megakaryocyte phenotype. Concomitant with these findings is the upregulation of *fli-1* expression observed with the foetal liver of E13.5 *Eklf* knockout mice using high-density oligonucleotide microarrays (Frontelo et al., 2007). Furthermore, the foetal livers of *Eklf* knockout mice were found to contain a greater number of megakaryocytes than wild-type mice (Frontelo et al., 2007), which again is consistent with the aforementioned observations. Moreover, functional cross-antagonism was previously reported to occur between Eklf and fli-1, where fli-1 was found to repress the transcriptional targets of Eklf (for example, *ß-major globin* expression) and vice versa (for example, *GPIX* expression, a megakaryocyte-specific cell surface antigen that is regulated by fli-1) (Starck et al., 2003). Thus, the results described in the current and previous chapter strongly suggest that Eklf expression leads to the inhibition of megakaryopoiesis (through the repression of *fli-1*) and the promotion of erythropoiesis, which is consistent with the work described by Frontelo et al, 2007.

Also in this chapter it was demonstrated that *fli-1* transcripts were increased in the bone marrow of *Bklf* knockout mice, suggesting that *fli-1* transcription could be directly regulated by Bklf, which is known to be positively regulated by Eklf (Funnell et al., 2007). Preliminary ChIP experiments were performed in order to determine whether any Bklf presence could be detected at the *fli-1* promoter but were inconclusive (data not shown).

Chapter 6 – Discussion and Conclusions

6.1 Eklf is an RNA-binding protein

Prior to the experiments described in this thesis, Klf proteins were known to function as DNA-binding transcription factors (Miller and Bieker, 1993). In chapter 3 we demonstrated, using various independent *in vitro* assays, that Eklf possesses RNA-binding activity and is capable of binding to A and U homoribopolymers. In addition, we showed that Eklf has the capacity to interact with cellular RNA, suggesting that this RNA-binding activity may be of biological relevance and may be relevant to its role in the cytoplasm (Quadrini et al., 2008; Shyu et al., 2007), particularly if it has a role in mRNA subcellular localisation and/or the regulation of translation. Hence, we have now identified yet another RNA-binding classical ZF protein that can be added to the relatively small number of proteins that have been documented in the literature (presented in Table 1.1). This is the first report of a classical ZF protein that has a preference for strings of A or U RNA residues. It is evident that poly(A) and poly(U) may not be the natural targets of Eklf. Therefore, the true physiological targets of Eklf remain to be determined.

Eklf recognises both DNA and RNA and this dual binding capacity has been observed for a few other classical ZF proteins, namely TFIIIA, WT1 and Mok2 (Arranz et al., 1997; Caricasole et al., 1996; Romaniuk, 1985). These dual-function binding properties and

subcellular localisation suggest that Eklf plays roles in transcription as well as in the post-transcriptional regulation processes of specific genes. Due to the homology that exists in the ZF region across members of the Klf family, it is likely that other members of this family of transcription factors have the capacity to bind RNA also. Preliminary evidence suggests that this may be the case, as Bklf was found to possess RNA-binding activity also (Noelia Nunez, personal communication).

6.2 The ZF region of Eklf behaves as a dominant negative mutant

In this study we have shown that the ZF region of Eklf behaves as a dominant negative mutant when recombinantly expressed as a GST-fusion protein in haematopoietic cell lines. In chapter 4 we described the construction of the stable MEL cell line expressing GST-Eklf 261-376 and demonstrated the dominant negative activity through the reduction of *Bklf* mRNA transcript levels as well as through the block in erythroid differentiation. In addition, it was shown to cause a significant reduction in *ß-major globin* transcript levels (Alister Funnell, personal communication). These results are consistent with the work of Manwani et al, 2007, who recently re-designed the Eklf transcription factor by replacing the activation domain with a repression domain from either the *Drosophila* engrailed protein (ENG) or the Sin 3a/Histone deacetylase interaction domain (HID). Although this protein differed from our dominant negative mutant, in that the ZFs were tethered to a repression domain, the effects exhibited by both proteins were similar, as *ß-major globin* transcript levels were found to be significantly decreased in both cases. Thus, we have developed a useful tool for

experiments where a knockdown of Eklf activity is required. In addition, the dominant negative mutant may aid in proteomics studies examining the ZFs, such as the isolation and identification of protein partners, due to the facilitation of protein purification due to the presence of a tag. The proposed mechanism of the dominant negative mutant activity is illustrated in Figure 6.1A and B.

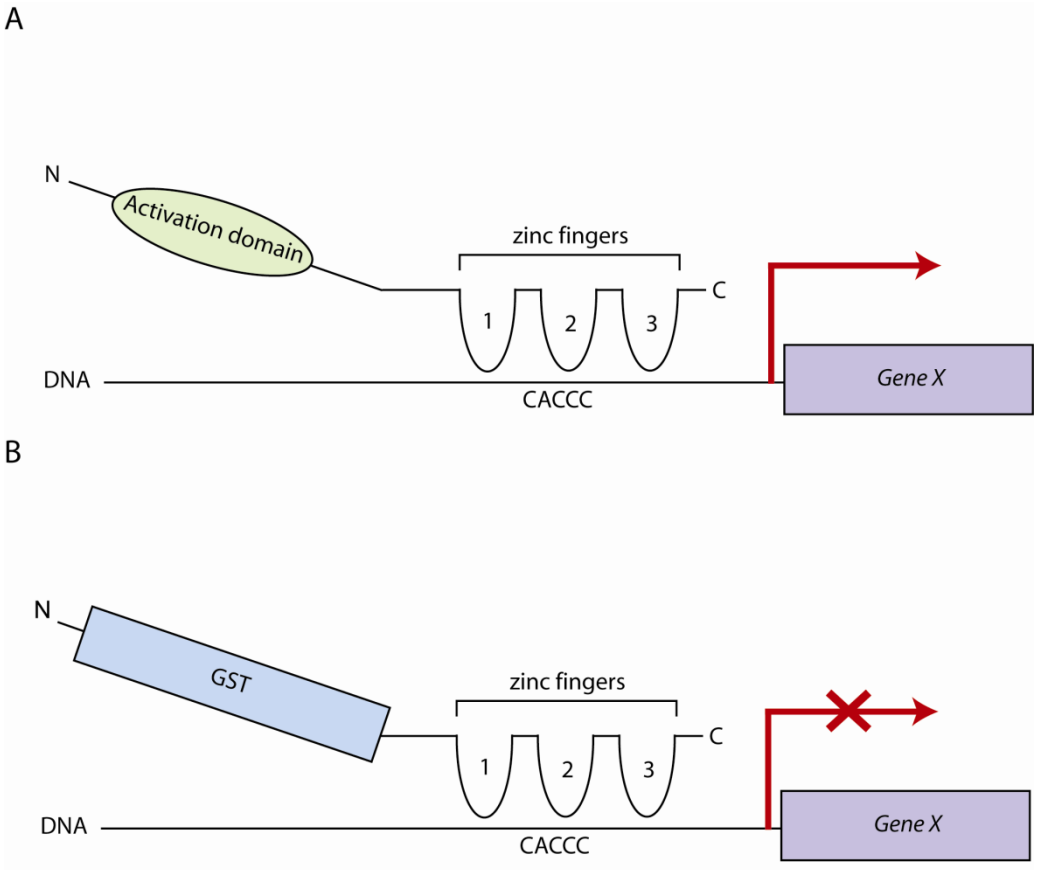


Figure 6.1 Proposed molecular mechanism of Eklf dominant negative mutant activity.

A. Eklf activation through binding of its ZFs to CACCC-box elements present within the promoter sequences of its target genes (*Gene X* - indicated by a purple box). The activation domain (light green oval) allows recruitment of the co-activator complex and the subsequent activation of gene expression. **B.** Competition by the dominant negative mutant (GST-Eklf 261-376) for the Eklf DNA-binding sites results in a marked reduction of activation, presumably due to lack of a functional activation domain, as was replaced by the GST tag (light blue box).

6.3 Eklf may play an important role in haematopoietic lineage commitment decisions

In this study we showed that Eklf can inhibit megakaryocyte development through the use of the dominant negative mutant in the multipotent haematopoietic K562 cell line. We found that stably expressing GST-Eklf 261-376 in these cells results in drastic changes in cellular morphology and an increase in both cellular adherence and *fli-1* expression. All of these changes resemble the response of K562 cells to the phorbol ester 12-*O*-tetradecanoylphorbol-13-acetate (TPA) (Athanasiou et al., 1996) and are consistent with activation of a megakaryocytic or platelet-specific response. Our results are consistent with the recent work of Frontelo et al, 2007, where a significant increase in the number of megakaryocyte colonies, cultured using the foetal liver cells of E13.5 *Eklf* knockout embryos, was observed. The promotion of erythropoiesis and inhibition of megakaryopoiesis by Eklf suggests that this transcription factor plays a critical role in lineage decisions from the MEP.

In addition, we demonstrated that *fli-1* transcript levels were modestly increased in the bone marrow cells of *Bklf* knockout mice, suggesting that *fli-1* expression may be directly regulated by the transcriptional repressor Bklf. Preliminary ChIP assays were unsuccessful at detecting direct binding of Bklf or Eklf to the *fli-1* promoter, suggesting the assay may not have worked in this system. Alternatively, Klfs may not bind to this region of the promoter and Klf DNA-binding sites may exist at other regions that were not examined in this study. In any case, further work is necessary in order to clearly

understand the mechanism of *fli-1* regulation. Our present understanding of the regulation of *fli-1* gene expression is depicted in Figure 6.2.

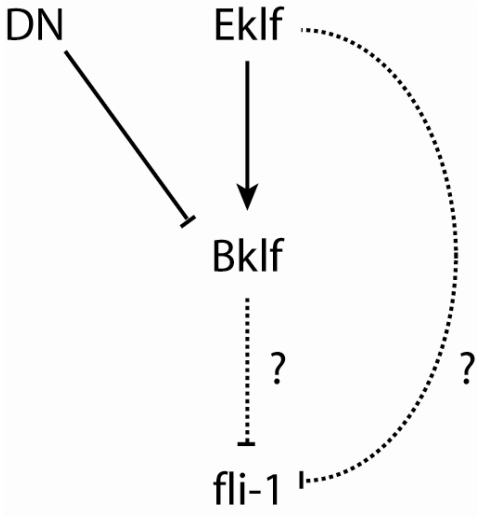


Figure 6.2 Model of fli-1 gene regulation by Klfs.

Study of the dominant negative mutant (DN) implicates the pathway as indicated. Eklf activates Bklf expression, which results in a repression of *fli-1* levels, either through Bklf or directly through Eklf. Dotted lines with a question mark indicate that further examination is necessary in order to clearly establish the mechanism of *fli-1* repression.

6.4 Future studies

The *in vivo* data presented in this study was acquired through the use of the mammalian MEL and K562 cell lines. One apparent limitation that arises from employing such cell lines is the fact that these cells are aberrant and do not possess the normal cell differentiation machinery. Therefore, findings obtained using these cells may not necessarily be representative of haematopoiesis. Further confirmation of the *in vivo* results could be obtained through the use of primary murine cells. For instance, RNA-binding experiments could be pursued further using primary cells through RNA-ChIP and analysed using microarrays. For instance, the RNA-binding forms of Eklf that lack the ability to bind DNA could be used to rescue primary murine B1.6 cells and then microarrays conducted to identify the RNA species bound.

6.5 Final summary

In summary, we have shown that Eklf, the founding member of the Klf family of transcription factors, is an RNA-binding protein using its classical ZFs. It has a preference of strings of A and U RNA residues, suggesting it may bind to ARE elements in mRNA transcripts in a cellular context. Taking into account that approximately 3% of genes in the human genome encode classical ZF proteins along with the results presented in this study, it appears that a substantial number of classical ZF proteins with RNA-binding activity remain to be identified and characterised. With the growing realisation of the importance of protein-RNA interactions in gene regulation at the post-

transcriptional level, this will undoubtedly remain a field of intense research in the years to come.

In addition, we have shown that the ZF region of Eklf behaves as a dominant negative mutant when expressed as a GST-fusion protein in haematopoietic cell lines. Furthermore, using the dominant negative mutant, we have confirmed that Eklf is critical for the promotion of erythroid differentiation and the inhibition of megakaryocyte differentiation.

Chapter 7 – References

Adachi, J., and Hasegawa, M. (1996). Programs for molecular phylogenetics based on maximum likelihood. In Computer Science monographs, Volume 28 (Tokyo, Institute of Statistical Mathematics).

Amero, S.A., Matunis, M.J., Matunis, E.L., Hockensmith, J.W., Raychaudhuri, G., and Beyer, A.L. (1993). A unique ribonucleoprotein complex assembles preferentially on ecdysone-responsive sites in Drosophila melanogaster. Mol Cell Biol *13*, 5323-5330.

Anderson, K.P., Kern, C.B., Crable, S.C., and Lingrel, J.B. (1995). Isolation of a gene encoding a functional zinc finger protein homologous to erythroid Kruppel-like factor: identification of a new multigene family. Mol Cell Biol *15*, 5957-5965.

Andreazzoli, M., De Lucchini, S., Costa, M., and Barsacchi, G. (1993). RNA binding properties and evolutionary conservation of the Xenopus multifinger protein Xfin. Nucleic Acids Res *21*, 4218-4225.

Armstrong, J.A., Bieker, J.J., and Emerson, B.M. (1998). A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLF in vitro. Cell *95*, 93-104.

Arning, S., Gruter, P., Bilbe, G., and Kramer, A. (1996). Mammalian splicing factor SF1 is encoded by variant cDNAs and binds to RNA. RNA *2*, 794-810.

Arranz, V., Harper, F., Florentin, Y., Puvion, E., Kress, M., and Ernoult-Lange, M. (1997). Human and mouse MOK2 proteins are associated with nuclear ribonucleoprotein components and bind specifically to RNA and DNA through their zinc finger domains. Mol Cell Biol *17*, 2116-2126.

Athanasiou, M., Clausen, P.A., Mavrothalassitis, G.J., Zhang, X.K., Watson, D.K., and Blair, D.G. (1996). Increased expression of the ETS-related transcription factor FLI-

1/ERGB correlates with and can induce the megakaryocytic phenotype. Cell Growth Differ *7*, 1525-1534.

Banerjee, S.S., Feinberg, M.W., Watanabe, M., Gray, S., Haspel, R.L., Denkinger, D.J., Kawahara, R., Hauner, H., and Jain, M.K. (2003). The Kruppel-like factor KLF2 inhibits peroxisome proliferator-activated receptor-gamma expression and adipogenesis. J Biol Chem *278*, 2581-2584.

Bardoni, B., Schenck, A., and Mandel, J.L. (1999). A novel RNA-binding nuclear protein that interacts with the fragile X mental retardation (FMR1) protein. Hum Mol Genet *8*, 2557-2566.

Barreau, C., Paillard, L., and Osborne, H.B. (2005). AU-rich elements and associated factors: are there unifying principles? Nucleic Acids Res *33*, 7138-7150.

Bashirullah, A., Halsell, S.R., Cooperstock, R.L., Kloc, M., Karaiskakis, A., Fisher, W.W., Fu, W., Hamilton, J.K., Etkin, L.D., and Lipshitz, H.D. (1999). Joint action of two RNA degradation pathways controls the timing of maternal transcript elimination at the midblastula transition in Drosophila melanogaster. EMBO J *18*, 2610-2620.

Bastian, L.S., Kwiatkowski, B.A., Breininger, J., Danner, S., and Roth, G. (1999). Regulation of the megakaryocytic glycoprotein IX promoter by the oncogenic Ets transcription factor Fli-1. Blood *93*, 2637-2644.

Basu, P., Morris, P.E., Haar, J.L., Wani, M.A., Lingrel, J.B., Gaensler, K.M., and Lloyd, J.A. (2005). KLF2 is essential for primitive erythropoiesis and regulates the human and murine embryonic beta-like globin genes in vivo. Blood *106*, 2566-2571.

Beerli, R.R., and Barbas, C.F., 3rd (2002). Engineering polydactyl zinc-finger transcription factors. Nat Biotechnol *20*, 135-141.

Bentley, D. (1999). Coupling RNA polymerase II transcription with pre-mRNA processing. Curr Opin Cell Biol *11*, 347-351.

Bernstein, J.A., Khodursky, A.B., Lin, P.H., Lin-Chao, S., and Cohen, S.N. (2002). Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. Proc Natl Acad Sci U S A *99*, 9697-9702.

Bieker, J.J. (2001). Kruppel-like factors: three fingers in many pies. J Biol Chem *276*, 34355-34358.

Bieker, J.J., and Southwood, C.M. (1995). The erythroid Kruppel-like factor transactivation domain is a critical component for cell-specific inducibility of a beta-globin promoter. Mol Cell Biol *15*, 852-860.

Bird, A.J., McCall, K., Kramer, M., Blankman, E., Winge, D.R., and Eide, D.J. (2003). Zinc fingers can act as Zn2+ sensors to regulate transcriptional activation domain function. EMBO J *22*, 5137-5146.

Bishop, D.C. (2001). Investigating the Ability of FHL3 to act as a co-repressor of BKLF. In Department of Biochemistry (Sydney, University of Sydney).

Bittel, D.C., Smirnova, I.V., and Andrews, G.K. (2000). Functional heterogeneity in the zinc fingers of metalloregulatory protein metal response element-binding transcription factor-1. J Biol Chem *275*, 37194-37201.

Black, A.R., Black, J.D., and Azizkhan-Clifford, J. (2001). Sp1 and kruppel-like factor family of transcription factors in cell growth regulation and cancer. J Cell Physiol *188*, 143-160.

Blackwell, T.K., and Weintraub, H. (1990). Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. Science *250*, 1104-1110.

Brenner, H.R., Witzemann, V., and Sakmann, B. (1990). Imprinting of acetylcholine receptor messenger RNA accumulation in mammalian neuromuscular synapses. Nature *344*, 544-547.

Brown, V., Small, K., Lakkis, L., Feng, Y., Gunter, C., Wilkinson, K.D., and Warren, S.T. (1998). Purified recombinant Fmrp exhibits selective RNA binding as an intrinsic property of the fragile X mental retardation protein. J Biol Chem *273*, 15521-15527.

Calnan, B.J., Tidor, B., Biancalana, S., Hudson, D., and Frankel, A.D. (1991). Arginine-mediated RNA recognition: the arginine fork. Science *252*, 1167-1171.

Carballo, E., Lai, W.S., and Blackshear, P.J. (1998). Feedback inhibition of macrophage tumor necrosis factor-alpha production by tristetraprolin. Science *281*, 1001-1005.

Caricasole, A., Duarte, A., Larsson, S.H., Hastie, N.D., Little, M., Holmes, G., Todorov, I., and Ward, A. (1996). RNA binding by the Wilms tumor suppressor zinc finger proteins. Proc Natl Acad Sci U S A *93*, 7562-7566.

Chen, C.Y., Gherzi, R., Ong, S.E., Chan, E.L., Raijmakers, R., Pruijn, G.J., Stoecklin, G., Moroni, C., Mann, M., and Karin, M. (2001). AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. Cell *107*, 451-464.

Chen, C.Y., and Shyu, A.B. (1995). AU-rich elements: characterization and importance in mRNA degradation. Trends Biochem Sci *20*, 465-470.

Chen, C.Y., Xu, N., and Shyu, A.B. (2002). Highly selective actions of HuR in antagonizing AU-rich element-mediated mRNA destabilization. Mol Cell Biol *22*, 7268-7278.

Chen, X., and Bieker, J.J. (2004). Stage-specific repression by the EKLF transcriptional activator. Mol Cell Biol *24*, 10416-10424.

Chen, X., Chu, M., and Giedroc, D.P. (1999). MRE-Binding transcription factor-1: weak zinc-binding finger domains 5 and 6 modulate the structure, affinity, and specificity of the metal-response element complex. Biochemistry *38*, 12915-12925.

Cheng, A.C., Calabro, V., and Frankel, A.D. (2001). Design of RNA-binding proteins and ligands. Curr Opin Struct Biol *11*, 478-484.

Choo, Y., and Isalan, M. (2000). Advances in zinc finger engineering. Curr Opin Struct Biol *10*, 411-416.

Crossley, M., Whitelaw, E., Perkins, A., Williams, G., Fujiwara, Y., and Orkin, S.H. (1996). Isolation and characterization of the cDNA encoding BKLF/TEF-2, a major CACCC-box-binding protein in erythroid cells and selected other cells. Mol Cell Biol *16*, 1695-1705.

Dang, D.T., Pevsner, J., and Yang, V.W. (2000). The biology of the mammalian Kruppel-like family of transcription factors. Int J Biochem Cell Biol *32*, 1103-1121.

Dang, D.T., Zhao, W., Mahatan, C.S., Geiman, D.E., and Yang, V.W. (2002). Opposing effects of Kruppel-like factor 4 (gut-enriched Kruppel-like factor) and Kruppel-like factor 5 (intestinal-enriched Kruppel-like factor) on the promoter of the Kruppel-like factor 4 gene. Nucleic Acids Res *30*, 2736-2741.

Dean, A. (2004). Chromatin remodelling and the interaction between enhancers and promoters in the beta-globin locus. Brief Funct Genomic Proteomic *2*, 344-354.

Dean, J.L., Wait, R., Mahtani, K.R., Sully, G., Clark, A.R., and Saklatvala, J. (2001). The 3' untranslated region of tumor necrosis factor alpha mRNA is a target of the mRNA-stabilizing factor HuR. Mol Cell Biol *21*, 721-730.

Dominski, Z., Erkmann, J.A., Yang, X., Sanchez, R., and Marzluff, W.F. (2002). A novel zinc finger protein is associated with U7 snRNP and interacts with the stem-loop binding protein in the histone pre-mRNP to stimulate 3'-end processing. Genes Dev *16*, 58-71.

Donze, D., Townes, T.M., and Bieker, J.J. (1995). Role of erythroid Kruppel-like factor in human gamma- to beta-globin gene switching. J Biol Chem *270*, 1955-1959.

Dovat, S., Ronni, T., Russell, D., Ferrini, R., Cobb, B.S., and Smale, S.T. (2002). A common mechanism for mitotic inactivation of C2H2 zinc finger DNA-binding domains. Genes Dev *16*, 2985-2990.

Draper, D.E. (1995). Protein-RNA recognition. Annu Rev Biochem *64*, 593-620.

Ellenberger, T.E., Brandl, C.J., Struhl, K., and Harrison, S.C. (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex. Cell *71*, 1223-1237.

Felsenstein, J. (1989). PHYLIP -- Phylogeny Inference Package (Version 3.2). Cladistics *5*, 164-166.

Feng, W.C., Southwood, C.M., and Bieker, J.J. (1994). Analyses of beta-thalassemia mutant DNA interactions with erythroid Kruppel-like factor (EKLF), an erythroid cell-specific transcription factor. J Biol Chem *269*, 1493-1500.

Fialcowitz-White, E.J., Brewer, B.Y., Ballin, J.D., Willis, C.D., Toth, E.A., and Wilson, G.M. (2007). Specific protein domains mediate cooperative assembly of HuR oligomers on AU-rich mRNA-destabilizing sequences. J Biol Chem *282*, 20948-20959.

Finerty, P.J., Jr., and Bass, B.L. (1997). A Xenopus zinc finger protein that specifically binds dsRNA and RNA-DNA hybrids. J Mol Biol *271*, 195-208.

Finerty, P.J., Jr., and Bass, B.L. (1999). Subsets of the zinc finger motifs in dsRBP-ZFa can bind double-stranded RNA. Biochemistry *38*, 4001-4007.

Fisher, C.L., and Pei, G.K. (1997). Modification of a PCR-based site-directed mutagenesis method. Biotechniques *23*, 570-571, 574.

Forrest, K.M., and Gavis, E.R. (2003). Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for nanos mRNA localization in Drosophila. Curr Biol *13*, 1159-1168.

Friend, C. (1957). Cell-free transmission in adult Swiss mice of a disease having the character of a leukemia. J Exp Med *105*, 307-318.

Friend, C., Scher, W., Holland, J.G., and Sato, T. (1971). Hemoglobin synthesis in murine virus-induced leukemic cells in vitro: stimulation of erythroid differentiation by dimethyl sulfoxide. Proc Natl Acad Sci U S A *68*, 378-382.

Friesen, W.J., and Darby, M.K. (1997). Phage display of RNA binding zinc fingers from transcription factor IIIA. J Biol Chem *272*, 10994-10997.

Frontelo, P., Manwani, D., Galdass, M., Karsunky, H., Lohmann, F., Gallagher, P.G., and Bieker, J.J. (2007). Novel role for EKLF in megakaryocyte lineage commitment. Blood *110*, 3871-3880.

Funnell, A.P. (2008). Identification of a regulatory network within the Kruppel-like factor family. In School of Molecular and Microbial Biosciences (Sydney, University of Sydney).

Funnell, A.P., Maloney, C.A., Thompson, L.J., Keys, J., Tallack, M., Perkins, A.C., and Crossley, M. (2007). Erythroid Kruppel-like factor directly activates the basic Kruppel-like factor gene in erythroid cells. Mol Cell Biol *27*, 2777-2790.

Gaston, K., and Jayaraman, P.S. (2003). Transcriptional repression in eukaryotes: repressors and repression mechanisms. Cell Mol Life Sci *60*, 721-741.

Gebauer, F., and Hentze, M.W. (2004). Molecular mechanisms of translational control. Nat Rev Mol Cell Biol *5*, 827-835.

Georgopoulos, K. (2002). Haematopoietic cell-fate decisions, chromatin regulation and ikaros. Nat Rev Immunol *2*, 162-174.

Ghaleb, A.M., Nandan, M.O., Chanchevalap, S., Dalton, W.B., Hisamuddin, I.M., and Yang, V.W. (2005). Kruppel-like factors 4 and 5: the yin and yang regulators of cellular proliferation. Cell Res *15*, 92-96.

Gherzi, R., Lee, K.Y., Briata, P., Wegmuller, D., Moroni, C., Karin, M., and Chen, C.Y. (2004). A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. Mol Cell *14*, 571-583.

Grigull, J., Mnaimneh, S., Pootoolal, J., Robinson, M.D., and Hughes, T.R. (2004). Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors. Mol Cell Biol *24*, 5534-5547.

Grondin, B., Bazinet, M., and Aubry, M. (1996). The KRAB zinc finger gene ZNF74 encodes an RNA-binding protein tightly associated with the nuclear matrix. J Biol Chem *271*, 15458-15467.

Hall, M.A., Curtis, D.J., Metcalf, D., Elefanty, A.G., Sourris, K., Robb, L., Gothert, J.R., Jane, S.M., and Begley, C.G. (2003). The critical regulator of embryonic hematopoiesis, SCL, is vital in the adult for megakaryopoiesis, erythropoiesis, and lineage choice in CFU-S12. Proc Natl Acad Sci U S A *100*, 992-997.

Hall, T.M. (2005). Multiple modes of RNA recognition by zinc finger proteins. Curr Opin Struct Biol *15*, 367-373.

Harju, S., McQueen, K.J., and Perterson, K.R. (2000). Chromatin structure and control of beta-like globin gene switching. Exp Biol Med (Maywood) *227*, 683-700.

Heyduk, T., Ma, Y., Tang, H., and Ebright, R.H. (1996). Fluorescence anisotropy: rapid, quantitative assay for protein-DNA and protein-protein interaction. Methods Enzymol *274*, 492-503.

Hobbs, S., Jitrapakdee, S., and Wallace, J.C. (1998). Development of a bicistronic vector driven by the human polypeptide chain elongation factor 1alpha promoter for creation of stable mammalian cell lines that express very high levels of recombinant proteins. Biochem Biophys Res Commun *252*, 368-372.

Holcik, M., and Sonenberg, N. (2005). Translational control in stress and apoptosis. Nat Rev Mol Cell Biol *6*, 318-327.

Honda, B.M., and Roeder, R.G. (1980). Association of a 5S gene transcription factor with 5S RNA and altered levels of the factor during cell differentiation. Cell *22*, 119-126.

Irwin, N., Baekelandt, V., Goritchenko, L., and Benowitz, L.I. (1997). Identification of two proteins that bind to a pyrimidine-rich sequence in the 3'-untranslated region of GAP-43 mRNA. Nucleic Acids Res *25*, 1281-1288.

Iuchi, S., and Kuldell, N., eds. (2005). Zinc finger proteins: from atomic contact to cellular function (New York, Landes Bioscience; Kluwer Academic/Plenum Publishers).

Jackers, P., Szalai, G., Moussa, O., and Watson, D.K. (2004). Ets-dependent regulation of target gene expression during megakaryopoiesis. J Biol Chem *279*, 52183-52190.

Jensen, K.B., Musunuru, K., Lewis, H.A., Burley, S.K., and Darnell, R.B. (2000). The tetranucleotide UCAY directs the specific recognition of RNA by the Nova K-homology 3 domain. Proc Natl Acad Sci U S A *97*, 5740-5745.

Joho, K.E., Darby, M.K., Crawford, E.T., and Brown, D.D. (1990). A finger protein structurally similar to TFIIIA that binds exclusively to 5S RNA in Xenopus. Cell *61*, 293-300.

Jorgensen, P., and Tyers, M. (2004). How cells coordinate growth and division. Curr Biol *14*, R1014-1027.

Kaczynski, J., Cook, T., and Urrutia, R. (2003). Sp1- and Kruppel-like transcription factors. Genome Biol *4*, 206.

Keene, J.D. (2007). RNA regulons: coordination of post-transcriptional events. Nat Rev Genet *8*, 533-543.

Klein, E., Ben-Bassat, H., Neumann, H., Ralph, P., Zeuthen, J., Polliack, A., and Vanky, F. (1976). Properties of the K562 cell line, derived from a patient with chronic myeloid leukemia. Int J Cancer *18*, 421-431.

Klocke, B., Koster, M., Hille, S., Bouwmeester, T., Bohm, S., Pieler, T., and Knochel, W. (1994). The FAR domain defines a new Xenopus laevis zinc finger protein subfamily with specific RNA homopolymer binding activity. Biochim Biophys Acta *1217*, 81-89.

Kornberg, R.D. (2005). Mediator and the mechanism of transcriptional activation. Trends Biochem Sci *30*, 235-239.

Kuersten, S., and Goodwin, E.B. (2003). The power of the 3' UTR: translational control and development. Nat Rev Genet *4*, 626-637.

Kumar, M., Gromiha, M.M., and Raghava, G.P. (2007). Prediction of RNA binding sites in a protein using SVM and PSSM profile. Proteins.

Kuo, C.T., Veselits, M.L., Barton, K.P., Lu, M.M., Clendenin, C., and Leiden, J.M. (1997a). The LKLF transcription factor is required for normal tunica media formation and blood vessel stabilization during murine embryogenesis. Genes Dev *11*, 2996-3006.

Kuo, C.T., Veselits, M.L., and Leiden, J.M. (1997b). LKLF: A transcriptional regulator of single-positive T cell quiescence and survival. Science *277*, 1986-1990.

Laity, J.H., Dyson, H.J., and Wright, P.E. (2000). DNA-induced alpha-helix capping in conserved linker sequences is a determinant of binding affinity in Cys(2)-His(2) zinc fingers. J Mol Biol *295*, 719-727.

Laity, J.H., Lee, B.M., and Wright, P.E. (2001). Zinc finger proteins: new insights into structural and functional diversity. Curr Opin Struct Biol *11*, 39-46.

Lakowicz, J.R. (1999). Principles of Fluorescence Spectroscopy, Second Edition (Kluwer Academic / Plenum Publishers).

Lania, L., Majello, B., and De Luca, P. (1997). Transcriptional regulation by the Sp family proteins. Int J Biochem Cell Biol *29*, 1313-1323.

Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M.J. (2001). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. EMBO J *20*, 4987-4997.

Lee, M.S., Gippert, G.P., Soman, K.V., Case, D.A., and Wright, P.E. (1989). Three-dimensional solution structure of a single zinc finger DNA-binding domain. Science *245*, 635-637.

Lemon, B., and Tjian, R. (2000). Orchestrated response: a symphony of transcription factors for gene control. Genes Dev *14*, 2551-2569.

Lomberk, G., and Urrutia, R. (2005). The family feud: turning off Sp1 by Sp1-like KLF proteins. Biochem J *392*, 1-11.

Lopez de Silanes, I., Galban, S., Martindale, J.L., Yang, X., Mazan-Mamczarz, K., Indig, F.E., Falco, G., Zhan, M., and Gorospe, M. (2005). Identification and functional outcome of mRNAs associated with RNA-binding protein TIA-1. Mol Cell Biol *25*, 9520-9531.

Lu, D., Searles, M.A., and Klug, A. (2003). Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition. Nature *426*, 96-100.

Lu, J.Y., Sadri, N., and Schneider, R.J. (2006). Endotoxic shock in AUF1 knockout mice mediated by failure to degrade proinflammatory cytokine mRNAs. Genes Dev *20*, 3174-3184.

Luo, M.J., and Reed, R. (1999). Splicing is required for rapid and efficient mRNA export in metazoans. Proc Natl Acad Sci U S A *96*, 14937-14942.

Luo, Q., Ma, X., Wahl, S.M., Bieker, J.J., Crossley, M., and Montaner, L.J. (2004). Activation and repression of interleukin-12 p40 transcription by erythroid Kruppel-like factor in macrophages. J Biol Chem *279*, 18451-18456.

Mackay, J.P., and Crossley, M. (1998). Zinc fingers are sticking together. Trends Biochem Sci *23*, 1-4.

Maniatis, T., and Reed, R. (2002). An extensive network of coupling among gene expression machines. Nature *416*, 499-506.

Manwani, D., Galdass, M., and Bieker, J.J. (2007). Altered regulation of beta-like globin genes by a redesigned erythroid transcription factor. Exp Hematol *35*, 39-47.

Mata, J., Marguerat, S., and Bahler, J. (2005). Post-transcriptional control of gene expression: a genome-wide perspective. Trends Biochem Sci *30*, 506-514.

Matsumoto, N., Kubo, A., Liu, H., Akita, K., Laub, F., Ramirez, F., Keller, G., and Friedman, S.L. (2006). Developmental regulation of yolk sac hematopoiesis by Kruppel-like factor 6. Blood *107*, 1357-1365.

Matsumoto, N., Laub, F., Aldabe, R., Zhang, W., Ramirez, F., Yoshida, T., and Terada, M. (1998). Cloning the cDNA for a new human zinc finger protein defines a group of closely related Kruppel-like transcription factors. J Biol Chem *273*, 28229-28237.

Matthews, J.M., and Sunde, M. (2002). Zinc fingers--folds for many occasions. IUBMB Life *54*, 351-355.

McCarty, A.S., Kleiger, G., Eisenberg, D., and Smale, S.T. (2003). Selective dimerization of a C2H2 zinc finger subfamily. Mol Cell *11*, 459-470.

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M., and Bentley, D.L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature *385*, 357-361.

Mendez-Vidal, C., Wilhelm, M.T., Hellborg, F., Qian, W., and Wiman, K.G. (2002). The p53-induced mouse zinc finger protein wig-1 binds double-stranded RNA with high affinity. Nucleic Acids Res *30*, 1991-1996.

Mikkola, H.K., Klintman, J., Yang, H., Hock, H., Schlaeger, T.M., Fujiwara, Y., and Orkin, S.H. (2003). Haematopoietic stem cells retain long-term repopulating activity and multipotency in the absence of stem-cell leukaemia SCL/tal-1 gene. Nature *421*, 547-551.

Miller, I.J., and Bieker, J.J. (1993). A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins. Mol Cell Biol *13*, 2776-2786.

Miller, J., McLachlan, A.D., and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. EMBO J *4*, 1609-1614.

Mobarak, C.D., Anderson, K.D., Morin, M., Beckel-Mitchener, A., Rogers, S.L., Furneaux, H., King, P., and Perrone-Bizzozero, N.I. (2000). The RNA-binding protein HuD is required for GAP-43 mRNA stability, GAP-43 gene expression, and PKC-dependent neurite outgrowth in PC12 cells. Mol Biol Cell *11*, 3191-3203.

Morgan, B., Sun, L., Avitahl, N., Andrikopoulos, K., Ikeda, T., Gonzales, E., Wu, P., Neben, S., and Georgopoulos, K. (1997). Aiolos, a lymphoid restricted transcription factor that interacts with Ikaros to regulate lymphocyte differentiation. EMBO J *16*, 2004-2013.

Mori, T., Sakaue, H., Iguchi, H., Gomi, H., Okada, Y., Takashima, Y., Nakamura, K., Nakamura, T., Yamauchi, T., Kubota, N.*, et al.* (2005). Role of Kruppel-like factor 15 (KLF15) in transcriptional regulation of adipogenesis. J Biol Chem *280*, 12867-12875.

Morris, D.P., and Greenleaf, A.L. (2000). The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. J Biol Chem *275*, 39935-39943.

Muller, F., and Tora, L. (2004). The multicoloured world of promoter recognition complexes. EMBO J *23*, 2-8.

Myers, R.M., Tilly, K., and Maniatis, T. (1986). Fine structure genetic analysis of a beta-globin promoter. Science *232*, 613-618.

Narla, G., Heath, K.E., Reeves, H.L., Li, D., Giono, L.E., Kimmelman, A.C., Glucksman, M.J., Narla, J., Eng, F.J., Chan, A.M.*, et al.* (2001). KLF6, a candidate tumor suppressor gene mutated in prostate cancer. Science *294*, 2563-2566.

Nolte, R.T., Conlin, R.M., Harrison, S.C., and Brown, R.S. (1998). Differing roles for zinc fingers in DNA recognition: structure of a six-finger transcription factor IIIA complex. Proc Natl Acad Sci U S A *95*, 2938-2943.

Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R., and Grosveld, F. (1995). Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. Nature *375*, 316-318.

Oishi, Y., Manabe, I., Tobe, K., Tsushima, K., Shindo, T., Fujiu, K., Nishimura, G., Maemura, K., Yamauchi, T., Kubota, N.*, et al.* (2005). Kruppel-like transcription factor KLF5 is a key regulator of adipocyte differentiation. Cell Metab *1*, 27-39.

Otting, G., Qian, Y.Q., Billeter, M., Muller, M., Affolter, M., Gehring, W.J., and Wuthrich, K. (1990). Protein--DNA contacts in the structure of a homeodomain--DNA complex determined by nuclear magnetic resonance spectroscopy in solution. EMBO J *9*, 3085-3092.

Ouyang, L., Chen, X., and Bieker, J.J. (1998). Regulation of erythroid Kruppel-like factor (EKLF) transcriptional activity by phosphorylation of a protein kinase casein kinase II site within its interaction domain. J Biol Chem *273*, 23019-23025.

Pabo, C.O., Peisach, E., and Grant, R.A. (2001). Design and selection of novel Cys2His2 zinc finger proteins. Annu Rev Biochem *70*, 313-340.

Parker, R., and Song, H. (2004). The enzymes and control of eukaryotic mRNA turnover. Nat Struct Mol Biol *11*, 121-127.

Patikoglou, G., and Burley, S.K. (1997). Eukaryotic transcription factor-DNA complexes. Annu Rev Biophys Biomol Struct *26*, 289-325.

Pavletich, N.P., and Pabo, C.O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science *252*, 809-817.

Pelham, H.R., and Brown, D.D. (1980). A specific transcription factor that can bind either the 5S RNA gene or 5S RNA. Proc Natl Acad Sci U S A *77*, 4170-4174.

Peng, S.S., Chen, C.Y., Xu, N., and Shyu, A.B. (1998). RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein. EMBO J *17*, 3461-3470.

Perdomo, J., Verger, A., Turner, J., and Crossley, M. (2005). Role for SUMO modification in facilitating transcriptional repression by BKLF. Mol Cell Biol *25*, 1549-1559.

Perkins, A. (1999). Erythroid Kruppel like factor: from fishing expedition to gourmet meal. Int J Biochem Cell Biol *31*, 1175-1192.

Perkins, A.C., Sharpe, A.H., and Orkin, S.H. (1995). Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. Nature *375*, 318-322.

Pevny, L., Simon, M.C., Robertson, E., Klein, W.H., Tsai, S.F., D'Agati, V., Orkin, S.H., and Costantini, F. (1991). Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. Nature *349*, 257-260.

Picard, B., and Wegnez, M. (1979). Isolation of a 7S particle from Xenopus laevis oocytes: a 5S RNA-protein complex. Proc Natl Acad Sci U S A *76*, 241-245.

Piecyk, M., Wax, S., Beck, A.R., Kedersha, N., Gupta, M., Maritim, B., Chen, S., Gueydan, C., Kruys, V., Streuli, M.*, et al.* (2000). TIA-1 is a translational silencer that selectively regulates the expression of TNF-alpha. EMBO J *19*, 4154-4163.

Quadrini, K.J., Gruzglin, E., and Bieker, J.J. (2008). Non-random subcellular distribution of variant EKLF in erythroid cells. Exp Cell Res.

Query, C.C., Bentley, R.C., and Keene, J.D. (1989). A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein. Cell *57*, 89-101.

Raich, N., and Romeo, P.H. (1993). Erythroid regulatory elements. Stem Cells *11*, 95-104.

Raineri, I., Wegmueller, D., Gross, B., Certa, U., and Moroni, C. (2004). Roles of AUF1 isoforms, HuR and BRF1 in ARE-dependent mRNA turnover studied by RNA interference. Nucleic Acids Res *32*, 1279-1288.

Roeder, R.G. (2005). Transcriptional regulation and the role of diverse coactivators in animal cells. FEBS Lett *579*, 909-915.

Romaniuk, P.J. (1985). Characterization of the RNA binding properties of transcription factor IIIA of Xenopus laevis oocytes. Nucleic Acids Res *13*, 5369-5387.

Saleque, S., Cameron, S., and Orkin, S.H. (2002). The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages. Genes Dev *16*, 301-306.

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

Searles, M.A., Lu, D., and Klug, A. (2000). The role of the central zinc fingers of transcription factor IIIA in binding to 5 S RNA. J Mol Biol *301*, 47-60.

Selinger, D.W., Saxena, R.M., Cheung, K.J., Church, G.M., and Rosenow, C. (2003). Global RNA half-life analysis in Escherichia coli reveals positional patterns of transcript degradation. Genome Res *13*, 216-223.

Shatkin, A.J., and Manley, J.L. (2000). The ends of the affair: capping and polyadenylation. Nat Struct Biol *7*, 838-842.

Shields, J.M., Christy, R.J., and Yang, V.W. (1996). Identification and characterization of a gene encoding a gut-enriched Kruppel-like factor expressed during growth arrest. J Biol Chem *271*, 20009-20017.

Shivdasani, R.A., Fujiwara, Y., McDevitt, M.A., and Orkin, S.H. (1997). A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. EMBO J *16*, 3965-3973.

Shyu, Y.C., Lee, T.L., Wen, S.C., Chen, H., Hsiao, W.Y., Chen, X., Hwang, J., and Shen, C.K. (2007). Subcellular transport of EKLF and switch-on of murine adult beta maj globin gene transcription. Mol Cell Biol *27*, 2309-2323.

Siatecka, M., Xue, L., and Bieker, J.J. (2007). Sumoylation of EKLF promotes transcriptional repression and is involved in inhibition of megakaryopoiesis. Mol Cell Biol *27*, 8547-8560.

Singer, D., Cooper, M., Maniatis, G.M., Marks, P.A., and Rifkind, R.A. (1974). Erythropoietic differentiation in colonies of cells transformed by Friend virus. Proc Natl Acad Sci U S A *71*, 2668-2670.

Siomi, H., Matunis, M.J., Michael, W.M., and Dreyfuss, G. (1993a). The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. Nucleic Acids Res *21*, 1193-1198.

Siomi, H., Siomi, M.C., Nussbaum, R.L., and Dreyfuss, G. (1993b). The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein. Cell *74*, 291-298.

St Johnston, D. (2005). Moving messages: the intracellular localization of mRNAs. Nat Rev Mol Cell Biol *6*, 363-375.

St Johnston, D., Brown, N.H., Gall, J.G., and Jantsch, M. (1992). A conserved double-stranded RNA-binding domain. Proc Natl Acad Sci U S A *89*, 10979-10983.

Starck, J., Cohet, N., Gonnet, C., Sarrazin, S., Doubeikovskaia, Z., Doubeikovski, A., Verger, A., Duterque-Coquillaud, M., and Morle, F. (2003). Functional cross-antagonism between transcription factors FLI-1 and EKLF. Mol Cell Biol *23*, 1390-1402.

Starck, J., Doubeikovski, A., Sarrazin, S., Gonnet, C., Rao, G., Skoultchi, A., Godet, J., Dusanter-Fourt, I., and Morle, F. (1999). Spi-1/PU.1 is a positive regulator of the Fli-1 gene involved in inhibition of erythroid differentiation in friend erythroleukemic cell lines. Mol Cell Biol *19*, 121-135.

Sun, L., Liu, A., and Georgopoulos, K. (1996). Zinc finger-mediated protein interactions modulate Ikaros activity, a molecular control of lymphocyte development. EMBO J *15*, 5358-5369.

Takizawa, P.A., Sil, A., Swedlow, J.R., Herskowitz, I., and Vale, R.D. (1997). Actin-dependent localization of an RNA encoding a cell-fate determinant in yeast. Nature *389*, 90-93.

Terribilini, M., Lee, J.H., Yan, C., Jernigan, R.L., Honavar, V., and Dobbs, D. (2006). Prediction of RNA binding sites in proteins from amino acid sequence. RNA *12*, 1450-1462.

Terribilini, M., Sander, J.D., Lee, J.H., Zaback, P., Jernigan, R.L., Honavar, V., and Dobbs, D. (2007). RNABindR: a server for analyzing and predicting RNA-binding sites in proteins. Nucleic Acids Res *35*, W578-584.

Thiel, G., Lietz, M., and Hohl, M. (2004). How mammalian transcriptional repressors work. Eur J Biochem *271*, 2855-2862.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res *22*, 4673-4680.

Tsai, R.Y., and Reed, R.R. (1998). Identification of DNA recognition sequences and protein interaction domains of the multiple-Zn-finger protein Roaz. Mol Cell Biol *18*, 6447-6456.

Tsang, A.P., Visvader, J.E., Turner, C.A., Fujiwara, Y., Yu, C., Weiss, M.J., Crossley, M., and Orkin, S.H. (1997). FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. Cell *90*, 109-119.

Tsiftsoglou, A.S., Pappas, I.S., and Vizirianakis, I.S. (2003). Mechanisms involved in the induced differentiation of leukemia cells. Pharmacol Ther *100*, 257-290.

Tuerk, C., and Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science *249*, 505-510.

Tupler, R., Perini, G., and Green, M.R. (2001). Expressing the human genome. Nature *409*, 832-833.

Turner, J., and Crossley, M. (1998). Cloning and characterization of mCtBP2, a co-repressor that associates with basic Kruppel-like factor and other mammalian transcriptional regulators. EMBO J *17*, 5129-5140.

Turner, J., Nicholas, H., Bishop, D., Matthews, J.M., and Crossley, M. (2003). The LIM protein FHL3 binds basic Kruppel-like factor/Kruppel-like factor 3 and its co-repressor C-terminal-binding protein 2. J Biol Chem *278*, 12786-12795.

Valeur, B. (2001). Molecular Fluorescence: Principles and Applications (Wiley-VCH Verlag).

Van Loo, P.F., Bouwman, P., Ling, K.W., Middendorp, S., Suske, G., Grosveld, F., Dzierzak, E., Philipsen, S., and Hendriks, R.W. (2003). Impaired hematopoiesis in mice lacking the transcription factor Sp3. Blood *102*, 858-866.

van Vliet, J., Crofts, L.A., Quinlan, K.G., Czolij, R., Perkins, A.C., and Crossley, M. (2006). Human KLF17 is a new member of the Sp/KLF family of transcription factors. Genomics *87*, 474-482.

van Vliet, J., Turner, J., and Crossley, M. (2000). Human Kruppel-like factor 8: a CACCC-box binding protein that associates with CtBP and represses transcription. Nucleic Acids Res *28*, 1955-1962.

Wang, B.S., Grant, R.A., and Pabo, C.O. (2001). Selected peptide extension contacts hydrophobic patch on neighboring zinc finger and mediates dimerization on DNA. Nat Struct Biol *8*, 589-593.

Wang, L., and Brown, S.J. (2006). BindN: a web-based tool for efficient prediction of DNA and RNA binding sites in amino acid sequences. Nucleic Acids Res *34*, W243-248.

Wang, X., and Zhao, J. (2007). KLF8 transcription factor participates in oncogenic transformation. Oncogene *26*, 456-461.

Wang, Y., Liu, C.L., Storey, J.D., Tibshirani, R.J., Herschlag, D., and Brown, P.O. (2002). Precision and functional specificity in mRNA decay. Proc Natl Acad Sci U S A *99*, 5860-5865.

Wani, M.A., Means, R.T., Jr., and Lingrel, J.B. (1998). Loss of LKLF function results in embryonic lethality in mice. Transgenic Res *7*, 229-238.

Wani, M.A., Wert, S.E., and Lingrel, J.B. (1999). Lung Kruppel-like factor, a zinc finger transcription factor, is essential for normal lung development. J Biol Chem *274*, 21180-21185.

Watson, J.D., ed. (1983). Structures of DNA. Cold Spring Harbor Symposia on Quantitative Biology (New York, Cold Spring Harbor Laboratory).

Wei, H., Wang, X., Gan, B., Urvalek, A.M., Melkoumian, Z.K., Guan, J.L., and Zhao, J. (2006). Sumoylation delimits KLF8 transcriptional activity associated with the cell cycle regulation. J Biol Chem *281*, 16664-16671.

West, A.G., and Fraser, P. (2002). Remote control of gene transcription. Hum Mol Genet *14*, 101-111.

Williamson, J.R., Raghuraman, M.K., and Cech, T.R. (1989). Monovalent cation-induced structure of telomeric DNA: the G-quartet model. Cell *59*, 871-880.

Wilusz, C.J., and Wilusz, J. (2004). Bringing the role of mRNA decay in the control of gene expression into focus. Trends Genet *20*, 491-497.

Wolfe, S.A., Nekludova, L., and Pabo, C.O. (2000). DNA recognition by Cys2His2 zinc finger proteins. Annu Rev Biophys Biomol Struct *29*, 183-212.

Woychik, N.A., and Hampsey, M. (2002). The RNA polymerase II machinery: structure illuminates function. Cell *108*, 453-463.

Wright, W.E., Binder, M., and Funk, W. (1991). Cyclic amplification and selection of targets (CASTing) for the myogenin consensus binding site. Mol Cell Biol *11*, 4104-4110.

Yang, E., van Nimwegen, E., Zavolan, M., Rajewsky, N., Schroeder, M., Magnasco, M., and Darnell, J.E., Jr. (2003). Decay rates of human mRNAs: correlation with functional characteristics and sequence attributes. Genome Res *13*, 1863-1872.

Yang, M., May, W.S., and Ito, T. (1999). JAZ requires the double-stranded RNA-binding zinc finger motifs for nuclear localization. J Biol Chem *274*, 27399-27406.

Zhang, W., and Bieker, J.J. (1998). Acetylation and modulation of erythroid Kruppel-like factor (EKLF) activity by interaction with histone acetyltransferases. Proc Natl Acad Sci U S A *95*, 9855-9860.

Zhang, W., Kadam, S., Emerson, B.M., and Bieker, J.J. (2001). Site-specific acetylation by p300 or CREB binding protein regulates erythroid Kruppel-like factor transcriptional activity via its interaction with the SWI-SNF complex. Mol Cell Biol *21*, 2413-2422.

Zhao, J., Bian, Z.C., Yee, K., Chen, B.P., Chien, S., and Guan, J.L. (2003). Identification of transcription factor KLF8 as a downstream target of focal adhesion kinase in its regulation of cyclin D1 and cell cycle progression. Mol Cell *11*, 1503-1515.

Zhou, Z., Luo, M.J., Straesser, K., Katahira, J., Hurt, E., and Reed, R. (2000). The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. Nature *407*, 401-405.