A Molecular Analysis of Two Related, c-Myb-Binding Proteins; p160 and p67

Fiona Jane Tavner B.Sc. (Hons)

Hanson Centre for Cancer Research, Division of Human Immunology,

Institute of Medical and Veterinary Science, Adelaide, South Australia,

and Department of Genetics, University of Adelaide, South Australia.

A thesis submitted for the Degree of Doctor of Philosophy in the Faculty of Science,
University of Adelaide, South Australia.

January, 1998.

TABLE OF CONTENTS

Chapter One

Introduction

Chapter Two

Materials and Methods

Chapter Three

Cloning and Characterisation of p160 and p67

Chapter Four

Determination of the Relationship Between p160 and p67

Chapter Five

SUMMARY

The c-*myb* proto-oncogene plays a central role in vertebrate haemopoiesis and encodes a DNA-binding transcription factor (c-Myb) that functions to maintain proliferation of immature haemopoietic cells. Both c-Myb and oncogenically activated forms, for example, the viral Myb proteins encoded by the AMV and E26 avian acute leukaemia viruses, are able to directly bind DNA in a sequence-specific manner and transcriptionally activate 'target' genes harbouring their cognate recognition sequence. The DNA-binding and transactivation activities of Myb are necessary for the ability to transform haemopoietic cells.

The activities of c-Myb are regulated in a number of ways, encompassing transcriptional and post-translational mechanisms. The negative regulatory domain (NRD) of c-Myb, located in the carboxyl-terminal region, is implicated as functionally important region required for the down-regulation of c-Myb activity. One identified component of the NRD is a leucine zipper motif, a structure known to mediate protein-protein interactions. Therefore, proteins that interact with the c-Myb leucine zipper may represent crucial regulators of c-Myb activity.

Two related proteins, termed p160 and p67, have been detected within nuclear extracts of murine cells on the basis of their ability to bind to a GST fusion protein containing the c-Myb leucine zipper motif, but not to similar fusion proteins containing mutated leucine zipper motifs (Favier and Gonda, 1994). These proteins are related both structurally and functionally, but possess distinct properties, namely their ability to interact with the c-Jun basic-leucine zipper (bZIP) region, and their distribution amongst the murine cell lines examined.

A partial murine cDNA clone had been isolated using sequence information derived from tryptic peptides generated from purified p67. This thesis describes the complete cloning of cDNA encoding both p160 and p67, and characterisation of these proteins, including the determination of their relationship.

A full-length murine cDNA clone (4.1 kb) corresponding to p160 mRNA was obtained. Database searches indicate that the p160 sequence is novel, and in addition, contains all five of the tryptic peptide sequences derived from p67. Interestingly, the p67-derived sequences are clustered in the amino-terminal region of p160. Northern analysis with p160 cDNA sequences revealed one, ubiquitously expressed mRNA species of 4.5 kb amongst the murine cell lines examined. Based on several lines of evidence, it was hypothesised that p160 and p67 have a precursor-product relationship, whereby p67 is derived from the amino-terminal region of p160 by proteolytic cleavage. In order to further study the properties of p67, a carboxyl-truncated p160 clone (termed p67*) was generated, and used to approximate the properties of endogenous p67.

The authenticity of the p160 cDNA clone was demonstrated by comparative peptide mapping of endogenous p160 and p67 with *in vitro* translated p160 and p67*. An investigation of the binding properties of p160 and p67*, using proteins synthesised *in vitro* and *in vivo*, revealed the same binding specificities as endogenous p160 and p67. The ability of p160 and p67* to interact with wild type c-Myb within the intracellular environment was examined by coimmunoprecipitation of ectopically expressed proteins.

Polyclonal antisera were raised against the amino- and carboxyl-terminal regions of p160. Both endogenous p160 and p67 were detected with antiserum raised against the amino-terminal region, whereas p160, but not p67, was detected with antiserum raised against the carboxyl-terminal region. Detection of endogenous p160 and p67 with the p160 antisera revealed the same subcellular localisation and distribution of these proteins among murine cell lines as previously described binding data.

The subcellular localisation of p67* prompted a further investigation of this protein. Unlike endogenous p67, which was detected within nuclear extracts, p67* was detected only within cytoplasmic extracts. Additionally, p67* lacks carboxyl-terminal residues present within p67. Two approaches were undertaken in order to target p67* to the nucleus; (1) The carboxyl terminus of p67* was extended by generating two additional p160 carboxyl-truncated proteins. (2) The SV40 large T-antigen nuclear localisation signal was incorporated into the carboxyl terminus of p67*. Analysis of the subcellular localisation of the p160 carboxyl-truncated proteins and nuclear-targeted p67* fusion proteins has

revealed potentially important sites/domains within p160 that contribute to transport between subcellular compartments.

The structural relatedness of p160 and p67 was established on the basis of common peptides, as shown by peptide mapping (Favier and Gonda, 1994). Examination of the p160 cDNA sequence and detection of a single mRNA species suggested a precursor-product relationship. This possible relationship was investigated by pulse-chase analyses, and proteolysis assays using total cell extracts and *in vitro* translated p160. Proteolytic cleavage of p160 was directly demonstrated by proteolysis assays. These assays showed that cells in which p67 was detected, contain a proteolytic activity that was able to cleave *in vitro* translated p160 and rapidly generate four major cleavage products, including one of 67 kDa that maps to the amino-terminal region of the p160 precursor protein, in which the p67-derived peptide sequences are located. Furthermore, this proteolytic activity can be specifically inhibited, and is composed of a serine protease(s). Pulse-chase analysis of p160 and preparation of nuclear extracts (from cells in which both p160 and p67 were detected) in the presence of this specific protease inhibitor, indicated that p67 can be generated during the nuclear extraction procedure. Subsequently, preliminary investigations were undertaken to examine whether p67 could be generated *in vivo*.

DECLARATION

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

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

Fiona J. Tavner

January, 1998.

ACKNOWLEDGEMENTS

ABBREVIATIONS

A	Adenine
AMV	Avian myeloblastosis virus
ATP	Adenosine Triphosphate
bp	Base pair
BSA	Bovine serum albumin
bZIP	Basic region and adjacent leucine zipper
C	Cytosine
CBP	CREB binding protein
CcN	Casein kinase II-cdc2-NLS motif of SV40 large T antigen
cDNA	Complementary deoxyribonucleic acid
CHAPS	3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulphonate
dATP	2'-deoxyadenosine-5'-triphosphate
DMEM	Dulbecco's modified Eagle's medium
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylene diamine tetra acetic acid
FCS	Foetal calf serum
G	Guanine
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GST	Glutathione S-transferase
HEPES	N-[2-hydroxyethyl] piperazine-N'-[2-ethane sulphonic acid]
ICE	Interleukin-1β converting enzyme
IPTG	Isopropyl-β-D-thiogalactoside
kb	Kilo base pair
kDa	Kilo Dalton
mRNA	Messenger RNA
rRNA	Ribosomal RNA
NES	Nuclear export signal
NLS	Nuclear localisation signal

NP-40	Nonidet P-40
NRD	Negative regulatory domain
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PIPES	Piperazine-N,N'-bis[2-ethane-sulphonic acid]
PMSF	Phenyl-methyl sulfonyl fluoride
RACE	Rapid amplification of cDNA ends
RNA	Ribonucleic acid
RT	Room temperature
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SSC	Saline sodium citrate
SV40	Simian virus 40
T	Thymidine
T-ag	T antigen
TBS	Tris buffered saline
TEMED	N, N, N', N'-tetramethyl-ethenediamine
TPCK	*N*-tosyl-L-phenylalanine chloromethyl ketone
Tris	Tris (hydroxymethyl) amino methane
UV	Ultraviolet

CHAPTER ONE
Introduction

1.1 Overview

The c-Myb protein is a DNA-binding transcription factor that is expressed in a tissue-specific manner. Analyses of c-Myb have predominantly focused on its function within the vertebrate haemopoietic system. The haemopoietic system provides a developmental model in which to study regulatory aspects of cellular proliferation and differentiation. The c-Myb transcription factor functions as a key regulatory factor within the haemopoietic system. The molecular mechanisms by which c-Myb transcriptionally regulates gene expression are only beginning to be understood. The identification and analysis of cellular proteins that interact with c-Myb are providing insight into the mechanisms by which c-Myb regulates gene transcription. Analysis of the c-Myb negative regulatory domain, a region that down-regulates the transcriptional activities of c-Myb, reveals that it is a complex region comprised of multiple components that contribute to post-translational regulation. One component constitutes a leucine zipper motif, known to mediate protein-protein interactions. This thesis describes the cloning, identification and characterisation of two related nuclear proteins that associate with the c-Myb negative regulatory domain by interaction with the leucine zipper motif. This chapter represents a review of the literature that pertains primarily to the function and structure of the c-Myb transcription factor. It serves to highlight the current status of knowledge with emphasis on the molecular mechanisms that contribute to the activity and regulation of this transcription factor.

1.2 Cellular Proliferation and Differentiation

The study of tumour viruses, particularly the RNA tumour viruses (a subgroup of the retrovirus family), initiated the isolation and identification of numerous oncogenes and their cellular progenitors; proto-oncogenes. By virtue of their life cycle, retroviruses are mutagenic agents and are known to cause many types of neoplasia (Bishop, 1985). Elucidation of the functional activities of proto-oncogenes and viral oncogenes has had a profound impact on our understanding of the mechanisms underlying cellular proliferation and differentiation. The protein products of these genes constitute components of signal transduction pathways that lead from the cell surface to the transcriptional machinery

within the nucleus. In particular, oncoproteins comprise growth factors, growth factor receptors, intracellular signalling molecules and transcription factors. It is within the nucleus where growth signals are translated into programmes of gene expression that alter cell behaviour. Proto-oncogenes, together with tumour suppressor and cell survival genes form the molecular circuitry of growth control.

1.3 The *myb* Genes

1.3.1 The viral *myb* oncogenes

The viral *myb* (v-*myb*) oncogenes were first identified from two avian acute leukaemia, replication-defective retroviruses; avian myeloblastosis virus (AMV) and E26 leukaemia virus (E26) (Bister *et al.*, 1982; Roussel *et al.*, 1979). AMV induces acute myeloblastic leukaemia in chickens by transforming haemopoietic cells of the myelomonocytic lineage and is capable of transforming the same cell types *in vitro* (Moscovici *et al.*, 1981; Beug *et al.*, 1979). The E26 leukaemia virus predominantly induces erythroblastosis *in vivo* and transforms haemopoietic progenitor cells of both erythroid and myeloid lineages *in vitro* (Radke *et al.*, 1982; Moscovici *et al.*, 1981).

Molecular cloning and sequencing of the v-*myb* oncogenes from AMV and E26 has delineated the structure of these genes. Transduction of a subset of cellular *myb* (c-*myb*: myeloblast) sequences by AMV and E26 resulted in the replacement of some retroviral gene sequences, which consequently culminates in the expression of a fusion protein containing viral and c-*myb* encoded peptides (Leprince *et al.*, 1983; Nunn *et al.*, 1983; Klempnauer *et al.*, 1982). In addition to transduced c-*myb* sequences, E26 also carries sequences derived from the cellular *ets* gene and hence is expressed as a fusion protein of viral, *myb* and *ets* encoded peptides (Leprince *et al.*, 1983; Nunn *et al.*, 1983). The c-*myb* derived sequences of E26 represent a subset of those transduced by AMV (Nunn *et al.*, 1983). The expression of *ets* sequences from the E26 genome gives this retrovirus the capacitiy to transform progenitor cells of the erythroid lineage (Nunn and Hunter, 1989). The genetic factors that confer the cell type-specific transforming abilities of AMV and E26 are attributed to the *myb* sequences of AMV and both the *myb* and *ets* sequences of E26, not to differences in the viral long terminal repeats (LTR) (Engelke and Lipsick, 1994; Introna *et al.*, 1990).

Identification of v-*myb* lead firstly to the isolation and characterisation of its cellular progenitor, c-*myb*, from chicken (Klempnauer *et al.*, 1982). Soon after, murine and human c-*myb* homologues were isolated (Majello *et al.*, 1986; Gonda *et al.*, 1985). Characterisation of the nucleotide sequence of c-*myb* has allowed the precise relationship of viral and cellular *myb* genes to be determined (Klempnauer *et al.*, 1982). The v-*myb* encoded proteins of AMV and E26 represent amino and carboxyl truncations of the 75 kDa protein encoded by c-*myb* (c-Myb) (Figure 1.1B) (Gerondakis and Bishop, 1986). The 45 kDa fusion protein expressed by AMV (v-Myb^{AMV}), in addition to 370 residues derived from c-*myb*, contains 6 residues from the viral *gag* gene and 11 residues from the viral *env* gene, which constitute the amino and carboxyl termini, respectively (Klempnauer *et al.*, 1983; Klempnauer *et al.*, 1982). The *myb* encoded residues of the 135 kDa gag-Myb-Ets fusion protein expressed by E26 (v-Myb^{E26}) represent a further truncated form of v-Myb^{AMV}, lacking 9 and 79 residues from the amino and carboxyl termini, respectively (Nunn *et al.*, 1983). As well as the truncated nature of v-Myb proteins with respect to c-Myb, eleven point mutations in v-Myb^{AMV} and one point mutation in v-Myb^{E26} result in amino acid substitutions within these proteins (Nunn *et al.*, 1983; Klempnauer *et al.*, 1982). Truncation of c-Myb points to a mechanism for oncogenic activation.

1.3.2 The *myb* gene family

Elucidation of the c-*myb* DNA sequence provided several clues concerning the function of the encoded protein. The initial findings that both c-Myb and v-Myb proteins are able to bind to DNA in a sequence-specific manner pointed to a transcriptional role (Nakagoshi *et al.*, 1990; Biedenkapp *et al.*, 1988). Particular attention has been paid to a region of the c-Myb amino terminus that constitutes the DNA-binding domain. This region comprises three imperfect tandem repeats of 51-52 amino acids (Gonda *et al.*, 1985). A more detailed account of the c-Myb DNA-binding domain is presented in section 1.5.2. The c-Myb DNA-binding domain forms the basis for definition of *myb*-related genes.

1.3.2.1 A-*myb* and B-*myb*

In addition to c-*myb*, two other highly conserved *myb* genes exist in vertebrates, namely A-*myb* and B-*myb*. Collectively these proteins compose the vertebrate *myb* gene family. The proteins encoded by A-*myb* (A-Myb) and B-*myb* (B-Myb) are reported to play transcriptional roles (Lane *et al.*, 1997; Tashiro *et al.*, 1995; Trauth *et al.*, 1994; Foos *et al.*,

Figure 1.1 Schematic representation of murine c-Myb and viral Myb oncoproteins.

(A) Recognised functional domains (black bars beneath schematic) and regulatory sites of c-Myb (murine: 636 amino acids, 75 kDa) are shown. Numbers below the schematic represent amino acid positions. R1-R3 are imperfect direct repeats; LZ represents the leucine zipper motif and is expanded above the schematic to indicate heptad-repeated residues; Serine residues (shown in red) represent regulatory phosphorylation sites. Cellular proteins known to interact with c-Myb are indicated above the schematic. Black bars represent c-Myb interactive regions.

(B) The relationship of the viral Myb oncoproteins, encoded by the AMV and E26 acute leukaemia retroviruses, with respect to c-Myb. From left to right, numbers represent the positions of the first amino-terminal and last carboxyl-terminal amino acid residue, respectively. The amino acid substitutions of AMV v-Myb are indicated by vertical lines.

This figure is modified from Gonda (1991). See text for details.

A

Interacting
Proteins

C/EBPβ

HSF3

p100

p160, p67

c-Myb?

375-**M-I-L-L-L**-403
FAETL

c-Myb C-terminus (EVES)

CBP

c-Myb
N-terminus

NH₂

R1   R2   R3

LZ

COOH

100   200   300   400   500   600

Ser11 12

EVESer 528

DNA-Binding

NRD1   NRD2

Transactivation

Negative Regulatory
Domain

B

v-Myb — AMV   72 ──────── 436
         E26   81 ──────── 362
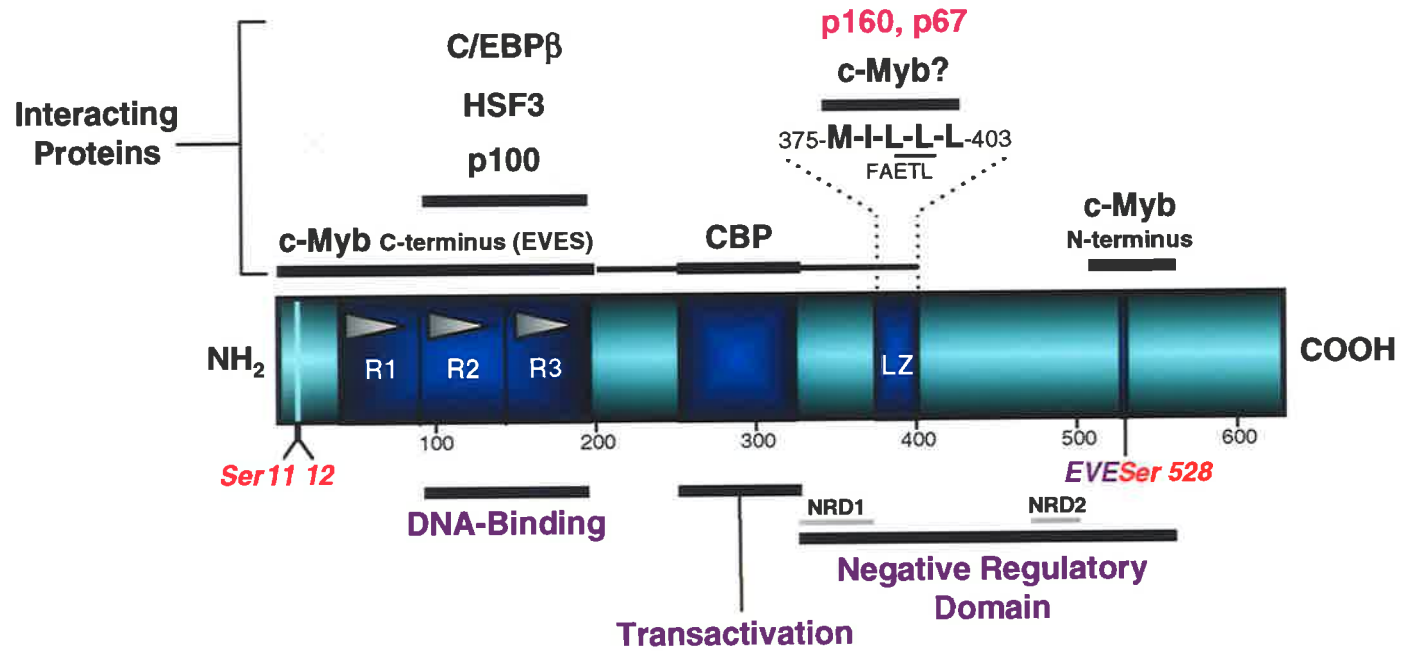
1994; Nakagoshi *et al.*, 1992; Foos *et al.*, 1992). Both A-Myb and B-Myb have highly conserved c-Myb DNA-binding domains and maintain sequence conservation within regions of the carboxyl terminus, however, both A-Myb and B-Myb lack the leucine zipper motif, located in the c-Myb negative regulatory domain (Lipsick, 1996; Nomura *et al.*, 1988). Additionally, the B-Myb transactivation domain shows no identity with that of c-Myb (Nakagoshi *et al.*, 1993).

Like c-*myb*, expression of A-*myb* is restricted to certain cell types. Expression of A-*myb* has been detected in proliferating, neuronal cells of certain lineages, adult male germ cells, B lymphocytes and female breast ductal epithelia (Toscani *et al.*, 1997; Trauth *et al.*, 1994). In addition to distinct expression patterns, the transactivational properties of A-Myb are similar to those of c-Myb, and suggest that A-Myb may represent a functional c-Myb homologue, with specific roles in cellular proliferation and differentiation (Trauth *et al.*, 1994). A recent report describes the developmental aberrations of mice lacking a functional A-*myb* gene (A-*myb* 'knockout' mice) (Toscani *et al.*, 1997). These mice survive to adulthood but their growth is comprised during embryogenesis. Adult male A-*myb* knockout mice are infertile due to disrupted spermatogenesis as a consequence of defective meiosis. Adult female A-*myb* knockout mice display abnormal breast development during and after pregnancy as a result of disrupted cellular proliferation.

In contrast to the expression of c-*myb* and A-*myb*, B-*myb* expression is ubiquitous and correlates with cellular proliferation (Arsura *et al.*, 1994; Foos *et al.*, 1992; Golay *et al.*, 1991). Indeed, B-Myb is regulated in a cell cycle-dependent manner and may represent an integral determinant of G1 to S phase progression (Ziebold *et al.*, 1997; Lane *et al.*, 1997; Sala *et al.*, 1997).

1.3.2.2 Genes related to *myb*

Other *myb*-related genes have been identified in a broad spectrum of eukaryotic species representative of fungi, slime moulds, yeast, diptera and plants (reviewed in Lipsick, 1996). These genes have been designated as '*myb*-related' because they all contain highly conserved sequences homologous to the c-Myb DNA-binding domain. This domain, termed the 'Myb domain' in *myb*-related genes, consists of a core of approximately 50 amino acids which is directly repeated, generally 2 or 3 times (Lipsick, 1996). The

Drosophila myb-related gene additionally contains sequences that show similarity to those within the carboxyl region of vertebrate Myb proteins (Peters *et al.*, 1987). One underlying theme is apparent from studies of *myb*-related genes; the products of such genes are transcriptional regulators that constitute crucial regulatory molecules with developmental roles (reviewed in Lipsick, 1996).

1.4   The Role of c-*myb*

1.4.1   Expression of c-*myb* in haemopoietic cells

Examination of c-*myb* mRNA expression among chicken tissues and human haemopoietic cells demonstrated that the highest levels of c-myb mRNA expression are found predominantly in immature haemopoietic cells and suggested that c-*myb* expression is associated with cellular proliferation, as mature, differentiated haemopoietic cells which cease to proliferate, do not express detectable levels of c-*myb* mRNA (Gonda and Metcalf, 1984; Westin *et al.*, 1982; Gonda *et al.*, 1982). However, expression is not seen in macrophages, a proliferating mature cell type, indicating that there is not a simple correlation between c-*myb* expression and cellular proliferation (Gonda, 1991; Gonda and Metcalf, 1984; Gonda *et al.*, 1982). In accordance with the detection of c-*myb* expression in particular cell types, haemopoietic tissues exhibit considerable c-*myb* expression (Thompson and Ramsay, 1995; Sheiness and Gardinier, 1984; Gonda *et al.*, 1982). Examination of primary and cultured haemopoietic cells representative of the myeloid, erythroid and lymphoid lineages from several species establish that c-*myb* expression is predominant among immature cells (eg Bender and Kuehl, 1986; Golay *et al.*, 1991; McClinton *et al.*, 1990; Todokoro *et al.*, 1988; Thompson *et al.*, 1986; Sheiness and Gardinier, 1984; Westin *et al.*, 1982).

Resting, mature T-lymphocytes do not express detectable levels of c-*myb*. However, when these cells are stimulated to proliferate *in vitro*, c-*myb* expression is induced in a biphasic pattern (Golay *et al.*, 1991, and references therein). The induction of c-*myb* expression within these cells corresponds firstly to early progression through the G1 stage of the cell cycle, then progression into S phase, indicating a cell cycle-dependent mode of expression, which is distinct from the cell cycle-independent expression in immature thymocytes (Thompson *et al.*, 1986). Hence c-*myb* may not only be functionally important in T-cell differentiation, but also T-cell activation. Similarly, in murine B-lymphoid tumour cells,

constitutive or cell-cycle regulated expression of c-*myb* is associated with a particular stage of cell differentiation (Catron *et al.*, 1992).

1.4.2   Expression of c-*myb* in non-haemopoietic cells

Further studies revealed that c-*myb* expression is not solely restricted to cells of haemopoietic origin. Expression of c-*myb* has been detected in adult human and murine colonic epithelia as well as human pre/malignant colonic epithelia (Rosenthal *et al.*, 1996; Ramsay *et al.*, 1992b; Torelli *et al.*, 1987). Expression is higher in colonic tumours than normal tissue and examination of the c-Myb protein reveals that it is not truncated, in contrast to the transforming v-Myb proteins of the avian acute leukaemia viruses (Ramsay *et al.*, 1992b; Torelli *et al.*, 1987). The colonic crypt is a specialised structure with zones of highly proliferating and non-proliferating cells. Immunohistochemical staining of the murine colonic crypt demonstrates that c-*myb* expression is not entirely restricted to proliferating cells, but is found in some differentiated, non-proliferating cell types (Rosenthal *et al.*, 1996). The expression of c-*myb* within colonic tissue implies that c-Myb may function outside the haemopoietic system, and is in some cases compatible with a differentiated, non-proliferative phenotype.

Lower levels of c-*myb* expression have been detected in chick embryo fibroblasts (CEF) (Thompson *et al.*, 1986). Maximal expression of c-*myb* in CEF cells coincides with the G1/S transition, indicating that c-*myb* expression is transient and cell cycle-dependent in these cells (Thompson *et al.*, 1986).

1.4.3   Human malignancies and c-*myb* expression

Human malignant cells of haemopoietic and colonic origin exhibit elevated levels of c-*myb* expression, as determined by examination of primary cells and cell lines (Ramsay *et al.*, 1992b; Torelli *et al.*, 1987; Slamon *et al.*, 1986; Slamon *et al.*, 1984). Expression of c-*myb* has been detected to varying extents in several other human neoplastic cell types derived from neural, breast and lung tissue (Raschella *et al.*, 1992; Zhou *et al.*, 1989; Thiele *et al.*, 1988; Griffin and Baylin, 1985). These studies have predominantly examined the expression of c-*myb* in cell lines, but not in primary cells. It is therefore uncertain whether c-*myb* expression in these cell lines is a consequence of prior genetic events in the

transformation process, or whether c-*myb* is expressed specifically for a functional requirement in these cells.

Amplification of the c-*myb* locus on chromosome 6q22-24 can contribute to overexpression of this gene, as in the reported cases of acute myelogenous leukaemia and a small percentage of breast carcinomas and pancreatic cancers (Wallrapp *et al.*, 1997; Zhou *et al.*, 1989; Pelicci *et al.*, 1984). Overexpression of c-*myb* is however associated with aberrant regulatory mechanisms other than gene amplification, as revealed by certain colon and breast carcinomas (Zhou *et al.*, 1989; Ramsay *et al.*, 1992b). Recently, overexpression of c-*myb* in particular colon carcinoma cell lines has been attributed to microsatellite deletions within a transcriptional attenuator region located in the first intron of this gene locus, which result in transcriptional dysregulation (Thompson *et al.*, 1997). In specific tissue and cell types where c-*myb* may indeed be functionally important (eg colonic epithelia), deregulation of c-*myb* expression could represent a central contributing genetic factor towards the generation of a neoplastic phenotype.

1.4.4 Differentiation and Myb

The *myb*-carrying avian acute leukaemia viruses, AMV and E26, exclusively transform cells of haemopoietic origin. Myeloid cells transformed by temperature-sensitive (ts) mutants of AMV and E26 differentiate into macrophage-like cells at the non-permissive temperature (Beug *et al.*, 1984; Moscovici and Moscovici, 1983). Differentiated, cycling macrophage-like cells transformed with an E26(ts) mutant retrovirus undergo 'retrodifferentiation' at the permissive temperature to cells that exhibit immature properties, that is the re-acquirement of a myeloblastic phenotype (Beug *et al.*, 1987). This study indicates that v-*myb* is able to enforce an immature phenotype in myeloid cells, which is also demonstrated by the cell type-specific transforming capacities of AMV and E26. Similarly, the induced differentiation of a murine myelomonocytic progenitor cell line with granulocyte colony-stimulating factor (G-CSF) is blocked by constitutive expression of v-Myb[AMV] (Patel *et al.*, 1993). Furthermore, immature cells continued to proliferate in the presence of G-CSF, suggesting that the activity of v-Myb is able to disassociate differentiation and proliferation, aspects of haemopoiesis that represent a continuum of intimate regulation (Patel *et al.*, 1993).

A large number of studies have examined the relationship between c-*myb* expression and differentiation using haemopoietic cell lines that can be chemically induced to terminally differentiate *in vitro*. As suggested by the predominant expression of c-*myb* in immature, proliferating haemopoietic cells, the induced differentiation of myeloid and erythroid leukaemia cell lines progresses with a distinct decrease in c-*myb* expression (Todokoro *et al.*, 1988; Gonda and Metcalf, 1984; Craig and Bloch, 1984; Westin *et al.*, 1982). Furthermore, ectopic overexpression of c-*myb* can inhibit induced differentiation of myeloid and erythroid leukaemia cell lines, indicating that down-regulation of c-*myb* expression is a necessary requirement for terminal differentiation (Danish *et al.*, 1992; Yanagisawa *et al.*, 1991; McClinton *et al.*, 1990; Todokoro *et al.*, 1988; Clarke *et al.*, 1988).

The temporal expression of c-*myb* during haemopoietic cell differentiation is critical. Expression of c-*myb* is down-regulated at a late stage of cell maturation that directly precedes terminal differentiation. Myelomonocytic cells induced to differentiate express c-*myb* in promonocytes, representative of a later stage in this differentiation pathway (Gonda and Metcalf, 1984). Moreover, constitutive expression of c-*myb* in these same cells inhibits terminal differentiation to macrophages, but allows maturation of cells to the promonocyte stage (Yanagisawa *et al.*, 1991). Likewise, induced terminal differentiation of a myelomonocytic progenitor cell line (32D cl3) by a physiological inducer (G-CSF) was inhibited by overexpression of c-*myb* (Bies *et al.*, 1995). However, cell maturation proceeded beyond the early and intermediate stages of granulocyte differentiation, indicating that terminal differentiation was blocked at cells in a relatively late stage of maturation (Bies *et al.*, 1995).

Constitutive expression of wild type c-*myb* in primary murine haemopoietic cells transforms these cells in a growth factor- and density-dependent manner *in vitro*, suggesting the involvement of an autocrine factor in c-*myb* enforced transformation of these cells (Ferrao *et al.*, 1995). Primary murine haemopoietic cells transformed by c-*myb* display a partial inhibition of differentiation and a highly proliferative state with characteristics of immature myelomonocytic cells (Ferrao *et al.*, 1995). In a similar study, constitutive expression of wild type c-*myb* was shown to transform avian myelomonocytic cells, isolated from embryonic yolk sacs, with heterogeneous phenotypes that resemble the

spectrum of maturation stages of both the monocyte and granulocyte lineages. This is in contrast to the homogeneous cell phenotypes enforced by v-myb^{AMV} (monoblast) and v-myb^{E26} (myeloblast) *in vitro* (Fu and Lipsick, 1997). Both studies indicate that cellular transformation enforced by constitutive expression of wild type c-*myb* is less efficient than that by v-*myb*. This may be a reflection of the relative presence of target cell-types, representative of particular stages of myelomonocytic maturation (Fu and Lipsick, 1997).

In situations where cellular proliferation can be uncoupled from differentiation, c-*myb* expression does not solely correlate with the proliferative status of the cell (Arsura *et al.*, 1994). Taken together, the studies of *myb* expression and differentiation indicate that c-*myb* functions to maintain an immature phenotype in myeloid and erythroid cells, of which proliferative capacity is an integral characteristic, and that late down-regulation of c-*myb* expression is required for terminal maturation. Studies of enforced *myb* expression in myelomonocytic and embryonic stem cells suggest a role for c-*myb* in lineage committment (Fu and Lipsick, 1997; Melotti and Calabretta, 1996).

1.4.5 Loss of c-*myb* function

Antisense oligonucleotides directed against c-*myb* have been used in an effort to gain insight into the role of c-*myb* by examining the consequence of loss of function in several cell types *in vitro*, including human primary and cultured leukaemia cells, erythroid progenitors, colon carcinoma, peripheral blood and bone marrow mononuclear cells (Ramsay *et al.*, 1992b; Valtieri *et al.*, 1991; Calabretta *et al.*, 1991; Caracciolo *et al.*, 1990; Anfossi *et al.*, 1989; Gerwitz and Calabretta, 1988). Essentially, the introduction of antisense c-*myb*, but not control sense, oligonucleotides resulted in a marked decrease in proliferation of actively cycling cells. In stimulated proliferating T-cells, the introduction of antisense c-*myb* causes a block specifically in the G1/S transition phase (Gerwitz *et al.*, 1989). Proliferating cells at particular stages of maturation display differing sensitivities to c-*myb* antisense oligonucleotides which may be an indication of the differential requirement for c-Myb function at specific stages during differentiation (Valtieri *et al.*, 1991; Calabretta *et al.*, 1991; Caracciolo *et al.*, 1990; Anfossi *et al.*, 1989). The interpretation of investigations using antisense oligonucleotides to probe the potential functions of c-*myb* must be treated with caution in view of recent studies outlining the

growth-retarding effects of oligonucleotides in a complementarity-independent context (Burgess *et al.*, 1995).

Compelling evidence for the importance of c-*myb* function in haemopoietic development has been provided by inactivating both murine c-*myb* alleles by homologous recombination in embryonic stem cells to generate c-*myb* 'knockout' mice (Mucenski *et al.*, 1991). By embryonal day 15, homozygous c-*myb* mutant foetuses display abnormal liver haemopoiesis and consequently die at this stage *in utero* from severe anaemia. Furthermore, loss of c-*myb* function resulted in a substantial decrease in haemopoietic progenitor cells of most lineages, indicating impaired proliferation, however cellular differentiation still occured, suggesting that c-*myb* plays a pivotal role in haemopoietic cell proliferation and differentiation (Lin *et al.*, 1996; Mucenski *et al.*, 1991). Foetal haemopoiesis in the yolk sac was not affected by loss of c-*myb* function (Mucenski *et al.*, 1991). This is consistent with the lack of c-*myb* expression in the yolk sac up to embryonal day 10, at which stage c-*myb* expression is detected in the foetal liver, indicating that yolk sac haemopoiesis is independent of c-*myb* function (Sitzmann *et al.*, 1995; Mucenski *et al.*, 1991).

Interestingly, up to embryonal day 15, no other obvious developmental abnormalities were found, despite the observations that c-*myb* expression is more wide-spread during embryogenesis (Rosenthal *et al.*, 1996; Sitzmann *et al.*, 1995; Mucenski *et al.*, 1991). Embryogenic expression of c-*myb* is detected in liver, developing thymus, neural retina, respiratory tract, colon and olfactory epithelia (Rosenthal *et al.*, 1996; Sitzmann *et al.*, 1995). The absence of non-haemopoietic developmental abnormalities in c-*myb* 'knockout' embryos may be explained by the embryonic lethality, due to aberrant haemopoiesis, preceding the potential manifestation of non-haemopoietic abnormalities, for example abnormal colon development (Rosenthal *et al.*, 1996).

The embryonic lethal nature of c-*myb* 'knockout' mice does not allow the further investigation of developmental processes beyond this stage. An alternative approach to study the consequences of loss of c-*myb* function involves the analysis of dominant negative forms of c-Myb. Transgenic mice expressing dominant negative forms of c-Myb specifically in the T-cell lineage demonstrate disrupted thymopoiesis with significantly

reduced numbers of thymocytes and a partial inhibition of T-cell differentiation (Badiani *et al.*, 1994). Futhermore, the severely depleted mature T-cells were unable to proliferate upon mitogen stimulation *in vitro* (Badiani *et al.*, 1994). Loss of c-*myb* function, in the context of T-cell maturation and activation, implicates c-*myb* as a critical regulator of cellular proliferation which is consistent with the studies previously discussed.

Each dominant negative c-Myb protein contains the DNA-binding domain. One form additionally contains a strong repressor domain derived from the *Drosophila* Engrailed protein and was found to be significantly more effective than the competitive inhibitor containing only the DNA-binding domain, signifying that Myb functions in a DNA-binding-dependent transcriptional context (Badiani *et al.*, 1994). A report describing the further analysis of Myb dominant negative transgenic mice shows that cells at all stages of T-cell development are more susceptible to apoptosis, which is attributed to the altered expression of specific Myb target genes (see section 1.6 for further discussion) (Taylor *et al.*, 1996).

1.5   The c-Myb Protein

Early studies demonstrated that the Myb proteins are located within the nucleus, and are associated specifically with chromatin, the nuclear matrix and nucleoplasm (Klempnauer *et al.*, 1986; Boyle *et al.*, 1985; Klempnauer *et al.*, 1984; Boyle *et al.*, 1984). Three functional domains and several regulatory phosphorylation sites have been identified within the c-Myb protein (Figure 1.1A). These domains represent hallmarks of transcriptional regulators and confer DNA-binding, transactivation and negative regulatory activities (Sakura *et al.*, 1989). Analyses of the structure and function of these domains identify c-Myb as a sequence-specific DNA-binding protein that functions by cooperative transcriptional regulation of specific target genes. A brief overview of RNA polymerase II transcription will be presented in order to discern the fundamental features of eukaryotic gene expression. The structure and function of each c-Myb domain, and regulatory phosphorylation sites, will be discussed separately.

1.5.1   Eukaryotic gene regulation and RNA polymerase II transcription

The initiation of transcription by RNA polymerase II (pol II) involves the cooperative interaction of a myriad of proteins that assemble on the promoter sequences of genes. This

assembly of proteins is commonly referred to as the 'basal transcription machinery' (reviewed in Ptashne and Gann, 1997; Zawel and Reinberg, 1995; Koleske and Young, 1995). Formation of the pol II pre-initiation complex proceeds in a defined succession and begins with recognition of the TATA element within the promoter by TFIID, a protein complex comprised of TATA-binding protein (TBP) and TBP-associated factors (TAFs) (Zawel and Reinberg, 1995; Tjian and Maniatis, 1994; Buratowski, 1994). Binding and formation of TFIID on the promoter provides a favourable molecular scaffold for subsequent interaction with TFIIA and the 'bridging protein' TFIIB, which is able to associate with TFIIF and recruit pol II (Ptashne and Gann, 1997; Zawel and Reinberg, 1995). Following this sequence of interactions are several other proteins, for example TFIIE and TFIIH, that contribute to the stable formation of the complete transcription pre-initiation complex and consequent activation of gene transcription from the initiation site by pol II (Zawel and Reinberg, 1995; Buratowski, 1994). Initiation of transcription can alternatively proceed from promoters that do not contain a TATA element (Sachs and Buratowski, 1997).

The precise activation of gene expression is in part mediated by transcription factors with distinct DNA-binding specificities in conjunction with the basal transcription machinery. The DNA sequences that direct the binding of transcription factors are located within the promoters of genes and in addition, can be situated in distal regulatory DNA elements termed enhancers, which can exist at vast distances (kilobases) either upstream or downstream of the promoter and transcription initiation site (Tjian and Maniatis, 1994). Transcriptional activators may either interact directly or indirectly with the basal transcription machinery (Ptashne and Gann, 1997; Tjian and Maniatis, 1994). Indirect interaction with the basal transcription machinery occurs through the association of DNA-binding transcription factors with a class of molecular 'bridging' proteins termed coactivators, which show no ability *per se* to directly bind DNA; examples of these are the CREB binding protein (CBP)/p300 family and SRC-1 (steroid receptor coactivator-1) family (Anzick *et al.*, 1997; Kamei *et al.*, 1996; Onate *et al.*, 1995; Shikama *et al.*, 1997). By interacting with DNA-binding transcription factors, transcriptional coactivators can recruit the basal transcription machinery to the promoter and thus facilitate activation of gene expression (Ptashne and Gann, 1997).

The silencing of gene expression is mediated by transcriptional repressors. Transcriptional repression can occur through a variety of *cis*- and *trans*-acting mechanisms that directly or indirectly interfere with the transcription machinery (eg Auble and Hahn, 1993). For example, direct competition between transcriptional repressors and DNA-binding components of the transcription pre-initiation complex can occur through the existence of overlapping cognate binding sites (Mitchell and Tjian, 1989).

The combination of *cis*-acting regulatory elements that constitute distinct transcription factor binding sites, together with the interaction of *trans*-acting factors provide the specificity of gene expression. However, it is becoming clear that chromatin structure plays an equally significant regulatory role and adds another level of complexity to the mechanisms of transcriptional regulation (Marcand *et al.*, 1996; Wolffe, 1994a). The accessibility of DNA to transcriptional regulators is determined by the state of chromatin condensation. Nucleosomes form the structural basis of condensed chromatin and are comprised of histones, around which DNA is packaged. Condensed chromatin is associated with transcriptional repression/silencing (Marcand *et al.*, 1996).

The acetylation status of histones is emerging as an important mechanism of controlling gene expression (Wolffe, 1994b; Pazin and Kadonaga, 1997; Sternglanz, 1996). Histone acetylation is associated with active transcription. The CBP/p300 coactivators possess intrinsic histone acetyltransferase (HAT) activity and can additionally interact with a HAT protein termed P/CAF (Ogryzko *et al.*, 1996; Bannister and Kouzarides, 1996). It is therefore apparent that the CBP/p300 coactivators not only contribute to recruitment of the transcription machinery, but also facilitate transcriptional activation by controlling the state of chromatin condensation.

1.5.2 DNA-binding domain

The c-Myb DNA-binding domain is located close to the amino terminus and constitutes a unique structure comprised of three imperfect tandem repeats of 51-52 amino acids, designated R1-3 (Figure 1.1A) (Sakura *et al.*, 1989; Gonda *et al.*, 1985). Within each repeat, three highly conserved tryptophan residues are present, spaced by 18 or 19 amino acids (Anton and Frampton, 1988). These periodically repeated tryptophans represent highly conserved residues among Myb-related proteins and are similarly found within the

DNA-binding domain of Ets transcription factors (Lipsick, 1996; Laget *et al.*, 1993). Mutation analysis of the c-Myb tryptophans by substitution for hydrophilic residues abolished or reduced DNA-binding activity, indicating that the tryptophan residues are essential for DNA-binding (Kanei-Ishii *et al.*, 1990).

Both c-Myb and v-Myb proteins bind to DNA monomerically in a sequence-specific manner, and recognise the core consensus sequence 5'-PyACC^{G}/$_{T}$G-3' (Weston, 1992; Howe and Watson, 1991; Garcia *et al.*, 1991; Nakagoshi *et al.*, 1990; Howe *et al.*, 1990; Nishina *et al.*, 1989; Biedenkapp *et al.*, 1988). The second (R2) and third (R3) repeats are essential for DNA-binding (Howe *et al.*, 1990; Sakura *et al.*, 1989). Each repeat is predicted to form three α-helices, which in turn are predicted to fold into a helix-turn-helix-related structure similar to those found in eukaryotic homeodomain proteins and bacterial repressors (Ogata *et al.*, 1992; Frampton *et al.*, 1989; Gonda *et al.*, 1985). Confirmation of this prediction has been provided by nuclear magnetic resonance (NMR) spectroscopy of R2 and R3 (Ogata *et al.*, 1994). The conserved tryptophan residues are located within a hydrophobic core as predicted, and serve to maintain the positions of each α-helix (Ogata *et al.*, 1992; Kanei-Ishii *et al.*, 1990). Sequence recognition is achieved cooperatively by specific contacts made by residues within the third helix of R2 and R3 (Ogata *et al.*, 1995; Ogata *et al.*, 1994). The R2R3 minimal binding domain lies within the major groove of DNA in a closely packed structure with interactions between the two repeats contributing to stability of the complex (Ogata *et al.*, 1995; Ogata *et al.*, 1994).

A highly conserved cysteine residue resides within R2 and is specifically located within the hydrophobic core when c-Myb is bound to DNA, which is in contrast to its exposed position when c-Myb is unbound (Lipsick, 1996; Ogata *et al.*, 1995; Myrset *et al.*, 1993). The R2 repeat is a structurally flexible domain that undergoes conformational changes upon DNA-binding (Ogata *et al.*, 1996; Ogata *et al.*, 1995; Myrset *et al.*, 1993). Mutation analysis reveals that the R2 cysteine residue influences DNA-binding activity. Substitution of this residue with serine reduces DNA-binding affinity (Myrset *et al.*, 1993; Guehmann *et al.*, 1992). Modification of the R2 cysteine may be an important regulatory event that influences the DNA-binding activities of c-Myb. In particular, the redox status of this cysteine may in part contribute to the regulation of DNA-binding as oxidation of this residue results in reduced DNA-binding activity by interfering with the conformational

changes of R2 that are associated with DNA-binding (Myrset *et al.*, 1993; Guehmann *et al.*, 1992).

The first repeat (R1) is structurally similar to R2 and R3, but in contrast is dispensible for sequence-specific DNA-binding (Ogata *et al.*, 1995; Howe *et al.*, 1990). Interestingly, R1 is highly conserved among vertebrate Myb proteins, but absent from most Myb-related proteins and partially absent from v-Myb (Lipsick, 1996). The role of R1 remains essentially unknown, however several hypotheses have been put forward. The first repeat may play a regulatory role and influence the overall binding properties of the DNA-binding domain such that the relative affinity for particular recognition sequences is altered. Examination of amino-truncated forms of Myb including v-Myb, suggest that removal of R1, and sequences amino-terminal to this repeat, decreases DNA-binding affinity, although these proteins retain transformation ability *in vitro* (Tanikawa *et al.*, 1993; Dini and Lipsick, 1993). Another study using amino-truncated proteins synthesised *in vitro* reported no difference in relative DNA-binding affinity, however this discrepency may be explained by specific protein modifications imposed by the nature of the system used to synthesise the protein (Ramsay *et al.*, 1995; Ramsay *et al.*, 1991). Deletion of R1, and proximal amino-terminal sequences which comprise known regulatory sites, may alter the expression pattern of genes regulated by c-Myb, resulting in a particular subset or wider array of genes being expressed. Consensus sequences with differing affinities for Myb binding may account for such a scenario (Tanikawa *et al.*, 1993; Ramsay *et al.*, 1992a; Howe and Watson, 1991; Garcia *et al.*, 1991; Ness *et al.*, 1989). The first repeat may influence DNA-binding by recognition of particular sequences that flank the core consensus sequence. In this context, recognition may either be direct, by interaction with the DNA, or indirect by association of cellular proteins with R1. Structural data indicate that R1 is only weakly associated with DNA and does not contribute significantly to Myb-binding by direct interaction with DNA (Ogata *et al.*, 1995; Ogata *et al.*, 1994). However, R1 does contribute to the formation of a stable, DNA-bound complex (Ogata *et al.*, 1995; Ogata *et al.*, 1994). One report highlights the potential association of cellular proteins with R1 that may influence direct DNA-binding mediated by R2R3, thus modulating the DNA-binding properties of Myb (Sala *et al.*, 1995).

1.5.3 Transactivation domain

The precise location of the c-Myb transcriptional activation domain has been defined by examination of c-Myb deletion mutants, which in some studies have been fused to a heterologous DNA-binding domain - that of the GAL4 yeast transcriptional activator (Kalkbrenner *et al.*, 1990; Weston and Bishop, 1989; Sakura *et al.*, 1989). The c-Myb transactivation domain is located centrally (Figure 1.1A) within a hydrophilic region containing a cluster of acidic residues, a common feature of activating regions (Tjian and Maniatis, 1994; Sigler, 1988).

The ability of c-Myb to transcriptionally activate gene expression was initially investigated by using reporter constructs containing multiple copies of the Myb-binding DNA sequence and minimal promoter elements, which creates a somewhat artificial context with respect to target sequences found *in vivo* (Nakagoshi *et al.*, 1990; Ibanez and Lipsick, 1990; Sakura *et al.*, 1989; Nishina *et al.*, 1989). Subsequently, natural promoters containing Myb-binding sites have been identified and used in transactivation assays to examine the properties and characteristics of Myb-mediated transcriptional regulation (eg Dudek *et al.*, 1992; Evans *et al.*, 1990; Ness *et al.*, 1989). Although this approach allows examination of transcriptional regulation in a more relevant context, it still lacks many attributes of physiological conditions such as the state of chromatin condensation, the effects of distant DNA elements such as enhancers and silencers, and other cell-specific transcription factors that may be absent from the cell line used to assay transcriptional regulation. Nevertheless, this approach has contributed to an understanding of the mechanisms by which Myb regulates gene expression.

Substitution of the v-Myb transactivation domain with the herpes simplex virus-1 VP16 transactivation domain indicates that the Myb transactivation domain is functionally distinct with properties that affect cellular transformation (Engelke *et al.*, 1995; Frampton *et al.*, 1993). The transactivation domain is required for Myb-enforced cellular transformation and implies that Myb exerts its effects by regulating gene expression (Hu *et al.*, 1991; Lane *et al.*, 1990).

1.5.4 Negative regulatory domain

The negative regulatory domain (NRD) of c-Myb was identified and defined by deletion analysis. Deletion of the NRD, located near the carboxyl terminus (Figure 1.1A) can enhance DNA-binding and transactivation activities (Ramsay *et al.*, 1992a; Ramsay *et al.*, 1991; Kalkbrenner *et al.*, 1990; Sakura *et al.*, 1989). In addition to the negative regulatory region defined by Sakura *et al.* (1989) (residues 325-500), a carboxyl-adjacent region (chicken c-Myb residues 494-553) has been shown to suppress c-Myb transactivation activity, indicating that sub-domains exist within the carboxyl-terminal region that contribute to negative regulation (Dubendorff *et al.*, 1992).

Initial examination of *in vitro* synthesised (rabbit reticulocyte lysate) truncated forms of c-Myb showed that carboxyl-truncation, which disrupts or deletes the NRD, increased DNA-binding activity (Ramsay *et al.*, 1991). This analysis was extended to examine the relative binding affinities of bacterially-expressed full-length and carboxyl-truncated c-Myb to oligonucleotides containing various Myb-binding sites (Ramsay *et al.*, 1992a). Carboxyl-truncated c-Myb binds to specific DNA sites with significantly increased affinity than full-length c-Myb, which demonstrated a low DNA-binding affinity (Ramsay *et al.*, 1992a). In contrast, purified full-length c-Myb expressed in eukaryotic cells binds to a Myb-responsive DNA element with high affinity (Krieg *et al.*, 1995). In this system, the effect of carboxyl-truncation on DNA-binding activity is less pronounced with a small increase in activity exhibited (Oelgeschlager *et al.*, 1995; Krieg *et al.*, 1995). The differences in DNA-binding affinity observed between full-length c-Myb expressed in bacterial or eukaryotic cells is probably attributed to the preparation procedure of the respective proteins (Krieg *et al.*, 1995). Due to the insoluble nature of bacterially-expressed full-length c-Myb, preparation required a denaturation/renaturation procedure which could consequently affect the activity of the purified protein (Krieg *et al.*, 1995; Ramsay *et al.*, 1992a).

The effect of Myb carboxyl truncation on relative DNA-binding activity has been further investigated using a wider range of truncated proteins in an assay that utilises unpurified eukaryotic-expressed proteins (Tanaka *et al.*, 1997). This analysis has identified two sub-domains of the NRD, designated NRD1 (residues 326-372) and NRD2 (residues 471-500), that suppress DNA-binding activity (Figure 1.1A) (Tanaka *et al.*, 1997). The mechanisms

by which NRD1 and NRD2 exert their effects remain unknown, however several potential phosphorylation sites have been identified within these sub-domains (Tanaka *et al.*, 1997).

Importantly, expression of carboxyl-truncated Myb proteins lacking the NRD sustains the proliferative capacity of primary haemopoietic cells and leads to increased transforming ability *in vitro* (Hu *et al.*, 1991; Grasser *et al.*, 1991; Gonda *et al.*, 1989a). These observations implicate the NRD as a functionally important region involved in the down-regulation of c-Myb activity. The carboxyl terminal region of c-Myb is more sensitive to proteolysis, implying an open structure that could facilitate protein associations which may consequently modulate c-Myb activity by interaction with the NRD (Krieg *et al.*, 1995). Overexpression of the c-Myb carboxyl-terminus *in vivo* implicates the role of protein interactions in the activity of this region (Vorbrueggen *et al.*, 1994; Dubendorff *et al.*, 1992). Several components of the NRD have been identified that contribute to the function of this complex domain. Key regulatory phosphorylation sites, for example the EVES motif, constitute identified components of the NRD and are discussed elsewhere in this review (section 1.5.5).

Carboxyl truncation leads to increased protein stability *in vivo* with respect to the short-lived wild type c-Myb protein (Bies and Wolff, 1997; Boyle *et al.*, 1985). The prolonged presence of a carboxyl-truncated Myb protein may contribute to the oncogenicity of such a protein. Degradation of c-Myb occurs by proteolysis through the action of the ubiquitin/26S proteasome pathway (Bies and Wolff, 1997). A carboxyl-truncated form of Myb that impinges on a substantial region of the NRD is less efficiently ubiquitinated, which may account for its increased stability *in vivo* (Bies and Wolff, 1997). This observation suggests the presence of specific sequences within the carboxyl terminus that mediate protein stability. Three regions of c-Myb, including the NRD, contain 'PEST' sequences, which are known to mediate protein degradation via the ubiquitin pathway (Bies and Wolff, 1997).

1.5.4.1 The c-Myb leucine zipper

The c-Myb leucine zipper motif is one component of the NRD that has been analysed more extensively and is beginning to provide insight into how the NRD functions to modulate the activity of c-Myb. The canonical leucine zipper motif spans a region of approximately

30-40 amino acids with a leucine residue repeated at every seventh position (O'Shea *et al.*, 1989). However, it is evident that other hydrophobic residues may occupy the positions commonly held by leucine residues (Landschulz *et al.*, 1988). Indeed this is the case with the c-Myb leucine zipper, which in addition to leucine residues has methionine and isoleucine residues (Figure 1.1A). The location of the methionine residue within the c-Myb leucine zipper motif comprises a polymorphic site in the corresponding DNA sequence, which can alternatively encode an isoleucine residue at this position (Favier and Gonda, 1994; Gonda *et al.*, 1985). Leucine zippers mediate protein-protein interactions and represent a form of heptad repeat (reviewed in Lupas, 1996). Heptad repeats contain repetitive hydrophobic and hydrophilic residues (termed positions *a-g* in heptad repeat nomenclature) that structurally form a bundle of α-helices, referred to as a coiled coil (O'Shea *et al.*, 1989; reviewed in Lupas, 1996). Leucine zippers form an amphipathic α-helical structure, with the repeated hydrophobic residues (generally leucine) at every seventh position (position *d* in heptad repeat nomenclature) aligned on the interaction face (O'Shea *et al.*, 1989; Landschulz *et al.*, 1988). Contrary to the name 'leucine zipper', dimerization occurs with each α-helix in a parallel orientation resulting in a coiled coil dimer, whereby a leucine residue from one helix interacts with a hydrophobic residue in the opposite helix (O'Shea *et al.*, 1989). Charged amino acid residues that flank the leucine zipper interaction face (positions *e* and *g* in heptad repeat nomenclature) contribute to dimerization stability and specificity by formation of salt bridges (Vinson *et al.*, 1993; Krylov *et al.*, 1994). The c-Myb leucine zipper predominantly contains hydrophobic residues in the positions that flank the interaction face, which may have effects on dimerization stability (Ebneth *et al.*, 1994). It is evident that the primary structure of the c-Myb leucine zipper motif contains several features that differ from the extensively studied classical leucine zipper motifs of the transcriptional regulators; GCN4, Jun and Fos (Ebneth *et al.*, 1994). Crystalisation of the c-Myb leucine zipper motif will be required to determine its structure.

The leucine zipper motif was first identified in the transcriptional regulators CCAAT/enhancer binding protein (C/EBP), GCN4, Jun, Fos and Myc (Landschulz *et al.*, 1988). These regulators are all DNA-binding proteins termed 'bZIP' proteins with reference to a region of basic residues located in the vicinity of the leucine zipper. In these proteins the leucine zipper indirectly contributes to DNA-binding by mediating protein

dimerization necessary for DNA-binding (reviewed in Pabo and Sauer, 1992). In contrast to the bZIP transcription factors, the c-Myb leucine zipper is distinct from the DNA-binding domain and is dispensible for DNA-binding (Figure 1.1A). The Myb proteins bind to DNA monomerically by virtue of a DNA-binding domain that functions as an intramolecular dimer (Lipsick, 1996).

Mutational analysis of the c-Myb leucine zipper has revealed that it is a crucial component of the NRD. Substitution of single and multiple hydrophobic heptad repeated residues for alanine or proline residues can affect the transactivating and transformation abilities of full-length c-Myb (Kanei-Ishii et al., 1992). Substitution of leucine residues 3 and 4 for proline markedly enhances transactivation of a reporter construct containing six concatemerised Myb-binding sites, and increased transforming capacity as assessed by haemopoietic cell colony assays (Kanei-Ishii et al., 1992). Substitution of leucine with proline is expected to completely disrupt the α-helical structure of the leucine zipper and consequently the ability to mediate protein-protein interactions. In addition, the severe effect of introducing proline substitutions into the leucine zipper motif may also structurally interfere with other regions of the protein. A carboxyl-truncated form of c-Myb that retains the leucine zipper has an increased binding affinity for DNA when the leucine zipper motif is mutated by substitution with proline residues, suggesting that an intact leucine zipper may contribute to the modulation of DNA-binding activity (Nomura et al., 1993). Substitution of leucine residues 3 and 4 for alanine, which is expected to maintain the α-helical structure, but disrupt critical hydrophobic interactions, also increased transactivation capacity, although to a lesser extent than the proline substituted mutant (Kanei-Ishii et al., 1992). Together these findings indicate that the leucine zipper can negatively regulate the activity of full-length c-Myb and implicate the association of interacting proteins as the mechanism of action.

A region spanning 10 amino acids and constituting part of the leucine zipper motif has been identified as a component required for the transactivation activity and transformation capacity of v-Myb[AMV] (Fu and Lipsick, 1996). This component, termed the 'FAETL' (single letter amino acid code) motif (Figure 1.1A), does not require an intact leucine zipper or the modulation of nearby phosphorylation sites to function. Disruption of the leucine zipper by substitution of heptad repeated residues with proline or alanine, similar to

20

the zipper mutants used by Kanei-Ishii *et al.* (1992), have differing effects on the transforming capacity of v-Myb^{AMV} (Fu and Lipsick, 1996). Substitution with prolines reduces, but does not abolish the ability to transform avian haemopoietic cells, indicating that the leucine zipper may specifically contribute to Myb function in a positive manner, or that the structural conformation of the protein itself is an important regulatory factor. It is proposed that the FAETL motif contributes to functional protein conformation by potential interaction with the DNA-binding domain (Fu and Lipsick, 1996). Furthermore, the FAETL motif may contribute to the transactivating activity of full-length c-Myb, but in this context, could be subject to negative regulation by other components of the carboxyl-terminus (that are absent from v-Myb^{AMV}) (Fu and Lipsick, 1996). Assessment of the transactivating and transforming capacities of a carboxyl-truncated Myb protein (CT3), lacking the leucine zipper and FAETL motifs, indicates that the FAETL motif is dispensible for transcriptional activation of reporter gene expression and transformation of murine haemopoietic cells *in vitro* (Hu *et al.*, 1991).

1.5.5 Phosphorylation and c-Myb activity

Protein phosphorylation constitutes an important and efficient method of regulatory control (Karin, 1994; Hunter and Karin, 1992). The c-Myb protein contains numerous phosphorylation sites, a few of which are known to modulate c-Myb activity (Figure 1.1A). Tryptic phosphopeptide analysis of c-Myb indicates common and lineage-specific phosphorylation sites, suggesting that differential phosphorylation may modulate the activity of c-Myb in a lineage-specific manner (Aziz *et al.*, 1995). The c-Myb protein is hyperphosphorylated during mitosis and subsequently binds to DNA less efficiently, which may affect transcriptional activity (Luscher and Eisenman, 1992).

A number of phosphorylation sites are clustered near the amino-terminus of c-Myb (Ramsay *et al.*, 1995). Serines 11 and 12 (serines 11/12) are phosphorylated *in vivo* and are phosphorylated by casein kinase II (CKII) *in vitro* (Oelgeschlager *et al.*, 1995; Luscher *et al.*, 1990). Phosphorylation of seines 11/12 by CKII can modulate DNA-binding activity. Phosphorylation of serines 11/12 is reported to have an inhibitory effect on DNA-binding of c-Myb to low affinity, but not to high affinity sites *in vitro* (Luscher *et al.*, 1990). Furthermore, substitution of serines 11/12 with alanine (non-phosphorylatable, conservative substitution) or asparagine (negatively-charged residue that mimics

phosphorylated state) residues indicates that phosphorylation at these sites reduces DNA-binding affinity and subsequently decreases transactivation activity (Oelgeschlager *et al.*, 1995). Another report describes the effect of phosphorylation of serines 11/12 as positively modulating DNA-binding activity (Ramsay *et al.*, 1995). In addition, phosphorylation at these sites was found to over-ride the negative effect of a wild type leucine zipper on DNA-binding activity (Ramsay *et al.*, 1995). These conflicting reports of the effect of phosphorylation of serines 11/12 on DNA-binding activity may reflect the nature of the proteins used in each study (ie full-length versus carboxyl-truncated, which lacks carboxyl-terminal phosphorylation sites) and/or the extraction conditions used to obtain the protein (Ramsay *et al.*, 1995). The modulation of DNA-binding activity by phosphorylation of serines 11/12 may be of particular importance, as oncogenically activated forms of Myb lack part of the amino-terminus encompassing serines 11/12.

Numerous conserved phosphorylation sites have been identified within and surrounding the NRD (Aziz *et al.*, 1993; Bading *et al.*, 1989a). These sites correspond to serine and threonine residues that are phosphorylated *in vivo* and have been demonstrated to be substrates for $p42^{MAPK}$ and GSK-3 (glycogen synthase kinase) *in vitro* (Aziz *et al.*, 1993; Woodgett, 1991). The consequential outcome of phosphorylation at these sites on c-Myb activity is poorly understood, however the absence of a number of these sites in v-Myb indicates that they may contribute to the action of the NRD, possibly by mediating inter- and intramolecular interactions. A number of serine and threonine GSK-3 phosphorylation sites are clustered near the leucine zipper. Substitution of these residues with alanine did not significantly affect the transactivation or transforming abilities of $v-Myb^{AMV}$ (Fu and Lipsick, 1996).

Murine c-Myb serine 528 (corresponding to chicken c-Myb serine 533) is phosphorylated by $p42^{MAPK}$ *in vitro* and can modulate the transcriptional activity of c-Myb on specific promoters (Miglarese *et al.*, 1996; Aziz *et al.*, 1993). Substitution of serine 528 with alanine increased transactivation of a reporter construct containing concatemerised *mim*-1A (a Myb target gene) binding sites (Aziz *et al.*, 1995). Examination of the transcriptional activity of this mutant using a broader range of natural promoters containing Myb-binding sites revealed that enhanced transactivation ability was dependent on the context of the promoter (Miglarese *et al.*, 1996). Therefore phosphorylation of serine 528 is implicated in

differentially repressing c-Myb transactivation of particular promoters and hence may consequently alter the array of target genes expressed. Differential repression of gene expression by phosphorylation of serine 528 may represent a regulatory mechanism that in part determines lineage-specific gene expression by c-Myb. Phosphorylation of serine 528 did not affect the ability of c-Myb to bind to a Myb-responsive element (Miglarese *et al.*, 1996).

The serine 528 phosphorylation site also constitutes part of the EVES motif (discussed in section 1.8) which can interact intramolecularly with the c-Myb DNA-binding domain. The mechanism by which phosphorylation of serine 528 down-regulates transcriptional activation is not understood, but it is hypothesised that a transcriptionally inactive protein conformation is formed by interaction of the phosphorylated EVES motif with the DNA-binding domain (Dash *et al.*, 1996). Alternatively, interacting proteins that contribute to transactivation activity, either as coactivators or repressors, may associate with the EVES motif. The differential responses to phosphorylation of specific residues within the c-Myb NRD implies that this region may serve to integrate cellular signals through the activity of protein kinases.

1.6 Genes regulated by Myb

Initial examination of the DNA-binding and transcriptional properties of the c- and v-Myb proteins implicated c-Myb as a transcription factor that functions to regulate the expression of specific genes (generically termed 'Myb target genes') involved in cellular proliferation and differentiation. The identification of *bona fide* target genes regulated by the Myb proteins has been difficult. To date, over 30 genes are recognised as Myb target genes with the majority designated as putative (Ness, 1996). Particular examples of Myb target genes will be discussed in order to highlight identification strategies and the information gained from subsequent studies of Myb-regulated gene expression.

Direct and indirect strategies have been employed in attempts to identify Myb target genes. Direct approaches have utilised conditional Myb mutants and/or differential screening (Lin *et al.*, 1997; Burk *et al.*, 1997; Frampton *et al.*, 1996; Ness *et al.*, 1989) The utilisation of a temperature sensitive mutant of v-Myb^{E26} in conjunction with a differential screen led to the identification of the first Myb target gene, a novel chicken promyelocyte-specific gene

termed *mim*-1 (*myb*-induced myeloid protein 1) which encodes a protein of unknown function located in the granules of promyelocytes (Ness *et al.*, 1989). The *mim*-1 promoter contains three functional Myb-binding sites of differing affinities which can facilitate Myb-mediated transcriptional activation of a reporter gene (Ness *et al.*, 1993; Ness *et al.*, 1989). Interestingly, *mim*-1 is not expressed in monoblasts transformed by v-Myb^{AMV}, implying that expression of *mim*-1 is not required for cellular transformation; rather the *mim*-1 gene product contributes to a particular stage of granulocyte differentiation (Ness *et al.*, 1989). The point mutations that reside primarily within the DNA-binding and transactivation domains of v-Myb^{AMV} (Figure 1.1B) may account for the lack of *mim*-1 expression in monoblasts transformed by AMV. The expression of *mim*-1 is restricted to a subset of cell types in which c-*myb* is expressed, thus implicating the action of cell type-specific factors in the regulation of *mim*-1 (Ness *et al.*, 1993; Ness *et al.*, 1989). The molecular mechanisms by which c-Myb achieves cell maturation-specific expression of *mim*-1 are discussed below (section 1.6.1).

Myeloblasts transformed by the v-Myb^{E26} temperature-sensitive mutant (ts21) undergo differentiation into macrophages when shifted to the non-permissive temperature (Beug *et al.*, 1987). The observation that a proportion of cells undergo apoptosis when shifted to the non-permissive temperature stimulated a further examination of the activity of the Myb-Ets fusion protein encoded by the E26 mutant and the expression of *bcl*-2, a survival gene that protects cells from apoptosis (Frampton *et al.*, 1996). In particular, cell death is attributed to inactivation of v-Myb$^{E26(ts21)}$ activity at the non-permissive temperature (Frampton *et al.*, 1996). The expression of *bcl*-2 is down-regulated when v-Myb$^{E26(ts21)}$ activity is abrogated by shifting to the non-permissive temperature. Furthermore, *bcl*-2 expression can be re-induced by reactivating v-Myb$^{E26(ts21)}$ by shifting back to the permissive temperature. The P2 promoter of the *bcl*-2 gene harbours several Myb-binding sites which are required for Myb-Ets-mediated transcriptional activation of a reporter construct linked to this promoter region (Frampton *et al.*, 1996).

Several reports confirm *bcl*-2 as a Myb target gene (Salomoni *et al.*, 1997; Hogg *et al.*, 1997; Taylor *et al.*, 1996). Induction of a dominant-negative form of Myb in a murine thyoma cell line correlates with down-regulation of *bcl*-2 expression and results in a large proportion of cells undergoing apoptosis (Taylor *et al.*, 1996). Similarly, a dominant-

negative Myb transgene expressed in the murine T-cell lineage conferred a greater susceptibility of these cells to undergo apoptosis (Badiani *et al.*, 1994). Collectively, these studies illustrate the emerging role of Myb proteins in the survival of particular cell types by transcriptional regulation of key survival genes such as *bcl*-2 and potentially other members of this family. Deregulation of such survival genes by oncogenically activated forms of Myb could contribute to cellular transformation.

Several Myb target genes have been indirectly identified by association of their functions with cellular proliferation. For example, the association of c-*myb* expression with the G1 to S phase transition in T-lymphocytes prompted investigations of genes known to function in this context, namely *cdc2* (cyclin-dependent serine/threonine kinase) and DNA polymerase α (Ku *et al.*, 1993; Sudo *et al.*, 1992; Venturelli *et al.*, 1990). Analysis of the promoter sequences of these genes identified putative Myb-binding sites. Subsequent binding analyses and transactivation studies ensued. Such studies showed that the *cdc2* gene is transcriptionally regulated by c-Myb in T-lymphocytes, however in the case of the DNA polymerase α gene, c-Myb was found to bind to the promoter sequence but did not transactivate reporter gene expression (Ku *et al.*, 1993; Sudo *et al.*, 1992).

Identified Myb target genes include CD34 (Melotti *et al.*, 1994; Melotti and Calabretta, 1994), neutrophil elastase (Nuchprayoon *et al.*, 1997; Oelgeschlager *et al.*, 1996), c-*kit* (stem cell factor receptor) (Hogg *et al.*, 1997; Melotti and Calabretta, 1996; Ratajczak *et al.*, 1992), c-*myc* (Cogswell *et al.*, 1993; Nakagoshi *et al.*, 1992; Evans *et al.*, 1990), T-cell receptor δ (Hernandez-Munain and Krangel, 1995; Hernandez-Munain and Krangel, 1994), CD13 (aminopeptidase) (Shapiro, 1995), *lck* (tyrosine kinase) (McCracken *et al.*, 1994), CD4 (Siu *et al.*, 1992), myeloperoxidase (Britos-Bray and Friedman, 1997), *GBX2* (homeobox gene) (Kowenz-Leutz *et al.*, 1997) and *Pax*-6 (quail; Pax-QNR) (Plaza *et al.*, 1995). Some of these Myb target genes have been more extensively studied than others which remain putative targets of Myb regulation. A small number of Myb target genes have been analysed in the context of direct regulation of the corresponding chromosomal genes *in vivo*. Most studies have identified Myb-binding sites within the promoter regions or enhancer elements of genes and performed analyses that aim to demonstrate c-Myb DNA-binding and transcriptional regulation. In this context, transcriptional regulation is usually assessed by reporter assays using ectopically expressed Myb in cell lines. The regulation of

putative Myb target genes has been further investigated by using a conditional Myb mutant in primary murine haemopoietic cells (Hogg *et al.*, 1997). This system utilises an oestrogen analogue to activate a carboxyl-truncated Myb protein fused to the ligand-binding domain of the oestrogen receptor. Examination of the expression of *cdc2* and c-*myc* in this system indicates that these genes may not be regulated by Myb in myeloid cells (Hogg *et al.*, 1997). Inactivation of Myb did however correlate with a decrease in the expression of c-*kit* and *bcl*-2. Examination of *bcl*-2 expression supports previous studies, although no apparent increase in cells undergoing apoptosis was observed upon inactivation of Myb (Hogg *et al.*, 1997). These observations suggest that Myb may not suppress apoptosis in myeloid cells by regulation of *bcl*-2 expression (Hogg *et al.*, 1997). Potential Myb-binding sites have been identified in the c-*kit* promoter, which can be transactivated by c-Myb in reporter assays, thus identifying c-*kit* as a Myb target gene (Hogg *et al.*, 1997).

A transcriptional repression function of c-Myb was initially suggested by reporter assays with an artificial promoter containing weak Myb-binding sites (Nakagoshi *et al.*, 1989). The promoters of c-*fms* (CSF-1 receptor) and c-*erbB*-2 (similarities to EGF receptor) direct c-Myb-mediated transcriptional repression of reporter constructs (Mizuguchi *et al.*, 1995; Reddy *et al.*, 1994). Repression of the c-*erbB*-2 promoter by c-Myb probably involves direct competition with TFIID binding, as one of the identified Myb-binding sites overlaps the TATA element. Furthermore, the c-Myb DNA-binding domain alone can mediate this repression (Mizuguchi *et al.*, 1995). Repression of the c-*fms* promoter by c-Myb implicates the action of repressor domain located within the carboxyl-terminal region, as carboxyl-truncated Myb lacking a significant proportion of the NRD, was unable to mediate transrepression of a reporter construct in a murine macrophage cell line (Reddy *et al.*, 1994).

A number of recognised Myb target genes represent markers of specific stages of haemopoietic cell maturation, for example *mim*-1 and CD34. A predictive analysis of potential Myb-regulated promoters based on Myb-binding affinity measurements indicates that many known genes including other transcription factors, signal transduction components and cell cycle regulators harbour potential Myb-binding sites and remain uninvestigated in the context of Myb regulation (Deng *et al.*, 1996).

1.6.1 Transcription factor cooperation with Myb

The *mim*-1 gene has provided a basis for many investigations into the molecular mechanisms underlying the regulation of gene expression by Myb proteins. The restricted expression of *mim*-1 in neutrophilic granulocytes is achieved by cooperation of Myb with NF-M (nuclear factor-myeloid/C/EBPβ/NF-IL6), a tissue-specific transcription factor expressed in myeloid cells. The *mim*-1 promoter additionally contains NF-M-binding sites, one of which is located in close proximity to the Myb-binding sites. Together these juxtaposed sites constitute a composite response element (Mink *et al.*, 1996). Both Myb and NF-M bind directly to the *mim*-1 promoter and cooperatively transactivate reporter gene expression (Ness *et al.*, 1993; Burk *et al.*, 1993). Ectopic coexpression of both Myb and NF-M in non-myeloid cells can induce expression of *mim*-1 and other myeloid-specific markers, indicating that the cooperativity of Myb with NF-M serves as a combinatorial signal for expression of myeloid-specific genes (Ness *et al.*, 1993; Burk *et al.*, 1993). A direct interaction between v-Myb^{AMV} and C/EBPβ has been demonstrated (Mink *et al.*, 1996). This interaction is mediated by the DNA-binding domains of each protein and is required for cooperative transactivation of the *mim*-1 promoter (Mink *et al.*, 1996). Although the *mim*-1 gene is not expressed in monoblasts transformed by AMV, v-Myb^{AMV} and NF-M are able to interact and cooperatively activate reporter gene expression (Mink *et al.*, 1996). The basis of this paradox remains obscure, however, the ability of NF-M and v-Myb^{AMV} to cooperatively activate reporter gene expression may be a consequence of the assay conditions (eg ectopic protein expression). Alternatively, these observations may imply that additional cell-specific transcription factors are involved in regulating the expression of *mim*-1. The combinatorial activation of *mim*-1 can be further enhanced by interaction with CBP, a transcriptional coactivator that interacts with the Myb transactivation domain (see section 1.8) (Oelgeschlager *et al.*, 1996; Dai *et al.*, 1996). Expression of the *mim*-1 gene therefore involves an activator complex, whereby Myb and NF-M bind directly to the *mim*-1 promoter and to each other, possibly to stabilise this protein organisation, and additionally recruit CBP which is able to interact with components of the transcription pre-initiation complex (Oelgeschlager *et al.*, 1996). Interestingly, the reduced c-Myb DNA-binding and transcativation activity on the *mim*-1 promoter, as a result of casein kinase II phosphorylation of serines 11 and 12, can be partially compensated by ectopic expression of NF-M (Oelgeschlager *et al.*, 1995).

A novel Myb target gene has recently been identified that is highly expressed in myeloid cells transformed by either AMV or E26, unlike the *mim-*1 gene which is not expressed in AMV-transformed myeloid cells (Burk *et al.*, 1997; Introna *et al.*, 1990). Sequence analysis of this novel gene, designated *tom-*1 (target of myb 1), reveals that it has two different promoters (A and B) that give rise to two transcripts with differing 5', but common 3' exons (Burk *et al.*, 1997). The *tom-*1A promoter contains several Myb-binding sites and an adjacent C/EBP binding site in a similar context to that of the *mim-*1 promoter. Mutation and transactivation analyses of the *tom-*1A promoter demonstrate that v-Myb and members of the C/EBP family cooperatively activate reporter gene expression (Burk *et al.*, 1997). Interestingly, activation of the *tom-*1 promoter is achieved by cooperation of v-Myb with C/EBPδ and to a lesser extent with C/EBPα, which is in contrast to activation of the *mim-*1 promoter. The activation of these promoters by cooperation of Myb with different C/EBP isoforms indicates that myeloid-specific gene expression may be regulated by a common association of Myb with C/EBP members, which individually show distinct patterns of expression during myeloid differentiation (Burk *et al.*, 1997).

Transduction of both Myb and Ets encoding sequences by the E26 acute leukaemia virus, and the ability of this retrovirus to predominantly induce erythroleukaemia *in vivo*, in contrast to AMV, prompted investigations of the functional relationship between the c-Myb and Ets haemopoietic transcription factors (Dudek *et al.*, 1992). Members of the Ets family of transcription factors have been shown to bind to the *mim-*1, *lck* type 1 and CD13/aminopeptidase promoters, and cooperate with c-Myb to synergistically activate transcription in reporter assays (Shapiro, 1995; McCracken *et al.*, 1994; Dudek *et al.*, 1992). A recent report demonstrates that cooperation of c-Myb with Ets is required to overcome transcriptional repression of the α4 integrin gene mediated by ZEB, a vertebrate zinc-finger/homeodomain protein with characterised functions in muscle and possibly neural tissue (Postigo *et al.*, 1997). In addition to its role in myogenesis, α4 integrin functions in haemopoiesis where it facilitates interaction with the stroma, and is involved in targeting lymphoid and myeloid cells to sites of inflammation (Lobb and Hemler, 1994). The mechanism by which c-Myb and Ets overcome ZEB-mediated transcriptional repression is distinct from that in myogenesis which involves the displacement of ZEB from the promoters of muscle-specific genes by myogenic basic helix-loop-helix proteins (Postigo *et al.*, 1997). Individually, neither c-Myb or Ets can relieve ZEB-mediated

28

repression of the α4 integrin promoter, but cooperatively, c-Myb and Ets can synergistically activate transcription from this promoter in reporter assays (Postigo *et al.*, 1997).

It is apparent that c-Myb can act either in a DNA-binding dependent or independent manner to regulate gene expression. Heat shock transcription factor (HSF) 3 can interact with R2R3 of the c-Myb DNA-binding domain (Kanei-Ishii *et al.*, 1997). This association results in activation of HSF3 and transcription of stress-responsive genes, in unstressed proliferating cells (Kanei-Ishii *et al.*, 1997). Activation of gene expression from the heat-shock responsive *hsp70* promoter occurs in a context that is independent of direct DNA-binding by c-Myb (Kanei-Ishii *et al.*, 1994). Direct DNA-binding is achieved by the interaction of HSF3 with its cognate recognition sequence (heat shock element) and c-Myb (Kanei-Ishii *et al.*, 1997). Deletion of the c-Myb transactivation domain did not affect transactivation of the *hsp70* promoter (Kanei-Ishii *et al.*, 1994). Mutation of the leucine zipper within the c-Myb negative regulatory domain can destabilise the interaction between HSF3 and c-Myb, and furthermore abolish transactivation of the *hsp70* promoter (Kanei-Ishii *et al.*, 1997; Kanei-Ishii *et al.*, 1994). These points suggest that c-Myb may recruit additional proteins through interaction with the leucine zipper, that contribute to the formation of a stable, transcriptionally active complex on the *hsp70* promoter.

Similarly, c-Myb can cooperate with BZLF1, an Epstein-Barr virus (EBV) transcriptional activator, in the absence of direct DNA-binding by Myb to synergistically activate the EBV and SV40 early promoters in lymphoid cells (Kenney *et al.*, 1992). Cooperative activation of these promoters is achieved in some cell types but not others, therefore implicating the activity of cell type-specific factors (Kenney *et al.*, 1992).

Transcription factor cooperativity with c-Myb is fastly becoming an emerging theme for myeloid- and lymphoid-specific gene expression (Nuchprayoon *et al.*, 1997; Britos-Bray and Friedman, 1997; Oelgeschlager *et al.*, 1996; Hernandez-Munain and Krangel, 1994). The mechanisms by which Myb can cooperate with other transcription factors to activate cell type-specific gene expression remain largely unresolved.

1.7   Oncogenic Activation

Both v-Myb of the AMV and E26 acute leukaemia viruses have truncated amino and carboxyl termini with respect to c-Myb (Figure 1.1B). Most of the first repeat (R1) is absent from v-Myb of AMV and E26. Almost the entire NRD is absent from v-Myb^{E26} whereas v-Myb^{AMV} has a partially deleted NRD, retaining the leucine zipper motif. In addition to the oncogenically activated Myb proteins of AMV and E26, retroviral insertional mutagenesis of the c-Myb locus has given rise to amino- and/or carboxyl-truncated forms of Myb that are associated with haemopoietic malignancies (reviewed in Gonda, 1991; Shen-Ong, 1990). Retroviral overexpression of a carboxyl-truncated Myb protein *in vivo* (in chickens) lead to the appearance of fibrosarcomas after a long period of latency, suggesting that other genetic lesions in addition to the overexpressed Myb protein were involved in promoting cellular transformation (Press *et al.*, 1994). Similarly, low-incidence avian B cell lymphomas have been induced by overexpression of an identical carboxyl-truncated Myb protein (Press *et al.*, 1995). A large proportion of lymphomas were not attributed to overexpression of carboxyl-truncated Myb, but rather were a consequence of retroviral insertional mutagenesis of the 5' end of the chromosomal c-*myb* gene that resulted in the expression of an amino-terminal truncated protein (Press *et al.*, 1995).

Truncation of c-Myb is correlated with transformation of murine and chicken haemopoietic cells *in vitro*, revealing the oncogenic potential of this protein (Dini and Lipsick, 1993; Hu *et al.*, 1991; Grasser *et al.*, 1991; Gonda *et al.*, 1989a; Gonda *et al.*, 1989b). Furthermore, the type (ie amino and/or carboxyl) and extent of c-Myb truncation are associated with distinct myeloid cell phenotypes representative of particular stages of maturation (Grasser *et al.*, 1991; Introna *et al.*, 1990). These observations indicate that differing oncogenically activated forms of c-Myb, that retain the minimal DNA-binding and transactivation domains, enforce a particular cell maturation phenotype. Oncogenic activation of c-Myb by protein truncation could conceivably result in the alteration of regulatory activities, by disruption of functional domains and sites. In view of the functional activities of c-Myb, it is probable that alteration of the regulatory activities of this protein may consequently impart changes in target gene expression that could account for the distinct cell phenotypes enforced by expression of truncated forms of Myb.

It has been hypothesised that c-Myb regulates two classes of genes, namely those that pertain to proliferation and to differentiation, on the basis of differential binding affinity (Dini and Lipsick, 1993). This hypothesis proposes that oncogenically activated forms of Myb alter gene expression by regulating a subset of genes, in particular, those that contribute to cellular proliferation rather than differentiation. Although very few Myb target genes have been identified, the majority of recognised Myb target genes specify markers of differentiation. Consistent with this observation, it has been proposed that cellular transformation by Myb (including c-Myb) is achieved by maintaining a particular state of cell maturation that is responsive to haemopoietic growth factors, which provide proliferative signals (Ness, 1996). It is evident that Myb-transformed cells are dependent on exogenous haemopoietic growth factors *in vitro* in order to sustain survival and proliferation (Ferrao *et al.*, 1995; Macmillan and Gonda, 1994; Gonda *et al.*, 1993; Grasser *et al.*, 1991; Gonda *et al.*, 1989a; Gonda *et al.*, 1989b; reviewed in Ness, 1996; Gonda, 1991). Therefore, it appears that cellular transformation by Myb is achieved in part by cooperation with proliferative signals transduced by haemopoietic growth factors.

The oncogenically activated v-Myb of AMV contains 11 point mutations, with respect to c-Myb, that give rise to amino acid substitutions (Figure 1.1B) (Klempnauer *et al.*, 1982). Several studies have investigated the effect of these mutations on the DNA-binding, transactivating and transforming abilities of Myb (Brendeford *et al.*, 1997; Dini *et al.*, 1995; Dini and Lipsick, 1993; Introna *et al.*, 1990). Examination of a series of v-Myb[AMV] chimeras constituting differing combinations of the spectrum of AMV substitutions reveals that mutations in both the DNA-binding and transactivation domains cooperatively contribute to oncogenicity (Dini *et al.*, 1995). These chimeric proteins display differing transactivating and transforming abilities, but similar relative DNA-binding activities on the *mim*-1 promoter (Dini *et al.*, 1995). These observations imply that the point mutations within v-Myb[AMV] potentially alter interactions with cellular proteins that contribute to Myb-mediated transcriptional activation (Dini *et al.*, 1995). In support of this notion, the substitutions in the DNA-binding domain reside on the solvent-exposed face and could therefore conceivably affect potential protein-protein interactions (Ogata *et al.*, 1995; Ogata *et al.*, 1994).

Reversion of the three substitutions that reside within R2 of the DNA-binding domain to wild type residues result in an altered transformed myeloid cell phenotype (promyelocyte) and expression of the promyelocyte-specific *mim-1* gene, which is not usually expressed in AMV-transformed cells (monoblasts) (Introna *et al.*, 1990). In particular, these mutations affect structural and functional properties of R2 (Brendeford *et al.*, 1997). The AMV-specific mutations impose a structurally compact conformation of R2 which interferes with the redox status of the putative regulatory cysteine residue located within this repeat (Brendeford *et al.*, 1997; Myrset *et al.*, 1993; Guehmann *et al.*, 1992). Furthermore, the stability of R2R3 (Myb minimal DNA-binding domain) DNA-bound complexes is affected by AMV-specific mutations within R2 that ultimately result in quantitative differences in DNA-binding to particular Myb-responsive elements (Brendeford *et al.*, 1997). Therefore, in the context of v-Myb^{AMV}, the point mutations located in R2 may affect the DNA-binding properties of this protein and consequently alter the repertoire of genes expressed with respect to c-Myb. Such possible changes in gene expression could partly account for the observed phenotype of AMV-transformed myeloid cells.

1.8 Cellular Proteins that Interact with Myb

Several proteins have been identified as Myb-interacting proteins (Figure 1.1A). Analyses of the interactions between Myb and cellular proteins are providing a mechanistic understanding of how c-Myb regulates gene expression. The Myb proteins are known to cooperatively activate gene expression (discussed in section 1.6.1). In some cases, cooperative activation of gene expression has been shown to involve a direct interaction between Myb and another transcription factor (ie C/EBPβ and HSF3) (see section 1.6.1). Both C/EBPβ and HSF3 interact with the Myb DNA-binding domain. The c-Myb DNA-binding, transactivation and negative regulatory domains each associate with specific cellular proteins. These interactions are discussed below.

A region spanning the first amino-terminal 192 residues of chicken c-Myb can interact in an intramolecular manner with the carboxyl terminus (Figure 1.1A). This interaction is mediated by the 'EVES (Glu-Val-Glu-Ser residues) motif' ('AVES' in murine c-Myb), a sequence located within the negative regulatory domain that also constitutes a phosphorylation site (Dash *et al.*, 1996). A subsequent search for proteins containing the EVES motif identified p100, a transcriptional coactivator known to interact with the viral

transcription factor EBNA-2, and TFIIE (a component of the RNA polymerase II transcription pre-initiation complex) (Dash *et al.*, 1996). The p100 coactivator can associate via the EVES motif with the c-Myb DNA-binding domain and influence reporter gene expression (Dash *et al.*, 1996). It has been proposed that c-Myb activity may in part be regulated by competition between intramolecular and intermolecular interactions in a phosphorylation-dependent manner, such that cooperative factors required for c-Myb transcriptional activity are excluded by intramolecular association of the DNA-binding domain with the carboxyl terminus (Dash *et al.*, 1996).

The identification of proteins that interact with the Myb transactivation domain provide insight as to how transcriptional activation is mediated. The transcriptional coactivators CBP/p300 are known to interact with a multitude of proteins that largely include transcription factors and viral oncoproteins (Shikama *et al.*, 1997). These coactivators function in a regulatory context and are believed to coordinate gene expression in response to a broad spectrum of signals. The CBP coactivator has been shown to interact with the c-Myb transactivation domain in a phosphorylation-independent manner, unlike the interaction with CREB, suggesting that the interaction of CBP with c-Myb may be constitutive (Oelgeschlager *et al.*, 1996; Dai *et al.*, 1996). Furthermore, this interaction with CBP enhances Myb-dependent transactivation, signifying that CBP is also a coactivator of c-Myb and presumably acts as a 'bridging' protein by promoting recruitment of the transcription machinery, a common theme in transcriptional activation (Ptashne and Gann, 1997; Oelgeshlager *et al.*, 1996; Dai *et al.*, 1996).

1.8.1 Proteins that interact with the c-Myb leucine zipper

The leucine zipper motif may mediate homodimerization. It is proposed that c-Myb is negatively autoregulated by the formation of homodimers which reduces transactivation ability, possibly by preventing DNA-binding (Nomura *et al.*, 1993). This study used a carboxyl-truncated form of Myb (retaining the leucine zipper motif, but lacking peptides from the NRD) to examine homodimerisation. Others have reported an inability of full-length c-Myb to homodimerize *in vitro* and propose that the action of the leucine zipper does not attenuate DNA-binding activity (Krieg *et al.*, 1995). Examination of a synthetic peptide encompassing the c-Myb leucine zipper indicates that this region can form an α-helical structure in solution, but only under conditions that stabilise secondary structure

(Ebneth *et al.*, 1994). Furthermore, the formation of dimers was not detected under conditions that facilitated formation of an α-helical structure and raise the possibility that the c-Myb leucine zipper motif may not be acting in the classical context of a coiled coil (Ebneth *et al.*, 1994). However, distant residues that flank the leucine zipper motif and absent from the peptide used in the aforementioned study may contribute to stabilisation of secondary and tertiary structure *in vivo* (Ebneth *et al.*, 1994).

Two related nuclear proteins that bind to the c-Myb leucine zipper, but not to mutated forms, have been detected in murine cell lines (Favier and Gonda, 1994). These proteins, termed p160 and p67, form the subject of this thesis and are discussed in more detail below and in Chapters 3, 4 and 5. Several other unidentified nuclear proteins of 96, 130, 90 and 160 kDa that specifically associate with a wild type c-Myb leucine zipper have been detected in HeLa cell extracts (Kanei-Ishii *et al.*, 1992).

1.8.1.1 The related p160 and p67 proteins

The p160 and p67 proteins were detected in radiolabelled nuclear extracts from murine cell lines using an affinity chromatography approach and are termed p160 and p67 on the basis of their observed molecular weights. Both p160 and p67 bound to a bacterially-expressed GST (Glutathione S-transferase) fusion protein containing a region of the NRD that encompasses the leucine zipper motif (Favier and Gonda, 1994). The interaction of p160 and p67 specifically with the c-Myb leucine zipper motif was indicated by the inability of p160 and p67 to interact with similar fusion proteins containing mutated leucine zipper motifs. These leucine zipper mutants contain proline or alanine substitutions of critical leucine residues in the third and fourth *d* positions of the heptad repeat, and correspond to those mutations described by Kanei-Ishii *et al.* (1992).

Comparative peptide mapping of p160 and p67 showed that these proteins are closely related (Favier and Gonda, 1994). However, each protein has distinct properties. The p160 and p67 proteins were initially detected in nuclear extracts from the murine FDC-P1 cell line, which expresses c-Myb and has a phenotype that resembles a myelomonocytic progenitor. Examination of a panel of predominantly murine haemopoietic cell lines demonstrated that p160 is expressed ubiquitously in a context that is not dependent on the presence of c-Myb (Favier and Gonda, 1994). Conversely, p67 is restricted to a subset of

immature myeloid cell lines with myelomonocytic progenitor characteristics, and was absent from some cell lines which do express c-Myb. Interestingly, neither p160 or p67 were detected in the limited number of human cell lines examined. This is somewhat surprising based on the large degree of sequence conservation between murine and human c-*myb* (Favier and Gonda, 1994; Majello *et al.*, 1986).

The p160 protein was additionally found to interact with a GST fusion protein containing the c-Jun basic-leucine zipper (bZIP) region (Favier and Gonda, 1994). In contrast, p67 was unable to bind to this region of c-Jun (Favier and Gonda, 1994). The ubiquitous distribution of p160 amongst murine cell lines, and the ability of this protein to interact with the c-Myb leucine zipper motif and c-Jun bZIP region suggests a potentially wider functional role for p160.

In order to identify the p160 and p67 Myb-binding proteins and further characterise their properties, a cloning approach was undertaken (Gonda *et al.*, 1996). Following purification of a quantity of p67 from FDC-P1 cells, amino acid sequence information was derived from five tryptic peptides generated from this protein (1. PDQETRLAT 2. VILRSPK 3. VLEEG 4. VDHLHLEK 5. ATPQIPETK; single letter amino acid code) (Gonda *et al.*, 1996) (Tavner *et al.*, 1998; see appendix). Degenerate oligonucleotides corresponding to four of the five tryptic peptide sequences were subsequently used to obtain p67 cDNA from FDC-P1 cells by PCR. Three overlapping cDNA clones containing open reading frames (ORFs) were isolated using the sequence information derived from p67 in conjunction with PCR and cDNA library screening (Tavner *et al.*, 1998; see appendix). A single 3135 bp cDNA clone was isolated from a WEHI-3B cDNA library, representative of a cell line in which both p160 and p67 were detected. Sequence analysis of this 3.1 kb clone revealed a long ORF encoding a polypeptide of 1010 amino acids with a predicted molecular mass of 114 kDa. This sequence contains a putative translation initiation site (Kozak consensus), termination codon, and of two of the five p67-derived tryptic peptide sequences. It was evident that this 3.1 kb cDNA clone could direct translation of a protein substantially larger than p67, but smaller than p160, as demonstrated by *in vitro* translation and predicted by sequence analysis (see Chapter 3.2). These observations suggested that this cDNA clone may be incomplete. A further analysis of this clone was required to ascertain its relationship with respect to the p67 and p160 proteins.

1.9   The Objectives of this Project

The association of p160 and p67 with the c-Myb negative regulatory domain by interaction with the leucine zipper motif implicates these proteins as potential regulators of c-Myb activity. It is hypothesised that p160 and p67 repress the functional activities of c-Myb, namely DNA-binding and/or transactivation, as disruption or deletion of the leucine zipper motif enhances these activities.

The aims of this project were:

(1) To complete the molecular cloning of p160 and p67, and characterise the properties of these proteins.

(2) To determine the relationship between p160 and p67.

1.10   Regulation of c-*myb*; A Summary

The regulation of c-*myb* occurs by both transcriptional and post-translational themes. This review has focused on post-translational mechanisms of regulation. Characterisation of the transcriptional control of c-*myb* gene expression demonstrates that several modes of action are involved. The *cis*-acting regulatory sequences controlling c-*myb* expression define two promoter regions that encompass 5' flanking sequences and the first intron of this locus (Jacobs *et al.*, 1994). Both positively and negatively acting regulatory elements have been identified within these regions. The down-regulation of c-*myb* expression during cell maturation to terminal differentiation is mediated by a transcriptional pause site located within the first intron and is responsible for blocking transcription elongation (Boise *et al.*, 1992; Reddy and Reddy, 1989; Watson, 1988; Bender *et al.*, 1987).

Several known transcription factors have been identified as regulators of c-*myb* expression. Expression of human c-*myb* is regulated by various transcription factors in haemopoietic cells of different lineages via distinct *cis*-acting elements located within the 5' promoter (Sullivan *et al.*, 1997). In lymphocytes, the WT1 (Wilm's tumour suppressor protein) transcriptional repressor has been reported to bind to a site within the 5' promoter of the c-*myb* gene and mediate repression of reporter constructs containing this region (McCann *et al.*, 1995). An inducible, unidentified protein termed CMAT (c-*myb* in activated T-cells) binds to a region within the 5' promoter of the c-*myb* gene in activated, but not resting T-cells (Phan *et al.*, 1996). Members of the NF-κB family bind to sites within the first intron

and have been shown to activate gene expression in reporter assays (Toth *et al.*, 1995). Interestingly, in fibroblasts, c-*myb* expression appears to be positively autoregulated by Myb-binding sites located within the 5' promoter (Nicolaides *et al.*, 1991). In contrast, these Myb-binding sites mediate negative autoregulation of c-*myb* expression in T-cells (Guerra *et al.*, 1995). The *cis*-acting elements and *trans*-acting regulatory factors that control c-*myb* expression remain largely uncharacterised.

CHAPTER TWO
Materials and Methods

2.1 MATERIALS

2.1.1 Enzymes

Restriction enzymes were purchased from several companies including Boehringer Mannheim, Germany; New England Biolabs, USA; Pharmacia Biotech, Sweden; Promega Corporation, USA.

Other enzymes were purchased from the following companies:

Calf intestinal alkaline phosphatase (CIP): Pharmacia Biotech, Sweden.

T4 DNA Ligase, Klenow DNA polymerase: New England Biolabs, USA.

T3, T7 and SP6 RNA polymerases: Promega Corporation, USA.

T7 DNA polymerase: Bresatec Ltd., Australia.

SuperScript reverse transcriptase: Gibco BRL, USA.

Taq DNA polymerase: Perkin Elmer, USA.

Pfu DNA polymerase (cloned): Stratagene, USA.

2.1.2 Radiochemicals

Radiochemicals were purchased from the following companies:

α-[^{32}P] dATP: Bresatec Ltd., Australia.

α-[^{33}P] dATP; Methionine, L-[^{35}S]: NEN Dupont, USA.

TRAN^{35}S-Label (containing 70% [^{35}S] L-methionine): ICN Pharmaceuticals Inc., USA.

2.1.3 General Chemicals and Reagents

All chemicals and reagents were of analytical research grade and purchased from a variety of companies as follows:

Acrylamide, Bisacrylamide, SDS: Bio-Rad, USA.

Agarose, DTT: Promega Corporation, USA.

Ampicillin, Complete protease inhibitor cocktail, IPTG, PMSF: Boehringer Mannheim, Germany.

Ammonium persulphate, β-mercaptoethanol, CHAPS, Coomassie brilliant blue R-250, Glutathione (reduced), HEPES, L-methionine, NP-40, PIPES, TEMED, Tris base, Trypan blue: Sigma Chemical Company, USA.

Deoxyribonucleotide Triphosphates, Glutathione sepharose 4B, Protein A sepharose CL-4B, RNase Inhibitor (RNAguard): Pharmacia Biotech, Sweden.

BSA, Triton X-100: Calbiochem, USA.

Phenol: Wako, Japan.

EDTA, Glycine, Tween 20, Urea: BDH Chemicals, Australia.

2.1.4 Antibodies

Anti-FLAG M2 monoclonal antibody was purchased from International Biotechnologies Inc., USA.

The anti-Myb 1.1 monoclonal antibody is raised against the carboxyl terminus of murine c-Myb (Ramsay *et al.*, 1989).

Polyclonal Myb antiserum (rabbit), raised against murine c-Myb residues 1-325, was a kind gift from Dr Linda Wolff, National Cancer Institute, Maryland, USA.

Peroxidase Conjugated Anti-mouse IgG (from goat) and Peroxidase Conjugated Anti-rabbit IgG (from goat) were purchased from Pierce, USA.

Peroxidase-linked Anti-mouse Ig (from sheep) and Peroxidase-linked Anti-rabbit Ig (from donkey) were purchased from Amersham Life Science, England.

2.1.5 Buffers

PBS: 136mM NaCl, 2.6mM KCl, 8mM Na_2HPO_4, KH_2PO_4, pH 7.3.

SSC: 150mM NaCl, 15mM sodium citrate, pH 7.1.

TAE: 40mM Tris-HCl pH 8.0, 20mM sodium acetate, 1mM EDTA.

TBE: 50mM Tris-HCl pH 8.3, 50mM boric acid, 1mM EDTA.

TBS: 50mM Tris-HCl pH 7.4, 135mM NaCl.

TTBS: TBS with the addition of 0.1% Tween 20.

2.1.6 Synthetic Oligonucleotides

Oligonucleotides were synthesised using an Applied Biosystems DNA Synthesiser (model 381A) within this department or by Bresatec, Australia. Oligonucleotides synthesised by Bresatec were received either as crude or purified lyophilized preparations.

Oligonucleotides were removed from synthesis columns by gently pushing 400-500 µl of ammonium hydroxide through the column several times with a 1ml syringe, ensuring that the column was full of this solution. The column was left for 15 minutes and the solution extracted using a 1 ml syringe. This procedure was repeated three more times. The total extracted solution was then incubated for 16 hours (overnight) at 56°C, cooled to RT and precipitated with nine volumes of butanol and spun at 3700 rpm (Beckman benchtop GPR centrifuge) for 20 minutes. The supernatant was discarded and the pellet washed twice with 70% ethanol, vacuum dried and resuspended in 200 µl sterile water. Some oligonucleotides required gel-purification. Such oligonucleotides were electrophoresed at 30 mA in a denaturing (8 M urea) 8% polyacrylamide gel with TBE buffer. Single stranded DNA was visualised by UV shadowing, excised and eluted overnight at 37°C in 400 µl sterile water, followed by ethanol precipitation and resuspension in 50 µl sterile water. The oligonucleotide concentration was determined by UV spectrophotometry at a wavelength of 260 nm.

2.1.7   Molecular Biology Kits

Kits were purchased from the following companies:

TNT Coupled Reticulocyte Lysate System, Wizard Plus Minipreps DNA Purification System: Promega Corporation, USA.

Qiagen DNA Purification (Midi) Kit: Qiagen, USA.

ABI Prism Dye Terminator Cycle Sequencing Ready Reaction Kit: Perkin Elmer, USA.

5'-AmpliFINDER RACE Kit: Clontech, USA.

Megaprime DNA Labelling Kit: Amersham Life Science, England.

BRESAclean DNA Purification Kit, SuperBase Sequencing Reagent Kit: Bresatec Ltd., Australia.

Detergent-Compatible Protein Assay Kit: Bio-Rad, USA.

2.1.8   Molecular Weight Markers

DNA molecular weight markers were purchased from the following companies:

*Eco*RI-digested SPP-1 bacteriophage DNA and *Hpa*II-digested pUC19: Bresatec Ltd., Australia.

Protein molecular weight markers were purchased from the following companies:

Prestained SDS-PAGE standards, low range and broad range: Bio-Rad, USA.

Prestained protein molecular weight standards (14,300-200,000 Da): Gibco-BRL, USA.

SeeBlue Pre-stained standards: Novex, Australia.

2.1.9 Bacterial Strains and Culture

The following *E.coli* strains were used to maintain cloned DNA:

BL-21(DE3): *hsdS, gal* (λcIts857 *ind1 Sam7 nin5 lac*UV-T7 gene 1)

DH1: *supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1*

DH5α: *supE44 ΔlacU169* (φ80 *lacZΔM15*) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*

LE392: *supE44 supF58 hsdR514 galK2 galT22 metB1 trpR55 lacY1*

BMH71-18: *mutS RecA*⁺

Bacterial cells were cultured in LB medium or on LB-agar, supplemented with 100 μg/ml ampicillin where required, as described in Sambrook *et al.* (1989).

2.1.10 Cell Culture Reagents

Reagents and equipment used for cell culture were obtained from the following sources:

Dulbecco's modified Eagle's Medium (DMEM): Gibco BRL, USA; Institute of Medical and Veterinary Science (IMVS) Media Production Unit, Adelaide, Australia.

Methionine-free DMEM, Trypsin: Trace Biosystems, Australia.

Foetal calf serum (FCS) and G418 (Geneticin): Gibco BRL, USA.

Puromycin: Calbiochem, USA.

Glutamine, penicillin, gentamicin, PBS, 7.5% sodium bicarbonate and sterile water: IMVS Media Production Unit, Australia.

Mouse GM-CSF was a kind gift from Dr Tracy Wilson, Walter and Eliza Hall Institute, Australia.

Human GM-CSF was a kind gift from Bronwyn Cambareri, this department.

Sterile plasticware used for cell culture was purchased from Corning, USA; Falcon, USA; Nunc, Denmark; Sterilin, USA.

2.1.11 RNA, Vectors and Cloned DNA Sequences

FDC-P1 poly A⁺ RNA was a kind gift from Elizabeth Macmillan, this department.

A blot containing poly A$^+$ RNA from murine and human cell lines was generated by Diane Favier, this department.

pBluscript SK was purchased from Stratagene, USA.

pGex-3X and pGex-4T-1 were purchased from Pharmacia Biotech, Sweden.

pGEM-3Z was purchased from Promega Corporation, USA.

pact-c-*myb* (described by Nishina *et al.*, 1989) was a kind gift from Dr Shunsuke Ishii, RIKEN, Japan.

Partial 3.1 kb p160 cDNA (encoding p120) was cloned into pBluescript SK (pBlusescript-120) by Diane Favier, this department.

pALTER-1 plasmids containing p160 cDNA or p67* cDNA were generated by Dr Rebecca Keough, this department.

pPR28 containing the SV40 large T-antigen CcN motif (Jans *et al.*, 1991) with mutated p34^{cdc2} phosphorylation sites, was a kind gift from Dr David Jans, Australian National University, Australia.

2.1.12 Construction of Expression Vectors

(I) Full-length p160 cDNA construct

A full-length p160 cDNA clone was constructed in pGEM-3Z by utilising the unique *Kpn*I site present in both the 5' RACE PCR product (see section 2.2.1.9) and the 3.1 kb partial p160 cDNA clone (nucleotides 992-4126). Partial p160 cDNA in pBluescript was digested with *Eco*RI/*Xho*I and subcloned into the *Eco*RI/*Sal*I sites of pGEM-3Z. This construct was subsequently digested with *Eco*RI/*Kpn*I and ligated with the 5' RACE PCR *Eco*RI-*Kpn*I fragment to yield a full-length p160 cDNA clone (pGEM3Z-160). Competent *E.coli* DH1 cells were transformed with pGEM3Z-160.

(II) p67* cDNA construct

The p67* cDNA clone was derived from p160 cDNA by introducing a termination codon after nucleotide position 1755 by PCR. The following PCR primers and conditions were used to generate the p67* cDNA clone from pGEM3Z-160:

5'-TATCGAATTCGTGACGTGTTTGGCTCAGC-3' and

5'-ACGTGGATCCTCATTCCTTCAGAGTACTC-3', incorporating a *Bam*HI site (94°C for 45 seconds; 60°C for 1 minute; 72°C for 90 seconds; for 30 cycles). The resulting PCR

product was digested with *Eco*RI/*Bam*HI and cloned into pGEM-3Z to yield pGEM3Z-67*. Competent *E.coli* DH1 cells were transformed with pGEM3Z-67*.

(III)   FLAG epitope-tagged constructs

The FLAG epitope was introduced into the p160 and p67* cDNA clones by site-directed mutagenesis (section 2.2.1.11). The FLAG epitope-tagged p160 and p67* expression vectors were constructed using pact-c-*myb* (Figure 2.1). The pact-c-*myb* expression vector contains the murine c-*myb* gene downstream of the 5' regulatory region of the chicken cytoplasmic β-actin gene (Nishina *et al.*, 1989). The c-*myb* gene was excised from pact-c-*myb* by digestion with *Nco*I/*Xba*I and replaced by an *Nco*I-*Xba*I fragment containing FLAG p67* cDNA   (from the p67* mutagenesis product: pALTER-67*FLAG) to yield pact-67*FLAG. The FLAG p160 cDNA expression vector, pact-160FLAG, was constructed by generating a fragment containing FLAG p160 cDNA by PCR from the p160 mutagenesis product (pALTER-160FLAG) using the SP6 promoter sequence and

5'-ACGTACTAGTTCAAGGTGTCTGCACTCTC-3',   incorporating   a   *Spe*I   site (conditions: 94°C for 45 seconds; 50°C for 1 minute; 72°C for 4 minutes; for 30 cycles). The resulting PCR product was digested with *Nco*I/*Spe*I, and subsequently used to replace the c-*myb* gene in the *Nco*I/*Xba*I sites of pact-c-*myb*. Competent *E.coli* DH5α cells were transformed with pact-160FLAG anf pact-67*FLAG.

The carboxyl-truncated p160 cDNA clones, pact-73FLAG and pact-80FLAG, were constructed by introducing a termination codon after nucleotide positions 1932 and 2118, respectively, by PCR from pALTER-160FLAG, using the SP6 promoter sequence and the following primers:

p73FLAG: 5'-ACGTACTAGTTCAGGCTTTGGCTCTAGAGCG-3', incorporating a *Spe*I site (conditions: 94°C for 45 seconds; 50°C for 1 minute; 72°C for 2 minutes; for 30 cycles).

p80FLAG: 5'-ACGTACTAGTTCAGTCATCCTCATCCTCGTTTG-3', incorporating a *Spe*I site (conditions: 94°C for 45 seconds; 50°C for 1 minute; 72°C for 3 minutes; for 30 cycles).

43

Figure 2.1   Schematic of the pact-c-*myb* expression vector.

The pact vector was originally constructed from pRSVCAT. The pact-c-*myb* expression vector is described by Nishina *et al.* (1989). Murine c-*myb* is cloned into the *Nco*I and *Xba*I unique restriction sites. The *Nco*I site also constitutes the c-*myb* translation start site.

AmpR (2866bp)

Xho I

pact-c-*myb*

(6306bp)

β-actin
5' promoter
(1310bp)

Bam HI

Xba I

SV40
Poly A
signal
(140bp)

Nco I

CCATGG

Met
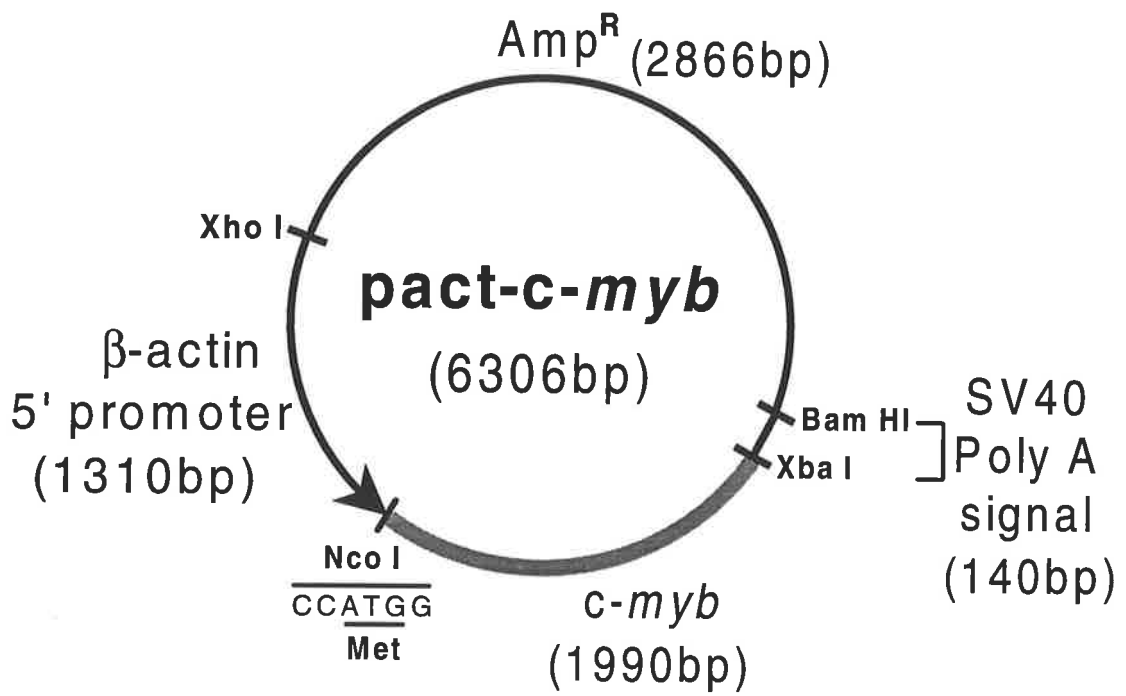
c-*myb*
(1990bp)

The resulting PCR products were digested with NcoI/SpeI and used to replace the c-myb gene in the NcoI/XbaI sites of pact-c-myb. Competent E.coli DH1 cells were transformed with pact-73FLAG and pact-80FLAG. The presence of an in-frame FLAG epitope and termination codon was verified by sequencing the 5' and 3' ends of each clone.

(IV)   p67* nuclear localisation signal (NLS) constructs

Two pact-67*FLAG, NLS containing constructs were made as follows:

pact-67*FLAG NLS (m) contains the minimal SV40 large T-antigen NLS (^{126}PKKKRKV^{132}) and was constructed by generating a p67*FLAG cDNA fragment (from pALTER-67*FLAG) with the NLS incorporated at the 3' end, by PCR using the T7 promoter sequence and

5'-ACGTACTAGTTCACACCTTGCGCTTCTTCTTAGGGGCTCCTTCCTTCAGAGT---ACTCATCATC-3' (purified), incorporating a termination codon and a SpeI site, in addition to the NLS (conditions: 94°C for 45 seconds; 59°C for 1 minute; 72°C for 2.5 minutes; for 5 cycles, then 94°C for 45 seconds; 65°C for 1.5 minutes; 72°C for 2.5 minutes; for 30 cycles). The resulting PCR product was digested with NcoI/SpeI and used to replace the c-myb gene in the NcoI/XbaI sites of pact-c-myb (Figure 2.1). Codon preference for each residue of the NLS is based on codon usage as described by Aota et al. (1988). Competent E.coli DH1cells were transformed with pact-67*FLAG NLS (m). The presence of an in-frame FLAG epitope, NLS and termination codon was verified by sequencing the 5' and 3' ends of this clone.

pact-67*FLAG NLS (e) contains the SV40 large T-antigen CcN motif (residues 110-132) (Jans et al., 1991), with mutated p34^{cdc2} phosphorylation sites. A fragment containing the CcN motif was generated by PCR from pPR28 using the following primers and conditions: 5'-AGCTCGGTACGTACCCGG-3' and 5'-ACGTACTAGTTCATACCTTTCTCTTCTTTTTTGG-3', incorporating a termination codon and SpeI site (94°C for 45 seconds; 54°C for 45 seconds; 72°C for 30 seconds; for 30 cycles). This fragment, containing the CcN motif, was digested with BamHI/SpeI. A p67*FLAG cDNA fragment containing the termination codon replaced with a BamHI site was generated by PCR from pALTER-67*FLAG using the T7 promoter sequence and 5'-CAGTGGATCCTTCCTTCAGAGTACTCATC-3' (conditions: 94°C for 45 seconds; 59°C for 1 minute; 72°C for 2.5 minutes; for 30 cycles). This fragment was digested with

*Nco*I/*Bam*HI. A pact-67*FLAG NLS(e) clone was generated by tri-ligation of the CcN motif *Bam*HI-*Spe*I and p67*FLAG *Nco*I-*Bam*HI fragments with a pact *Nco*I-*Xba*I fragment (Figure 2.1). Competent *E.coli* DH1 cells were transformed with pact-67*FLAG NLS (e). The presence of an in-frame FLAG epitope, CcN motif and termination codon was verified by sequencing the 5' and 3' ends of this clone.

(V)   GST fusion constructs

Glutathione S-transferase (GST) fusion constructs containing 5' or 3' p160 cDNA sequences were made as follows:

GST-160N, containing p160 amino-terminal residues 1-326, was generated by digestion of pGEM3Z-160 with *Eco*RI/*Kpn*I (nucleotides 1-1083), blunt-ended and ligated into the *Sma*I site of pGex 4T-1.

GST-160C, containing p160 carboxyl-terminal residues 1046-1344, was generated by digestion of partial p160 cDNA in pBluescript with *Bam*HI/*Ssp*I (nucleotides 3151-4101), blunt-ended and ligated into the *Sma*I site of pGex-3X.

Competent *E.coli* DH5α, BL-21 and LE392 cells were transformed with GST-160N. Competent *E.coli* DH5α and BL-21 cells were transformed with GST-160C. In-frame GST-p160 sequences were verified by DNA sequencing.

2.2   METHODS

2.2.1   DNA methods

2.2.1.1   Purification of plasmid DNA

Small scale purification (miniprep) of plasmid DNA was carried out using the Wizard Plus Miniprep Purification System in accordance with the manufacturer's instructions. Larger scale purification (midiprep) of plasmid DNA was carried out using Qiagen columns in accordance with the manufacturer's instructions. DNA concentration was determined by UV spectrophotometry at a wavelength of 260 nm.

2.2.1.2   Restriction enzyme digestion

Restriction enzyme digests were essentially conducted in accordance with the manufacturer's instructions, except that all digests were carried out in the presence of Super Duper Buffer (33 mM Tris-acetate pH 7.8, 62.5 mM potassium acetate, 10 mM

magnesium acetate, 4 mM spermidine, 0.5 mM dithioerythritol), with approximately 10 units of enzyme per µg of DNA. All digests were incubated for at least 1 hour.

2.2.1.3 Agarose gel electrophoresis

DNA was fractionated by electrophoresis in 1% agarose gels at 60-100 mA. Gels were submerged horizontally in TAE buffer, and DNA samples loaded in 2.5% Ficoll, 0.1% lauryl sarkosyl, 0.025% bromophenol blue, 0.025% xylene cyanol. DNA was visualised by staining with ethidium bromide (0.5 µg/ml) and exposure to UV light. Gels were documented by photography using Polaroid 667 film.

2.2.1.4 Isolation of DNA fragments from agarose gels

Restriction enzyme digested DNA fragments were recovered from agarose gels firstly by marking the band of interest with a cut, then excising the band of interest in the absence of UV light. Isolation of DNA from the gel slice was carried out using the BRESAclean Kit, in accordance with the manufacturer's instructions.

2.2.1.5 Ethanol Precipitation

Ethanol precipitation of DNA was performed as described in Sambrook *et al.* (1989). Approximately 20 µg of glycogen was added to facilitate precipitation of small amounts of DNA.

2.2.1.6 Phenol extraction

Extractions were performed firstly by mixing the aqueous solution with an equal volume of buffer equlibrated phenol (10 mM Tris-HCl, pH 8.0, 0.1 mM EDTA)/chloroform (1:1) and microcentrifuged at 13000 rpm. The aqueous phase was removed to a new tube and mixed with an equal volume of chloroform (49% chloroform, 1% isoamylalcohol). Following microcentrifugation at 13000 rpm, the aqueous phase was removed and retained.

2.2.1.7 DNA subcloning

(I) Preparation of vector

Vectors were linearised with the appropriate restriction enzyme(s). If a blunt-ended cut was generated, the 5' terminal phosphate groups were removed with calf intestinal phosphatase (CIP) as described in Sambrook *et al.* (1989).

(II) Endfilling of DNA fragments

Klenow DNA polymerase was used to endfill protruding 5' or 3' ends as described in Sambrook *et al.* (1989).

(III) DNA ligations

Approximately 50 ng of linearised vector DNA was ligated to a three molar excess of insert DNA in the presence of T4 DNA ligase (1 unit) and ligation buffer (50 mM Tris-HCl pH 7.8, 10 mM $MgCl_2$, 20 mM DTT, 1 mM ATP pH 7.0, 50 µg/ml BSA) in a total volume of 10 µl. Ligation reactions were incubated at 15°C overnight.

(IV) Bacterial cell transformation

E. coli BL21, DH1, DH5α and LE392 cells were made competent using the calcium chloride method described in Ausubel *et al.* (1994). Approximately 10-50 ng of DNA or an entire ligation reaction was added to a 500 µl aliquot of competent cells, incubated on ice for 10 minutes, heat shocked at 37°C for 1 minute and replaced on ice for 1.5 minutes. Cells were transferred to 2 ml of LB medium and incubated at 37°C for 20 minutes. Following centrifugation at 2800 rpm (Beckman benchtop GPR centrifuge), cell suspensions were plated on to LB agar containing antibiotics and incubated overnight at 37°C.

Electrocompetent *E. coli* BMH71-18 cells were prepared by the method described in Ausubel *et al.* (1994). Cells were electroporated in a 0.1 cm Gene Pulser cuvette (Bio-Rad), with an applied electric field strength of 18.0 kV/cm (Bio-Rad Gene Pulser unit). Recovery medium (SOC: 2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 10 mM $MgSO_4$, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM glucose) was added to cells immediately after electroporation. A 1ml aliquot of electroporated cells was inoculated to 4 ml of LB medium, containing ampicillin, and incubated overnight at 37°C.

2.2.1.8 DNA Sequencing

Sequencing was performed using the SuperBase Sequencing Reagent Kit, T7 DNA polymerase and α-[^{33}P], in accordance with the manufacturer's instructions. Sequencing reactions were electrophoresed at 55mA in denaturing (8 M urea) 6% polyacrylamide gels with TBE buffer. Polyacrylamide gels were pre-electrophoresed for 45 minutes at 55 mA

prior to loading. After electrophoresis, gels were dried at 80°C under vacuum onto 3MM chromotography paper (Whatman) and autoradiographed at RT overnight with BIOMAX MR-1 film (Kodak) and an intensifying screen (Dupont). Sequences were read manually.

Automated sequencing was performed using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction kit, in accordance with the manufacturer's instructions, and an Applied Biosystems Sequencer (model 377).

2.2.1.9 Polymerase Chain Reactions

Polymerase Chain Reactions (PCRs) were conducted in a total volume of 25 μl with *Pfu* DNA polymerase (2.5 units), *Pfu* reaction buffer (Stratagene), 0.2 mM deoxyribonucleotide triphosphates (dNTPs), 100 ng of each oligonucleotide primer and 10 ng of template DNA. The total reaction volume was overlayed with 20 μl mineral oil (Sigma Chemical Company, USA) and heated to 94°C for 5 minutes immediately before thermal cycling. Thermal cycling was performed using a programmable temperature controller (MJ Research Minicycler). Prior to sub-cloning, PCR products were phenol/chloroform extracted, digested with restriction enzymes and isolated from agarose gels.

The p67* cDNA PCR product (see section 2.1.12) was generated using *Taq* polymerase with *Taq* reaction buffer, 0.2 mM dNTPs, 1.5 mM $MgCl_2$, 50 pmol of each primer and 100 ng of template DNA.

(I) 5' RACE PCR

5' RACE PCR (Frohman *et al.*, 1988) was performed using the 5'-AmpliFINDER RACE Kit protocol. First strand cDNA was synthesised from FDC-P1 poly A^+ RNA (2 μg) using a p160 cDNA-derived oligonucleotide: 5'-TTGTAAAGGCCACCATCATAGTCAACTGCC-3' (nucleotides 1115-1144) and SuperScript reverse transcriptase, as described (Rayner and Gonda, 1994). 5' RACE PCR was undertaken with a nested p160 gene-specific primer, incorporating a *Bam*HI site: 5'-CGGATCCGGAAGGTCCCACGTATG-3' (nucleotides 1066-1087), and the AmpliFINDER anchor primer, using the following conditions: 94°C for 45 seconds; 60°C for 1 minute; 72°C for 2 minutes; for 30 cycles. The resulting PCR product was digested with *Eco*RI/*Bam*HI, cloned into pBluescript SK and verified by DNA sequencing.

2.2.1.10 Preparation of radiolabelled DNA

DNA fragments were oligolabelled by a modified random labelling method (Feinberg and Vogelstein, 1983) using the Megaprime DNA Labelling Kit and α-[^{32}P]-dATP, in accordance with the manufacturer's instructions. Unincorporated nucleotides were removed by centrifugation through Bio-Spin 6 Chromatography columns (Bio-Rad, USA), according to the manufacturer's instructions. Probe counts were determined using a radioisotope counter (Bioscan Quick Count 2000, USA).

2.2.1.11 Site-directed mutagenesis

In vitro site-directed mutagenesis was performed using the pALTER-1 mutagenesis system in accordance with the manufacturer's instructions (Promega Corporation, USA). The *E.coli* BMH71-18 (*mut*S) strain was used to propagate mutagenesis products. The FLAG epitope (DYKDDDDK) was introduced into the amino termini of p160 and p67* (see section 2.1.12) using double-stranded site-directed mutagenesis. Separate mutagenic oligonucleotides were required for each p160 and p67*cDNA clone (pALTER-160 contains a *Hind*III-*Eco*RI p160 cDNA fragment cloned into the *Sma*I site, SP6 orientation; pALTER-67* contains a *Eco*RI-*Bam*HI p67* cDNA fragment cloned into *Eco*RI/*Bam*HI, T7 orientation; section 2.1.11) and gel purified before use. Each oligonucleotide incorporated the FLAG octapeptide sequence downstream of a Kozak consensus translation initiation site (Kozak, 1986), which replaced the original translation site (nucleotides 13-19). The following mutagenic oligonucleotides were used:

p160: 5'-AGGGCTCTTCATCTCCGCCTTGTCATCGTCGTCCTTGTAGTCCATGGT...

...GATGCTGAGCCAAACACGTC-3'

p67*: 5'-ACGTGTTTGGCTCAGCATCACCATGGACTACAAGGACGACGATGAC...

...AAGGCGGAGATGAAGAGCCCT-3'

Mutagenesis products were screened by digestion with *Nco*I. The *Nco*I site was generated as part of the introduced Kozak consensus sequence. Mutagenised sequences were verified by DNA sequencing. The pALTER-160FLAG and pALTER-67*FLAG constructs are maintained in *E.coli* DH5α with selection in ampicillin (The pALTER tetracycline

resistance gene was inactivated and the ampicillin resistance gene repaired during the mutagenesis procedure, as described by the manufacturer).

2.2.2 Northern analysis

Radiolabelled DNA (p160 cDNA nucleotides 1-1083) was used to probe an RNA blot (section 2.1.11) by the method described in Sambrook *et al.* (1989), with the following stringency conditions; Hybridisation: 50% formamide, 5X SSC, 42°C. Wash: 2X SSC, 0.1% SDS, 42°C. The filter was autoradiographed with Hyperfilm-MP (Amersham, UK) at -80°C for 48 hours with an intensifying screen (Dupont).

2.2.3 Protein methods

2.2.3.1 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were electrophoresed in 8-12% SDS polyacrylamide (29.2:0.8, acrylamide:bisacrylamide) gels containing 375 mM Tris-HCl pH 8.8, 0.1% SDS, 0.1% ammonium persulphate, 0.08% TEMED. Prior to loading, samples in SDS sample buffer (10% glycerol, 2% SDS, 60 mM Tris pH 6.8, 0.01% (w/v) bromophenol blue, 0.72 M β-mercaptoethanol) were heated at 100°C for 3 minutes. SDS polyacrylamide gels (17x14cm) were electrophoresed in a vertical apparatus with SDS running buffer (0.025 M Tris, 0.192 M glycine, 0.1% (w/v) SDS) at 20 mA through the stacking gel (containing 5% polyacrylamide, 125 mM Tris-HCl pH 6.8, 0.1% SDS, 0.1% ammonium persulphate 0.1% TEMED), then at 30 mA through the separating gel until the dye front had run off the gel. Alternatively, gels were run overnight at 6 mA. Mini gels (7x10cm) were electrophoresed in an SDS-PAGE apparatus (Pharmacia Biotech, Sweden), firstly at 15 mA through the stacking gel, then at 25 mA through the separating gel.

(I) Detection of proteins

Following electrophoresis, SDS polyacrylamide gels containing radiolabelled proteins were fixed (40% methanol, 10% glacial acetic acid) for 15 minutes at RT and dried onto 3MM chromotography paper (Whatman) at 80°C under vacuum. Dried gels were either autoradiographed for a period between 16 hours and 1 week at RT with an intensifying screen (Dupont), or exposed to a phosphorscreen (Molecular Dynamics, USA) overnight. Detection and quantitative analyses of radiolabelled proteins exposed to a phosphorscreen were carried out using a Molecular Dynamics Phosphorimager with ImageQuant software.

Non-radiolabelled proteins were detected either by Western analysis/ECL (section 2.2.3.10), or by staining with Coomassie brilliant blue R-250 as follows: After electrophoresis, gels were stained (0.25% Coomassie brilliant blue R-250, 40% methanol, 10% glacial acetic acid), rocking at RT. Gels were destained (40% methanol, 10% glacial acetic acid) for at least 4 hours, rocking at RT in the presence of small foam blocks to bind the dye. Gels were air-dried overnight between 2 sheets of clear cellophane (Bio-Rad, USA).

2.2.3.2 Preparation of cell extracts

The required cells were harvested and washed twice with PBS prior to preparation of extracts (see section 2.2.4).

(I) Nuclear and cytoplasmic extracts

Nuclear and cytoplasmic extracts were prepared essentially as described by Favier and Gonda (1994). Cell pellets were resuspended in cold Buffer 1 (10 mM Tris-HCl pH 8.0, 25 mM KCl, 5 mM $MgCl_2$, 0.5% NP-40, 5 mM DTT, 2 mM PMSF, 40 µl/ml Complete protease inhibitor cocktail) at $5x10^7$ cells/ml and microcentrifuged at 1000 rpm for 5 minutes at 4°C. The supernatant (cytoplasmic fraction) was retained and the nuclei resuspended in cold Buffer 1A (Buffer 1 without NP-40) at $5x10^7$ cells/ml. After microcentrifugation at 1000 rpm for 5 minutes at 4°C, the supernatant was discarded, the nuclei washed in cold Buffer 2 (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 5 mM $MgCl_2$, 500 mM NaCl, 20% glycerol, 5 mM DTT, 2mM PMSF, 40 µl/ml Complete protease inhibitor cocktail) at $1x10^8$ cells/ml and incubated on ice for 10 minutes. DNA and cell debris were pelleted at 45000 rpm (TLA-45 rotor, Beckman TL-100 Ultracentrifuge) for 20 minutes at 4°C. The resulting supernatant was retained as the nuclear fraction. All extracts were stored as aliquots at -80°C.

(II) Whole cell SDS lysates

Cell pellets were resuspended in SDS lysis buffer (2% SDS, 10% glycerol, 60 mM Tris-HCl pH 6.8., 5 mM DTT) at $5x10^8$ cells/ml and sonicated with a microtip (Microson; Misonix, USA) for 10 seconds on ice. Sonication was repeated twice more. Lysates were heated at 100°C for 3 minutes and centrifuged at 45000 rpm (TLA-45 rotor, Beckman TL-

100 Ultracentrifuge) for 20 minutes at 4°C. The supernatant was retained and stored as aliquots at -80°C.

(III)   Total cell extracts

Cell pellets were resuspended in cold Buffer 1 (10 mM Tris-HCl pH 8.0, 25 mM KCl, 5 mM $MgCl_2$, 0.5% NP-40, 5 mM DTT, 2 mM PMSF, 40 μl/ml Complete protease inhibitor cocktail) at 1×10^8 cells/ml with the addition of 500 mM NaCl. Extracts were incubated on ice for 10 minutes, then centrifuged at 45000 rpm (TLA-45 rotor, Beckman TL-100 Ultracentrifuge) for 20 minutes at 4°C. The supernatants were retained and stored as aliquots at -80°C.

(IV)   Total cell extracts for proteolysis assays

Cell pellets were resuspended in assay buffer (25 mM HEPES pH 7.4, 1 mM EDTA pH 8.0, 0.1% CHAPS, 10% sucrose, 5 mM DTT) at 1×10^8 cells/ml. Cells were lysed by three successive freeze/thaw cycles whereby cells were snap frozen in liquid nitrogen and thawed at 37°C. Extracts were microcentrifuged at 1000 rpm for 5 minutes at 4°C, the supernatant removed and microcentrifuged at 13000 rpm for 15 minutes at 4°C. The supernatant was retained and stored as aliquots at -20°C.

2.2.3.3   Protein quantitation

The protein concentration of cell extracts was determined using the Detergent-Compatible Protein Assay kit, in accordance with the manufacturer's instructions (Bio-Rad, USA). Assays were performed in microtiter plates and measured with a microplate reader (Bio-Rad, USA).

2.2.3.4   *In vitro* transcription/translation

Proteins were synthesised *in vitro* using the TNT (transcription and translation) coupled reticulocyte lysate system (Promega Corporation, USA). Briefly, 1 μg of DNA expression vector (pGEM3Z-160, pGEM3Z-67*, pBluescript-120) was incubated at 30°C for 90 minutes in a 50 μl reaction volume containing the recommended quantities of kit reagents, including T7 RNA polymerase (for pGEM3Z-160, pGEM3Z-67*) or T3 RNA polymerase (for pBluescript-120), 4 μl of translation grade [^{35}S]-methionine (10.2 mCi/ml; NEN Dupont) and 40 units of RNase Inhibitor (Pharmacia Biotech, Sweden).

2.2.3.5 Purification of GST proteins (maxi prep)

A 100 ml overnight culture (incubated at 30°C), inoculated directly from the required fusion protein *E. coli* glycerol stock, was inoculated to 800 ml LB medium containing ampicillin and incubated at 30°C for 1-1.5 hours, until OD_{600} = 0.5. Synthesis of fusion proteins was induced by the addition of 0.1 mM IPTG and further incubation for 2.5-3 hours. Cells were harvested by centrifugation at 5000 rpm (JA-10 rotor, Beckman J2-21M/E centifuge) for 5 minutes at 4°C. Cell pellets were resuspended in 10 ml Tris/NaCl buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 5 mM EDTA, 5 mM DTT, 2 mM PMSF, 40 μl/ml complete protease inhibitor cocktail) and incubated on ice for 10 minutes. Cells were lysed by sonication for 15 seconds on ice (repeated twice more). A 0.5-1.0 ml aliquot of glutathione sepharose 4B (50% slurry in PBS) was added to the lysate and incubated with rotation for 30 minutes at RT. Glutathione sepharose-bound fusion protein (affinity matrix) was isolated by centrifugation at 2000 rpm (Beckman benchtop GPR centrifuge) for 3 minutes at RT, washed four times successively with 10 ml PBS/0.5% Triton X-100. The affinity matrix was resuspended 1:1 in PBS containing 10% glycerol, 2 mM DTT and stored at -20°C. GST fusion proteins were approximately quantitated by electrophoresis with BSA standards (1 mg/ml; 0.5 mg/ml; 0.25 mg/ml).

2.2.3.6 Binding assays

All binding assays were performed essentially as described by Favier and Gonda (1994).

(I) *In vitro* translated proteins

Rabbit reticulocyte lysates containing *in vitro* translated proteins were diluted 1/4 with nuclear extraction Buffer 2 (see section 2.2.3.2) and precleared firstly with 100 μl glutathione sepharose 4B for 1 hour with rotation at 4°C. Following microcentrifugation at 13000 rpm for 30 seconds at 4°C, the supernatant was retained and precleared with GST bound sepharose (~500 μg). A 40 μl aliquot of precleared lysate was added to the required GST fusion protein affinity matrix (~50-100 μg; GST-NRD2, GST-L3,4P Myb, GST L3,4A Myb, GST-Jun) and incubated with rotation for 1 hour at 4°C. Fusion protein amounts were approximately equalised. The bound sepharose was isolated by microcentrifugation (as above) and the supernatant retained as the unbound fraction. The bound sepharose was washed five times successively with NETN (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5% NP-40). Washed affinity matrix was added to SDS

sample buffer and electrophoresed in an 8% SDS-polyacrylamide gel with the unbound fraction. Proteins were detected by autoradiography or Phosphorimager.

(II)   Ectopically expressed proteins

Total cell extracts from HEK 293T cells transfected with p160FLAG and p67*FLAG expression vectors were precleared, a 70 µl aliquot of precleared extract added to 50µl of the required GST fusion protein affinity matrix (~50-100 µg), incubated and washed, as described above (section 2.2.3.6 (I)). The bound fraction and 50% of the unbound fraction were electrophoresed in an 8% SDS-polyacrylamide gel. Proteins were detected by Western analysis/ECL.

(III)   Endogenous proteins

Nuclear and cytoplasmic extracts from 1×10^8 FDC-P1 cells were precleared as described above (section 2.2.3.6 (I)) using 250 µl of glutathione sepharose, GST-sepharose (~1 mg) and GST-L3,4P Myb-sepharose (~1 mg). A 400-500 µl aliquot of precleared radiolabelled nuclear extract was added to GST-NRD2-sepharose (~500 µg) and incubated for 1 hour with rotation at 4°C. Alternatively, a 100 µl aliquot of precleared nuclear extract or 200 µl of precleared cytoplasmic extract was added to the required GST fusion protein (~100µg) and incubated for 1 hour with rotation at 4°C. Bound affinity matrix was washed as described above (section 2.2.3.6 (I)) and electrophoresed with unbound fractions in 8% SDS-polyacrylamide gels. Radiolabelled proteins were detected by autoradiography or Phosphorimager. Unlabelled proteins were detected by Western analysis/ECL.

2.2.3.7   Peptide mapping

Peptide mapping of endogenous and *in vitro* synthesised proteins was performed by partial digestion with *Staphylococcus aureus* V8 protease, as described by Cleveland (1983). Endogenous, radiolabelled p160 and p67 proteins were affinity purified with GST-NRD2 (see section 2.2.3.6 (III)). Protein bands were excised from an unfixed, dried 8% SDS-polyacrylamide gel and placed in the wells of a 5 cm stacking gel (5% polyacrylamide). Proteins were electrophoresed at 20 mA, approximately 1cm into the stacking gel and partially digested *in situ* with varying amounts of V8 protease (50-1000 ng) for 20 minutes. Peptides were electrophoresed at 20 mA in a 15% SDS-polyacrylamide gel and detected by autoradiography using BIOMAX MR-1 film (Kodak).

2.2.3.8 *In vitro* proteolysis assay

Proteolysis of *in vitro* translated p160 was performed using total cell extracts. Total cell extracts (10 μg) were added to 10 μl radiolabelled p160, and the total volume made up to 30 μl with assay buffer (section 2.2.3.2 (IV). Reactions were incubated at 37°C for varying times (1-90 minutes). Reactions were stopped by the addition of 2X SDS sample buffer (1:1), electrophoresed in 8% SDS-polyacrylamide gels and proteins detected by autoradiography or Phosphorimager.

(I) Immunoprecipitation of proteolysis products

Proteolysis reactions containing 20 μl of radiolabelled p160 and 50 μg of FDC-P1 total cell extract, in a total reaction volume of 30μl (as described above), were incubated for 1 and 10 minutes at 37°C. Following incubation, a 10 μl aliquot was removed and added to SDS sample buffer. The remaining 20 μl was immunoprecipitated in a total volume of 100 μl, containing 5 μl Complete protease inhibitor cocktail and 10 μl of p160 polyclonal antiserum (160N or 160C), assay buffer, and incubated for 2 hours with rotation at 4°C. A 30 μl aliquot of Protein A sepharose CL-4B was added to the immunoprecipitates and incubated for 2 hours with rotation at 4°C. After microcentrifugation at 13000 rpm for 30 seconds at 4°C, the sepharose beads were washed successively with 500 μl of the following washing solutions:

IP1: 10 mM Tris-HCl pH 8.0, 500 mM NaCl, 1% NP-40
IP2: 10 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% NP-40, 0.1% SDS
IP3: 10 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1% NP-40

In vitro translated p160 and p67* (10 μl) were immunoprecipitated in an identical manner, but in the absence of FDC-P1 total cell extract. Immunoprecipitates were electrophoresed in an 8% SDS-polyacrylamide gel and autoradiographed.

(II) Inhibitor assay

Inhibition of *in vitro* translated p160 proteolysis was performed using 10 μg FDC-P1 total cell extract, 5 μl radiolabelled p160, a specific protease inhibitor (see Table 2.1) and assay buffer in a total reaction volume of 30 μl (see above). Prior to adding radiolabelled p160, reactions containing the total cell extract and inhibitor were pre-incubated at 37°C for 20 minutes. Following the addition of radiolabelled p160, reactions were incubated for 1 hour

Table 2.1   Protease inhibitors and corresponding concentrations used in inhibitor assay.

Inhibitor	Working Concentration	Source
Antipaindihydrochloride	50 μg/ml	Boehringer Mannheim[a]
Aprotinin	2 μg/ml	"
Bestatin	40 μg/ml	"
Chymostatin	60 μg/ml	"
E-64	10 μg/ml	"
EDTA-Na$_2$	0.5 mg/ml	"
Leupeptin	0.5 μg/ml	"
Pefabloc	1 mg/ml	"
Pepstatin	0.5 mg/ml	"
Phosphoramidon	0.33 mg/ml	"
PMSF	1 mM	"
Ac-DEVD-CHO[b]	1 μM	Dr Sharad Kumar, Hanson Centre for Cancer Research
Ac-YVAD-CMK[b]	50 μM	"
Iodoacetamide	50 μM	Sigma Chemical Company
TLCK[c]	50 μg/ml	"
TPCK[c]	100 μg/ml	"
MG 132 (Aldehyde-active)[d]	1 μM	Boehringer Ingelheim
MG 132 (Alcohol-inactive)[d]	1 μM	"

[a] Substrate specificity of each protease inhibitor is described by the manufacturer.

[b] Inhibitors of ICE-like proteases, described in Kumar and Harvey (1995).

[c] TLCK: Nα-p-tosyl-L-lysine chloromethyl ketone; TPCK: N-tosyl-L-phenylalanine chloromethyl ketone. Substrate specificity is described by the manufacturer.

[d] 26S proteasome inhibitor, a kind gift from Dr M. Frances Shannon, this department.

at 37°C. Reactions were stopped by the addition of 30 µl 2X SDS sample buffer and analysed as described above.

2.2.3.9 Coimmunoprecipitation

Extracts from HEK 293T cells transfected with expression vectors (pact-160FLAG, pact-67*FLAG, pact-c-*myb*) were prepared and coimmunoprecipitations performed essentially as described by Harlow *et al.* (1986). Cell pellets were resuspended in 400 µl cold Harlow buffer (50 mM HEPES pH 7.5, 250 mM NaCl, 0.2 mM EDTA, 10 µM NaF, 0.5% NP-40) and incubated for 1 hour with rotation at 4°C. Following microcentrifugation at 13000 rpm for 20 minutes at 4°C, the supernatant was precleared with 50 µl Protein A sepaharose CL-4B for 1 hour with rotation at 4°C. Monoclonal antibodies were added to the precleared lysate (5 µl anti-FLAG M2 or an equal amount (3 µl) of an irrelevant antibody (1C1, raised against the GM-CSF receptor common β-chain)) and incubated for 2.5 hours with rotation at 4°C. A 50 µl aliquot of Protein A sepharose was added and incubated for 1 hour with rotation at 4°C. The sepharose beads were washed three times successively with 500 µl cold Harlow buffer containing 143 mM NaCl. Immunoprecipitates and a proportion (50 µl) of the unbound supernatant were electrophoresed in an 8% SDS-polyacrylamide gel. Proteins were analysed by Western analysis with the anti-FLAG M2 monoclonal antibody or the anti-Myb 1.1 monoclonal antibody.

2.2.3.10 Western analysis

After SDS-PAGE, proteins were transferred to a nitrocellulose membrane (Protran 0.2 µm; Schleicher and Schuell, Germany) using a semi-dry electro-transfer apparatus (Pharmacia Biotech, Sweden) with transfer buffer (48 mM Tris-HCl pH 7.4, 39 mM glycine, 0.0375% SDS, 20% methanol), in accordance with the manufacturer's instructions. Membranes were either blocked overnight at 4°C in TBS containing 5% (w/v) skim milk, or for 1-2 hours at RT in TTBS containing 5% (w/v) skim milk (TTBS+milk). Membranes were further blocked for 1 hour at RT in TTBS+milk, if previously blocked overnight. Membranes were rinsed with TTBS prior to adding the primary antibody. The primary antibody was incubated with the membrane in a minimal volume of TTBS (enough to cover the membrane) at RT for 45 minutes. Membranes were firstly rinsed with TTBS+milk, then washed twice for 15 minutes at RT in TTBS+milk, rocking. After rinsing the membrane with TTBS, the secondary antibody (peroxidase-conjugated anti-mouse or anti rabbit IgG

(Pierce) (1/10,000 or 1/20,000); peroxidase linked anti-mouse or anti-rabbit Ig (Amersham) (1/1000) - see section 2.1.4) was incubated with the membrane in TTBS for 1 hour at RT. Membranes were washed as described above. Proteins were detected by enhanced chemiluminescence (ECL) using SuperSignal Substrate (Pierce, USA) or ECL Western blotting detection reagents (Amersham, UK), in accordance with the manufacturer's instructions. Membranes were exposed to film (Hyperfilm-MP; Amersham, UK) for a period between 5 seconds and 15 minutes.

Primary and secondary antibodies were removed from membranes by incubation in stripping solution (100 mM β-mercaptoethanol, 2% SDS, 62.5 mM Tris-HCl, pH 6.7) at 56°C for 30 minutes. Membranes were rinsed with TTBS and washed twice in 200 ml TTBS for 15 minutes at RT. Membranes were blocked and reprobed, as described above.

2.2.3.11   Generation of p160 antibodies

Polyclonal antisera were raised against the amino- and carboxyl-terminal regions of p160, and are termed 160N and 160C, respectively. GST-p160 fusion proteins (see section 2.1.12 (V)), expressed from GST-160N (p160 amino-terminal region) and GST-160C (p160 carboxyl-terminal region) were purified and used to inoculate (in duplicate) adult New Zealand White rabbits. GST-p160 fusion proteins (250 μl bound sepharose) were eluted from the glutathione sepharose by incubation with 500 μl 5 mM glutathione (reduced)/50 mM Tris-HCl pH 8.0 for 15 minutes with rotation at RT. Elution was repeated once more. Approximately 80-100 μg of eluted fusion protein, mixed 1:1 with complete Freund's adjuvant (CSL, Australia), was inoculated subcutaneously into the rabbits, as described in Harlow and Lane (1988). The initial inoculation was followed by three more inoculations of 25-80 μg of eluted fusion protein with incomplete Freund's adjuvant (CSL, Australia), at 2-3 week intervals. Test bleeds were taken from the rabbits 12 days after the final inoculation and tested for reactivity. Rabbits were sacrificed 1 month after the final inoculation to obtain the antisera. Antisera were stored long-term as aliquots at -80°C. Working stocks were stored at 4°C.

A 1 ml aliquot of each p160 antiserum (160N and 160C) was depleted of anti-GST antibodies by absorption with GST-bound glutathione sepharose. Each antiserum was incubated with GST-sepharose (~1 mg) for 1 hour with rotation at 4°C. After

microcentrifugation at 13000 rpm for 30 seconds, GST-sepharose (~0.5 mg) was added to the supernatant and incubated for 3 hours at 4°C. Incubation with GST-sepharose was repeated once more, overnight at 4°C. The anti-GST-depleted p160 antiserum was retained after microcentrifugation and stored at 4°C. The anti-GST-depleted p160 antisera was used specifically in a binding assay with nuclear and cytoplasmic extracts prepared from FDC-P1 cells.

2.2.4 Cell culture and procedures

2.2.4.1 Cell lines and culture

The following cell lines were used in this study:

FDC-P1; murine myelomonocytic progenitor-like[a].

NIH3T3; murine fibroblast[a].

Human embryonic kidney (HEK) 293T; epithelial (DuBridge *et al.*, 1987).

WEHI-3B D[+]; murine myelomonocytic progenitor-like[a], expressing the human GM-CSF receptor α-chain and common β-chain (Smith *et al.*, 1997). This WEHI-3B cell line was a kind gift from Dr Nic Nicola, Walter and Eliza Hall Institute, Australia.

[a] Cell types described in Favier and Gonda (1994).

All cells were cultured in Dulbecco's modified Eagle's medium (DMEM) with 10 mM HEPES and supplemented with 10% foetal calf serum (heat inactivated), 2 mM glutamine, 12 µg/ml penicillin, 16 µg/ml gentamicin, 0.225% sodium bicarbonate. Additional supplements were added for the maintenance of the following cell lines:

FDC-P1: 50 units/ml mouse GM-CSF.

WEHI-3B D[+]: 2.5 µg/ml puromycin and 400 µg/ml G418.

All cell lines were cultured at 37°C, 5% CO_2, in a CO_2 incubator (Forma Scientific) and maintained at a density between 5×10^4 and 2×10^6 cells/ml.

Following removal of the culture medium and washing in PBS, adherent cells (NIH3T3 and HEK 293T) were removed from culture flasks by treatment with trypsin (0.5% trypsin, 0.1mM EDTA-PBS). Viable cells, determined by Trypan blue exclusion (1:1, 0.1% (w/v) Trypan blue), were counted using a haemocytometer.

2.2.4.2 Transient transfection

Transient transfection of NIH3T3 or HEK 293T cells with DNA expression constructs was performed firstly by seeding 1.2×10^6 293T cells or 5×10^5 NIH3T3 cells in 60 mm culture dishes and leaving them to adhere overnight, after which the medium was changed. Four hours after medium change (4.5 ml DMEM containing supplements), calcium phosphate precipitates containing 5-10 µg of expression constructs (see section 2.1.12) were added. Calcium phosphate precipitates were prepared by adding 250 µl of DNA/250 mM $CaCl_2$ solution dropwise to 250 µl of 2X HEPES/phosphate solution (250 mM NaCl, 1.5 mM sodium phosphate pH 7.0, 50 mM HEPES pH 7.1), mixed simultaneously by rapidly bubbling air through the solutions and left for 20 minutes to precipitate. Precipitates were left on HEK 293T cells for 4 hours, after which the cells were trypsinised, replated onto 100 mm culture dishes and incubated for 40-48 hours. Precipitates were left on NIH3T3 cells overnight, medium changed (5 ml) the following morning and cells incubated for a further 24 hours. After incubation, cells were harvested by trypsinisation, washed twice with PBS and extracts prepared. Cotransfection of HEK 293T cells with expression constructs was performed in an identical manner, except that 5 µg of each expression construct was used, with the final amount of transfected DNA equal to 10 µg.

2.2.4.3 Metabolic labelling

FDC-P1 cells were metabolically labelled with [^{35}S]-L-methionine (ICN, USA). Approximately 1×10^8 cells were pelleted at 1000 rpm (Beckman benchtop GPR centrifuge) and washed with PBS (repeated twice). The washed cell pellet was resuspended in 10 ml of methionine-free DMEM containing supplements, and incubated for 1 hour, after which 1.5 mCi of radiolabel was added, mixed, and left to incubate for a further 1.5 hours. After incubation with radiolabel, cells were harvested by centrifugation (as above), washed with PBS, and nuclear/cytoplasmic extracts prepared. The p160 and p67 proteins were affinity purified with GST-NRD2 (see section 2.2.3.6 (III)).

2.2.4.4 Pulse-chase treatment

FDC-P1 cells (1×10^8) were metabolically labelled with [^{35}S]-L-methionine (ICN, USA) for 1 hour (section 2.2.4.3), after which a 1000-fold molar excess of L-methionine was added, mixed and incubated for up to 4 hours. Approximately 2×10^7 cells (2 ml) were removed at the following time points: 0 (immediately after the addition of excess methionine) 0.5, 1, 2

59

and 4 hours. Cells were washed twice with PBS, and nuclear/cytoplasmic extracts prepared. The p160 and p67 proteins were affinity purified with GST-NRD2 (see section 2.2.3.6 (III)). Proteins were detected using a Molecular Dynamics Phosphorimager, and quantitative analyses performed using ImageQuant software (Molecular Dynamics, USA).

2.2.4.5   Induced differentiation of WEHI-3B D$^+$ cells

WEHI-3B D$^+$ cells expressing the human GM-CSF receptor α-chain and common β-chain were induced to terminally-differentiate into macrophages, essentially as described by (Smith *et al.*, 1997). Approximately 1×10^7 cells were added to a total volume of 100 ml DMEM containing supplements, including 2.5 µg/ml puromycin, 400 µg/ml G418 and 10 ng/ml human GM-CSF. Five separate culture flasks each containing 1×10^7 cells were incubated for up to 5 days. Uninduced cells were cultured in the absence of human GM-CSF, and used as a control. Whole cell SDS lysates were prepared at the following time points at 24 hour intervals: Day 0 (uninduced), Day 1, Day 2, Day 3, Day 4, Day 5. Prior to preparation of SDS lysates, approximately 2×10^4 induced or uninduced cells were removed for staining. The removed cells were centrifuged onto glass microscope plates at 500 rpm for 5 minutes (Shandon Cytospin-3 centrifuge) and stained with Giemsa. Differentiation was assessed on the basis of cell morphology.

CHAPTER THREE
Cloning and Characterisation of p160 and p67

3.1 Introduction

The c-Myb negative regulatory domain (NRD) is clearly a functionally complex domain. The mechanisms underlying the functions of the NRD remain mostly uncharacterised. The leucine zipper motif is one identified component of the NRD that modulates c-Myb transactivation and transformation ability (Kanei-Ishii *et al.*, 1992). The identification and characterisation of proteins that interact with the c-Myb leucine zipper should contribute towards a mechanistic understanding of how the NRD functions to regulate the activity of c-Myb.

The related p160 and p67 proteins were identified from murine cell lines on the basis of their ability to interact specifically with the c-Myb leucine zipper motif (Favier and Gonda, 1994). The binding properties of p160 and p67, and their distribution among particular cell lines has been discussed in Chapter 1.8.1.1. A 3.1 kb cDNA clone containing peptide sequences derived from p67 has been isolated (see section 1.8.1.1) (Gonda *et al.*, 1996; Tavner *et al.*, 1998; see appendix). This cDNA clone is predicted to encode a protein of 114 kDa and moreover, can direct translation (*in vitro*) of an approximately 120 kDa protein (see section 3.2). On the basis of the predicted and observed molecular masses of the protein encoded by the 3.1 kb cDNA clone, and the relatedness of p160 and p67, it was more likely that this clone corresponded to p160 mRNA. However, the difference in the molecular mass of p160 and the 120 kDa protein encoded by the 3.1 kb cDNA clone was unaccounted for. Furthermore, the relationship of this clone to p67 was unclear. This chapter describes the complete molecular cloning of p160/p67 cDNA and characterisation of the properties of these proteins.

3.2 Results

3.2.1 Isolation of a full-length p160 cDNA clone

A comparison of the relative sizes of endogenous p160 and p67, and the 120 kDa *in vitro* translated protein (termed p120) encoded by the 3.1 kb cDNA clone, clearly shows that the size of p120 does not coincide with that of either p67 or p160 (Figure 3.1). If the 3.1 kb

Figure 3.1 Size comparison of endogenous p160 and p67 with *in vitro* translated p120.

Endogenous p160 and p67 were affinity purified from nuclear extracts of metabolically labelled FDC-P1 cells using the GST-NRD2 fusion protein, which contains the c-Myb leucine zipper motif. The 120 kDa protein (p120) is the *in vitro* translation product of the 3.1 kb cDNA clone containing two of the five p67 tryptic peptide sequences. Proteins are identified by labels shown on the right. The positions of molecular weight markers (kDa) are shown on the left.
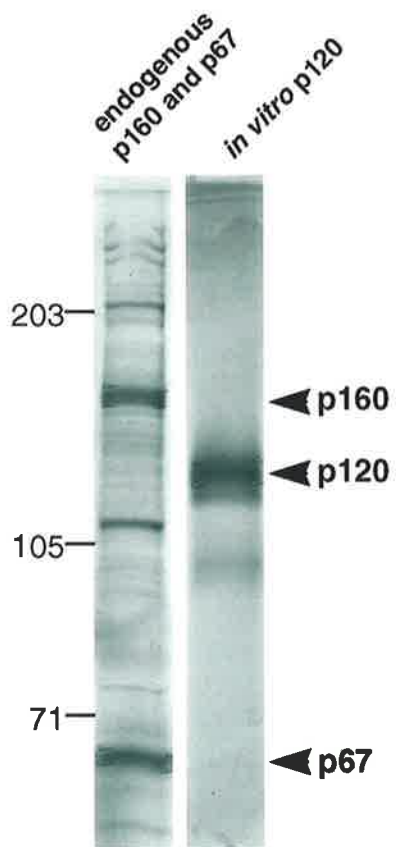
cDNA clone does indeed correspond to p160 mRNA, as proposed, there are at least two obvious explanations that may account for the observed differences in molecular weight. Firstly, the *in vitro* translated 120 kDa protein may represent an unmodified version of endogenous p160. That is, the *in vitro* translation system does not confer protein modifications such as glycosylation, which may account for the large (approximately 40 kDa) size difference. Alternatively, the 3.1 kb cDNA clone may represent a partial p160 cDNA clone, with the remaining sequences yet to be isolated. It was decided to pursue the possibility that the 3.1 kb cDNA clone represented a partial p160 cDNA clone.

Rapid amplification of cDNA ends (RACE) PCR (Frohman *et al.*, 1988) was undertaken to isolate putative remaining p160 cDNA sequences (Figure 3.2A). It was suspected that the 3' end of the 3.1 kb cDNA clone contained complete coding sequences due to the presence of multiple stop codons and a polyadenylation signal consensus sequence (nucleotides 4106-4111) (Figure 3.3). Therefore, a 5' RACE PCR approach was initiated to isolate putative 5' p160 cDNA sequences. Using gene-specific oligonucleotides derived from the 5' end of the 3.1 kb cDNA clone, a 5' RACE PCR-amplified product of approximately 1.1 kb was generated from FDC-P1 polyadenylated RNA (Figure 3.2B).

The ~1.1 kb 5' RACE PCR product was cloned and sequenced. Sequencing verified that this product contained 96 bp corresponding to the 5' end of the 3.1 kb cDNA clone, in addition to a further 991 bp of 5' coding sequence. The addition of the 5' RACE PCR-derived sequences to the existing 3.1 kb cDNA sequence resulted in an extended cDNA clone (4126 bp) comprising a long open reading frame encoding a polypeptide of 1344 amino acids with a predicted molecular mass of 152 kDa, and containing a putative translation initiation site (Kozak consensus) (nucleotides 13-19) (Figure 3.3). Subsequently, the 5' RACE PCR product and the 3.1 kb cDNA clone were physically joined together by utilising a unique *Kpn*I site present in both clones (Chapter 2.1.12 (I)). A comparison of the *in vitro* translation products of the 3.1 kb and 4.1 kb cDNA clones shows that a larger, approximately 160 kDa protein is encoded by the extended 4.1 kb clone (Figure 3.4A). Furthermore, comparison of the 160 kDa *in vitro* translated protein with endogenous p160, affinity purified from radiolabelled FDC-P1 nuclear extracts, shows that these proteins comigrate in an SDS-polyacrylamide gel (Figure 3.4B). The authenticity of the extended 4.1 kb cDNA clone was confirmed by peptide mapping with

Figure 3.2 Schematic representation of 5' RACE PCR, and the p160 5' RACE PCR product.

(A) A schematic outline of the 5' RACE PCR cloning technique. Steps 1-3 denote the following: 1. First strand cDNA synthesis from polyadenylated RNA with a gene-specific oligonucleotide (the p160-specific oligonucleotide sequence is shown; nucleotides 1115-1144); 2. Ligation of the 5' anchor oligonucleotide to single-stranded cDNA; 3. PCR-amplifcation of 5' cDNA sequence with the 5' anchor primer and a nested gene-specific primer (the p160-specific oligonucleotide sequence is shown; nucleotides 1066-1087). This figure is modified from the Clontech (USA) 5'-AmpliFINDER RACE Kit protocol.

(B) The p160 5' RACE PCR amplification product (approximately 1.1 kb), generated from FDC-P1 polyadenylated RNA using the oligounucleotides shown in part (A). 'M' denotes DNA markers (*Eco*RI-digested SSP-1 bacteriophage DNA).

A

1.
```
                              5' end of p120
                              cDNA clone
mRNA ───────────────────▢─────────────────── A_n
antisense ◄───────────────────▢── gene-specific anchor oligonucleotide:
                              5'-TTGTAAAGGCCACCATCATAGTCAACTGCC-3'
```

2.
```
        EcoRI
cDNA    ▢━━━━━━━━━━━━━━━━━▢
        5' Anchor
```

3.
```
cDNA          ▢━━━━━━━━━━━━━━━━━▢
Anchor primer ➤━━━━━━━━━━ sense

              ━━━━━━━━━━━━◄ nested gene-specific oligonucleotide:
              5'-CGGATCCGGAAGGTACCCACGTATG-3'
                   ‾‾‾‾‾‾
                   BamHI

            EcoRI
amplified 5' RACE ┴━━━━━━━━━━━━━━━━
PCR product       ━━━━━━━━━━━━━━━━┬
                               BamHI
```

B
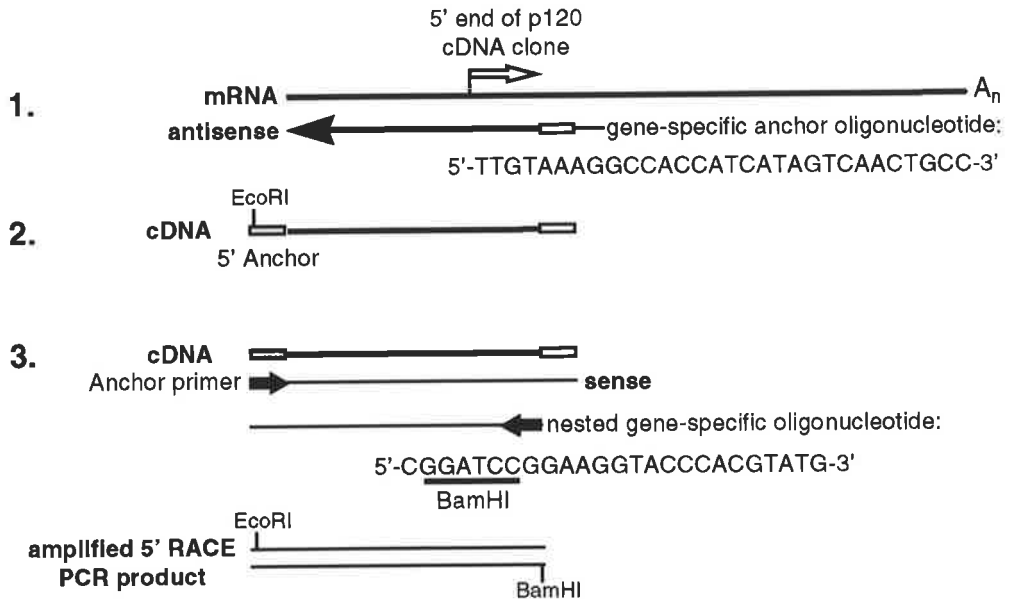
~ 1.1 kb ►
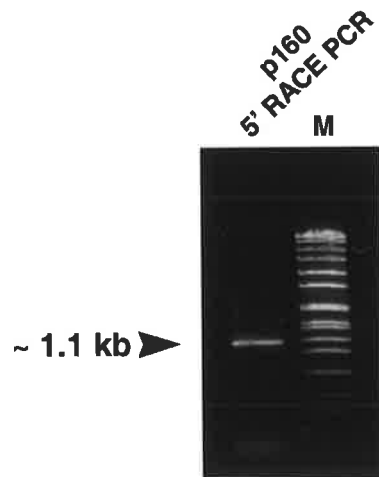
Figure 3.3   The p160 cDNA nucleotide sequence.

The entire p160 cDNA nucleotide sequence (1-4126) and predicted translation are shown. Underlined nucleotides (1-1087) correspond to the p160 5' RACE PCR product. The p67-derived tryptic peptide sequences are shown in bold. Residues in the *d* position of the two putative leucine zipper motifs are identified by a rectangle. Core peptide motifs of putative leucine charged domains (LCDs) are represented by a thick grey line above the sequence. The centrally-located acidic region is boxed. The carboxyl-terminal basic repeats are shown in bold and underlined. Note that the first nucleotide (A) of the p160 5' RACE PCR product is not shown.

```
  1                     MetAlaGluMetLysSerProThrLysAlaGluProAlaThrPro    15
  1   CGTGTTTGGCTCAGCATGGCGGAGATGAAGAGCCCTACGAAAGCTGAGCCTGCGACTCCC    60

 16   AlaGluAlaAlaGlnSerAspArgHisSerLeuLeuGluHisSerArgGluPheLeuAsp    35
 61   GCAGAAGCGGCGCAAAGCGACCGCCACAGCCTGCTGGAGCACAGCCGCGAGTTCTTGGAC   120

 36   PhePheTrpAspIleAlaLysProAspGlnGluThrArgLeuArgAlaThrGluLysLeu    55
121   TTCTTCTGGGACATTGCGAAACCGGATCAGGAAACGCGGCTCCGGGCCACGGAGAAGTTG   180

 56   LeuGluTyrLeuArgThrArgProAsnAspSerGluMetLysTyrAlaLeuLysArgLeu    75
181   TTGGAGTACTTGCGCACAAGGCCCAATGATTCGGAGATGAAATATGCCCTGAAGCGCCTA   240

 76   IleThrGlyLeuGlyValGlyArgGluAlaAlaThrAlaCysTyrSerLeuAlaLeuAla    95
241   ATCACTGGGCTTGGGGTGGGCCGAGAAGCCGCTACGGCCTGCTACAGCCTGGCGCTGGCA   300

 96   GlnLeuLeuGlnSerPheGluAspIleProLeuCysAspIleLeuAspGlnIleGlnGlu   115
301   CAGCTGTTGCAGTCTTTTGAAGACATCCCATTGTGTGACATCCTGGATCAGATACAAGAA   360

116   LysTyrSerLeuGlnAlaMetAsnLysAlaMetMetArgProSerLeuPheAlaAsnLeu   135
361   AAATACAGCCTACAAGCCATGAACAAGGCAATGATGAGACCTAGTCTCTTTGCAAACCTT   420

136   PheGlyValLeuAlaLeuPheGlnSerGlyArgLeuValLysAspLysGluAlaLeuMet   155
421   TTTGGAGTGCTAGCCCTCTTTCAGTCAGGCCGCCTAGTGAAGGACAAAGAGGCCCTGATG   480

156   LysSerValGlnLeuLeuLysIleLeuSerGlnHisProAsnHisLeuGlnGlyGlnPro   175
481   AAGTCCGTGCAATTGCTGAAGATCCTGTCCCAACACCCCAACCACTTACAGGGACAGCCA   540

176   IleLysAlaLeuValAspIleLeuSerGluValProGluSerMetPheGlnGluIleLeu   195
541   ATAAAGGCTCTGGTGGACATCCTCTCTGAGGTCCCAGAGTCCATGTTCCAGGAGATCCTG   600

196   ProLysValLeuLysGlyAsnMetLysValIleLeuArgSerProLysTyrLeuGluLeu   215
601   CCAAAGGTCCTCAAGGGTAACATGAAAGTGATCCTCCGCTCTCCCAAGTACTTGGAGCTC   660

216   PheLeuLeuAlaLysGlnArgValProThrLysLeuGluSerLeuMetGlySerValAsp   235
661   TTCCTCCTGGCTAAGCAGAGGGTGCCGACAAAGCTCGAGTCACTCATGGGCTCGGTTGAC   720

236   LeuPheSerGluAspAsnIleProSerLeuValAsnIleLeuLysValAlaAlaAsnSer   255
721   CTATTCTCAGAAGACAATATTCCCAGTCTGGTGAACATCCTGAAGGTGGCCGCCAACTCT   780

256   ValLysLysGluHisLysLeuProAsnValAlaLeuAspLeuLeuArgLeuAlaLeuLys   275
781   GTCAAGAAGGAGCACAAGCTGCCTAATGTGGCTCTGGACCTGCTCCGCCTGGCTCTCAAG   840

276   GluSerArgPheGluLeuPheTrpLysLysValLeuGluGluGlyLeuLeuLysAsnPro   295
841   GAGAGCAGATTCGAACTGTTCTGGAAGAAGGTTTTGGAGGAGGGGCTGCTGAAGAATCCG   900

296   SerTrpThrSerSerTyrMetCysPheArgLeuLeuGlyAlaSerLeuProLeuLeuSer   315
901   TCCTGGACATCCAGCTACATGTGCTTCCGCCTACTGGGTGCGTCTCTGCCGCTTCTGTCA   960

316   GluGluGlnLeuGlnLeuValMetArgGlyAspLeuIleArgHisPheGlyGluAsnMet   335
961   GAGGAGCAGTTGCAGTTGGTGATGCGAGGAGACTTGATCCGCCATTTTGGGGAGAACATG  1020

336   ValIleSerLysProGlnAsnLeuPheLysIleIleProGluIleSerThrTyrValGly   355
1021  GTTATTTCTAAGCCCCAAAACCTATTTAAGATCATCCCAGAGATAAGTACATACGTGGGT  1080

356   ThrPheLeuGluGlyCysGlnAspAspProLysArgGlnLeuThrMetMetValAlaPhe   375
1081  ACCTTCCTAGAGGGGTGCCAGGATGACCCTAAGCGGCAGTTGACTATGATGGTGGCCTTT  1140

376   ThrThrIleThrAsnGlnGlyLeuProValMetProThrPheTrpArgValThrArgPhe   395
1141  ACAACCATCACCAATCAAGGTCTCCCTGTCATGCCTACCTTCTGGCGTGTCACGCGGTTT  1200

396   LeuAsnAlaGluAlaLeuGlnSerTyrValAlaTrpLeuArgAspMetPheLeuGlnPro   415
1201  TTGAATGCTGAAGCCCTGCAGAGCTATGTGGCCTGGTTGCGGGACATGTTCCTGCAGCCT  1260

416   AspLeuAsnSerLeuValAspPheSerThrAlaAsnGlnLysArgAlaGlnAspAlaSer   435
1261  GACCTGAACTCCTTGGTTGACTTCAGCACTGCCAACCAGAAGAGAGCTCAGGACGCCTCG  1320

436   LeuAsnValProGluArgAlaValPheArgLeuArgLysTrpIleIleHisArgLeuVal   455
1321  TTGAATGTGCCTGAGCGAGCTGTATTCCGGCTCCGGAAGTGGATCATCCACCGCCTGGTC  1380

456   SerLeuValAspHisLeuHisLeuGluLysAspGluAlaValValGluGlnIleAlaArg   475
1381  AGCCTTGTGGATCATTTGCATCTGGAGAAGGATGAAGCTGTGGTTGAGCAAATAGCCAGG  1440
```

```
476   PheCysLeuPheHisAlaPhePheLysThrLysLys**AlaThrProGlnIleProGluThr**   495
1441  TTTTGCTTGTTCCATGCCTTCTTTAAGACGAAGAAGGCTACGCCCCAGATCCCAGAGACG      1500

496   **Lys**GlnHisPheSerPheProLeuAspAspArgAsnArgGlyValPheValSerAlaPhe   515
1501  AAGCAGCACTTCTCCTTCCCTTTGGACGACCGCAACCGTGGGGTCTTTGTCAGTGCCTTC      1560

516   PheSerLeuLeuGlnThrLeuSerValLysPheArgGlnThrProAspLeuAlaGluAsn      535
1561  TTCAGCCTACTGCAGACGCTCAGTGTGAAGTTCAGGCAGACACCAGACCTGGCTGAAAAT      1620

536   GlyLysProTrpThrTyrArgLeuValGlnLeuAlaAspMetLeuLeuAsnHisAsnArg     555
1621  GGGAAGCCTTGGACTTACCGCCTGGTTCAGTTGGCAGACATGCTGTTGAACCATAACCGC      1680

556   AsnValThrSerValThrSerLeuThrThrGlnGlnArgGlnAlaTrpAspGlnMetMet     575
1681  AATGTAACCAGTGTGACATCCTTGACAACACAGCAGCGTCAGGCCTGGGACCAGATGATG      1740

576   SerThrLeuLysGluLeuGluAlaArgSerSerGluThrArgAlaIleAlaPheGlnHis     595
1741  AGTACTCTGAAGGAATTAGAGGCCCGCTCCTCTGAGACCAGGGCCATTGCCTTCCAGCAC      1800

596   LeuLeuLeuLeuValGlyLeuHisIlePheLysSerProAlaGluSerCysAspValLeu     615
1801  CTGCTGCTTCTGGTGGGCCTCCACATCTTCAAGTCCCCTGCAGAAAGCTGTGATGTCCTA      1860

616   GlyAspIleGlnThrCysIleLysLysSerMetGluGlnAsnProArgArgSerArgSer     635
1861  GGAGACATTCAGACTTGCATCAAGAAAAGCATGGAGCAGAATCCCCGCCGATCACGCTCT      1920

636   ArgAlaLysAlaSerGlnGluProValTrpValGluValMetValGluIleLeuLeuSer     655
1921  AGAGCCAAAGCCTCCCAGGAGCCAGTATGGGTGGAGGTGATGGTGGAGATCTTGCTGTCC      1980

656   LeuLeuAlaGlnProSerAsnLeuMetArgGlnValValArgSerValPheGlyHisIle    675
1981  TTGCTGGCTCAACCCAGCAACTTGATGCGCCAGGTGGTCCGGAGTGTATTTGGTCATATC     2040

676   CysProHisLeuThrProArgCysLeuGlnLeuIleLeuAlaValLeuSerProValThr    695
2041  TGTCCCCACCTTACTCCACGTTGTCTGCAGCTAATCCTGGCTGTGCTCAGCCCTGTGACA     2100

696   AsnGluAspGluAspAspAsnValValValThrAspAspAlaAspGluLysGlnLeuGln    715
2101  AACGAGGATGAGGATGACAACGTGGTGGTCACTGATGACGCTGATGAGAAGCAGCTGCAG     2160

716   HisGlyGluAspGluAspSerAspAsnGluAspAsnLysAsnSerGluSerAspMetAsp    735
2161  CATGGAGAGGACGAAGACTCAGATAATGAGGACAATAAGAACTCAGAGAGTGACATGGAC     2220

736   SerGluAspGlyGluGluSerGluGluGluAspArgAspLysAspValAspProGlyPhe    755
2221  AGTGAGGATGGGGAAGAAAGTGAAGAGGAGGACCGTGACAAAGATGTGGACCCAGGCTTC     2280

756   ArgGlnGlnLeuMetGluValLeuLysAlaGlyAsnAlaLeuGlyGlyValAspAsnGlu    775
2281  CGTCAACAGTTGATGGAAGTGTTAAAAGCTGGGAATGCATTGGGTGGAGTGGACAACGAG     2340

776   GluGluGluGluLeuGlyAspGluAlaMetMetAlaLeuAspGlnAsnLeuAlaSerLeu    795
2341  GAGGAGGAGGAGCTTGGGGATGAGGCCATGATGGCCCTGGACCAGAACCTGGCCAGCCTA     2400

796   PheLysGluGlnLysMetArgIleGlnAlaArgAsnGluGluLysAsnLysLeuGlnLys    815
2401  TTTAAAGAGCAGAAGATGCGCATCCAGGCCCGGAATGAGGAGAAAAACAAGCTACAGAAG     2460

816   GluLysLysLeuArgArgAspPheGlnIleArgAlaLeuAspLeuIleGluValLeuVal    835
2461  GAGAAGAAGCTCCGACGGGACTTCCAAATCAGGGCACTAGACCTGATCGAGGTGCTGGTG     2520

836   ThrLysGlnProGluHisProLeuIleLeuGluLeuLeuGluProLeuLeuAsnValIle    855
2521  ACCAAGCAGCCTGAGCACCCCCTGATCCTGGAACTACTTGAGCCACTGCTGAACGTGATC     2580

856   GlnHisSerMetArgSerLysGlySerThrLysGlnGluGlnAspLeuLeuHisLysThr    875
2581  CAGCACAGCATGCGCAGCAAAGGCTCCACCAAGCAGGAGCAGGACCTCCTGCACAAGACC     2640

876   AlaArgIlePheMetHisHisLeuCysArgAlaArgArgTyrCysHisGluValGlyPro    895
2641  GCCCGCATCTTCATGCACCACCTGTGTCGTGCCCGCCGCTACTGCCACGAGGTGGGGCCG     2700

896   CysAlaGluAlaLeuHisAlaGlnValGluArgLeuValGlnGlnAlaGlySerGlnAla    915
2701  TGTGCAGAGGCTCTGCATGCCCAGGTGGAGAGGCTTGTGCAGCAGGCTGGCAGCCAGGCT     2760

916   AspAlaSerValAlaLeuTyrTyrPheAsnAlaSerLeuTyrLeuLeuArgValLeuLys    935
2761  GATGCCTCTGTCGCCCTATACTATTTCAATGCCTCTCTGTACCTGCTGCGAGTCCTCAAG     2820

936   GlyAsnThrAsnLysArgHisGlnAspGlyHisLysLeuHisGlyAlaAspThrGluAsp    955
2821  GGCAACACCAATAAGAGGCACCAAGATGGTCATAAGTTACACGGAGCTGACACAGAGGAC     2880
```

```
 956  SerGluAspGlnAlaAlaAsnCysLeuAspLeuAspPheValThrArgValTyrSerAla    975
2881  TCAGAGGACCAGGCTGCTAACTGCTTAGACTTGGACTTTGTGACCCGGGTGTATTCAGCA   2940

 976  SerLeuGluSerLeuLeuThrLysArgAsnSerSerLeuThrValProMetPheLeuSer    995
2941  TCACTGGAATCTCTTCTGACCAAGCGTAACAGCTCACTTACGGTCCCCATGTTCCTCAGC   3000

 996  LeuPheSerArgTyrProValIleCysLysAsnLeuLeuProValLeuAlaGlnHisVal   1015
3001  CTCTTCTCCAGATACCCAGTGATCTGTAAGAACCTGCTTCCCGTCCTGGCTCAGCATGTG   3060

1016  AlaGlyProSerArgProArgHisGlnAlaGlnAlaCysLeuMetLeuGlnLysThrLeu   1035
3061  GCTGGCCCATCTCGGCCCCGCCATCAGGCCCAGGCATGCCTGATGCTCCAGAAGACTCTG   3120

1036  SerAlaArgGluLeuArgValCysPheGluAspProGluTrpGluGlnLeuIleThrGln   1055
3121  TCTGCACGAGAGCTGAGAGTCTGTTTTGAGGATCCTGAGTGGGAGCAGCTGATTACCCAA   3180

1056  LeuLeuGlyLysAlaThrGlnThrLeuGlnThrLeuGlyGluAlaGlnSerLysGlyGlu   1075
3181  CTCTTGGGAAAGGCCACCCAGACCCTGCAGACTCTTGGGGAGGCACAGAGCAAGGGGGAG   3240

1076  HisGlnLysGluLeuSerIleLeuGluLeuLeuAsnThrLeuLeuArgThrValAsnHis   1095
3241  CACCAGAAGGAGCTATCCATCTTGGAGCTGCTGAACACTCTGTTGAGGACAGTCAATCAC   3300

1096  GluLysLeuSerValAspLeuThrAlaProLeuGlyValLeuGlnSerLysGlnGlnLys   1115
3301  GAGAAGCTGTCTGTGGACCTCACTGCTCCCTTGGGCGTGCTTCAGAGCAAGCAACAGAAG   3360

1116  LeuGlnGlnSerLeuGlnGlnGlyAsnHisSerSerGlySerAsnArgLeuTyrAspLeu   1135
3361  CTGCAGCAAAGCCTGCAGCAGGGGAATCACTCATCTGGCTCCAATCGCCTCTATGATCTC   3420

1136  TyrTrpGlnAlaMetArgMetLeuGlyValGlnArgProLysSerGlu**LysLysAsnAla**  1155
3421  TACTGGCAGGCCATGAGGATGCTAGGAGTCCAACGTCCAAAGTCAGAAAAGAAGAATGCC   3480

1156  **Lys**AspIleProSerAspThrGlnSerProValSerThr**LysArgLysLysLys**GlyPhe   1175
3481  AAGGATATTCCTAGTGACACCCAGAGCCCCGTCAGCACAAAGCGGAAGAAAAAGGGATTC   3540

1176  LeuProGluThr**LysLysArgLysLysLeuLys**SerGluGlyThrThrProGluLysAsn   1195
3541  TTGCCAGAGACCAAGAAGCGAAAGAAACTTAAATCTGAGGGCACCACACCAGAAAAGAAT   3600

1196  AlaAlaSerGlnGlnAspAlaValThrGluGlyAlaMetProAlaAlaThrGlyLysAsp   1215
3601  GCTGCGTCACAGCAGGATGCAGTGACAGAGGGTGCCATGCCTGCTGCCACTGGTAAAGAC   3660

1216  GlnProProSerThrGly**LysLysLysArgLysArgValLys**AlaSerThrProSerGln   1235
3661  CAGCCCCCCAGCACAGGCAAGAAGAAAAGGAAGAGGGTAAAGGCCAGCACCCCATCCCAG   3720

1236  ValAsnGlyIleThrGlyAlaLysSerProAlaProSerAsnProThrLeuSerProSer   1255
3721  GTGAATGGGATAACTGGGGCCAAGAGTCCAGCTCCCAGTAACCCCACCCTAAGCCCCAGC   3780

1256  ThrProAlaLysThrProLysLeuGln**LysLysLysGluLys**LeuSerGlnValAsnGly   1275
3781  ACCCCTGCCAAGACCCCCAAAACTGCAGAAGAAAAAGAGAAGCTGTCACAGGTGAATGGA   3840

1276  AlaThrProValSerProIleGluProGluSerLysLysHisHisGlnGluAlaLeuSer   1295
3841  GCCACTCCTGTGTCCCCCATAGAGCCTGAAAGCAAAAAGCATCATCAGGAGGCACTCAGC   3900

1296  ThrLysGluValIleArgLysSerProHisProGlnSerAlaLeuPro**LysLysArgAla**  1315
3901  ACAAAGGAGGTCATAAGAAAGTCCCCCCACCCCCAGTCTGCCCTGCCAAAGAAAAGAGCA   3960

1316  **Arg**LeuSerLeuValSerArgSerProSerLeuLeuGlnSerGlyVal**LysLysArgArg**  1335
3961  AGGCTGTCTCTGGTGAGCAGGAGCCCCAGCCTGTTACAGAGTGGGGTCAAGAAAAGGAGA   4020

1336  **Val**AlaSerArgArgValGlnThrPro***                                   1344
4021  GTGGCCAGCAGGAGAGTGCAGACACCTTGAGTGTGTATAGTCTGTTTCCCTGCCCTAGAG   4080

4081  ACTCCTATTTTTTCACCAATATTTTAATAAACAATCCATGATGCTA                 4126
```

Figure 3.4 Size comparison of *in vitro* translated p160 with p120 and endogenous p160.

(A) Comparison of the *in vitro* translation products of (1) the extended 3.1 kb cDNA clone (4.1 kb) containing nucleotides isolated by 5' RACE PCR and (2) the 3.1 kb cDNA clone. The corresponding approximate molecular weights of these proteins are shown and have been determined from molecular weight markers (not shown).

(B) Size comparison of the 160 kDa *in vitro* translation product with endogenous p160 and p67. Endogenous proteins were affinity purified from nuclear extracts of metabolically labelled FDC-P1 cells using the GST-NRD2 fusion protein, which contains the c-Myb leucine zipper motif. Proteins are identified by labels shown on the right. The positions of molecular weight markers (kDa) are shown on the left.
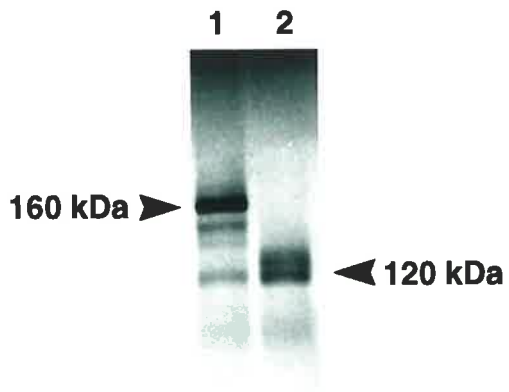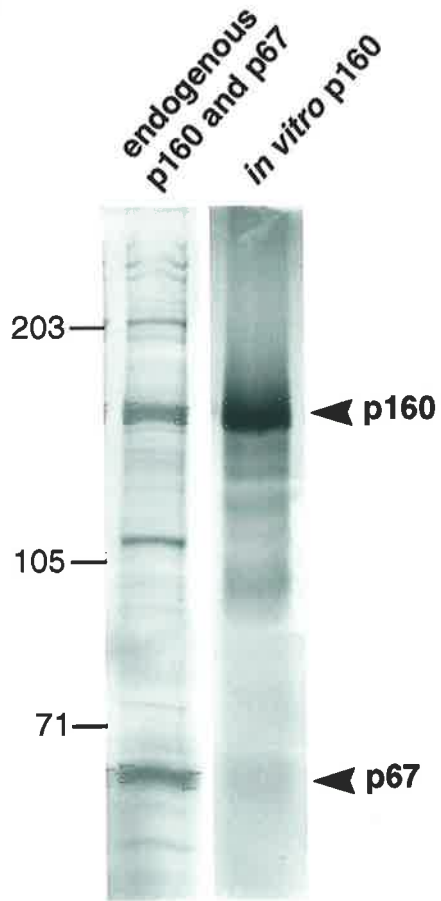
S.aureus V8 protease. Comparison of the peptide profiles of endogenous and *in vitro* translated p160 demonstrates identical peptide profiles (Figure 3.5). In addition, peptide mapping of endogenous p67 verifies the relatedness of p160 (endogenous and *in vitro* translated) and p67, as demonstrated by the similar peptide profiles (Figure 3.5). Collectively, these results support the identification of a full-length 4126 bp murine cDNA clone corresponding to p160, that contains peptide sequences derived from p67.

3.2.2   Sequence analysis of the p160 cDNA clone

Several prominent features are evident in the p160 cDNA sequence. One striking aspect of this sequence is the relatively large proportion of leucine residues (13%). The three remaining tryptic peptide sequences (1. PDQETRLAT 2. VILRSPK 3. VLEEG) derived from p67, which are absent from the 3.1 kb cDNA clone, are present in the p160 5' RACE PCR sequence (Figure 3.3). Interestingly, all five p67-derived peptide sequences are clustered in the amino-terminal region of p160 (Figure 3.6). Two putative leucine zipper motifs have been identified in the p160 sequence and contain the following residues in heptad repeat position *d*; ^{307}L-L-L-I-M^{335} (LZ1) and ^{900}L-L-Q-L-L^{928} (LZ2) (Figure 3.3). Structurally, heptad repeat position *d*, together with position *a*, form a hydrophobic interaction face (O'Shea *et al.*, 1989). In a leucine zipper motif, position *d* is usually held by a leucine residue. The *d* positions in p160 LZ1 and LZ2 predominantly contain hydrophobic residues, however a hydrophilic residue (glutamine 914) is present in the third *d* position of LZ2. The *a* positions of LZ1 and LZ2 contain 2/5 and 3/5 hydrophobic residues, respectively. The *e* and *g* positions of a heptad repeat contribute to leucine zipper dimerization specificity and stability, and are commonly held by charged residues. Neither LZ1 or LZ2 contain a predominance of charged residues in the *e* and *g* positions (3/10 and 1/10, respectively). Unusually, a proline (residue 312) is situated in LZ1. The presence of a proline residue is uncommon in α-helices, which constitute the secondary structure of heptad repeats (O'Shea *et al.*, 1989; Landschulz *et al.*, 1988) The p160 sequence contains multiple 'leucine charged domains' (LCDs), recently described as a protein-protein interaction motif of transcriptional coactivators and nuclear hormone receptors (Torchia *et al.*, 1997; Heery *et al.*, 1997). Leucine charged domains share a core peptide motif (LXXLL; single letter amino acid code) and are predicted to form an amphipathic α-helical structure, similar, but smaller to that of leucine zippers (Torchia *et al.*, 1997; Heery *et al.*, 1997). The p160 sequence contains five putative LCDs that conform to the core 'LXXLL'

Figure 3.5 Comparative peptide mapping of endogenous p160 and p67, and *in vitro* translated p160.

Endogenous p160 and p67 were affinity purified from nuclear extracts of metabolically labelled FDC-P1 cells using the GST-NRD2 fusion protein, which contains the c-Myb leucine zipper motif. The p160 *in vitro* translated protein was generated from the full-length 4.1 kb cDNA clone (see section 3.2.2 for details). Each protein was partially digested with the following varying amounts of *Staphylococcus aureus* V8 protease; endogenous p160: 1000, 500, 100 ng; *in vitro* translated p160: 1000, 500, 100 ng; endogenous p67: 500, 300, 50 ng. The peptide profiles of each partially digested protein are shown. The positions of molecular weight markers (kDa) are shown on the left.
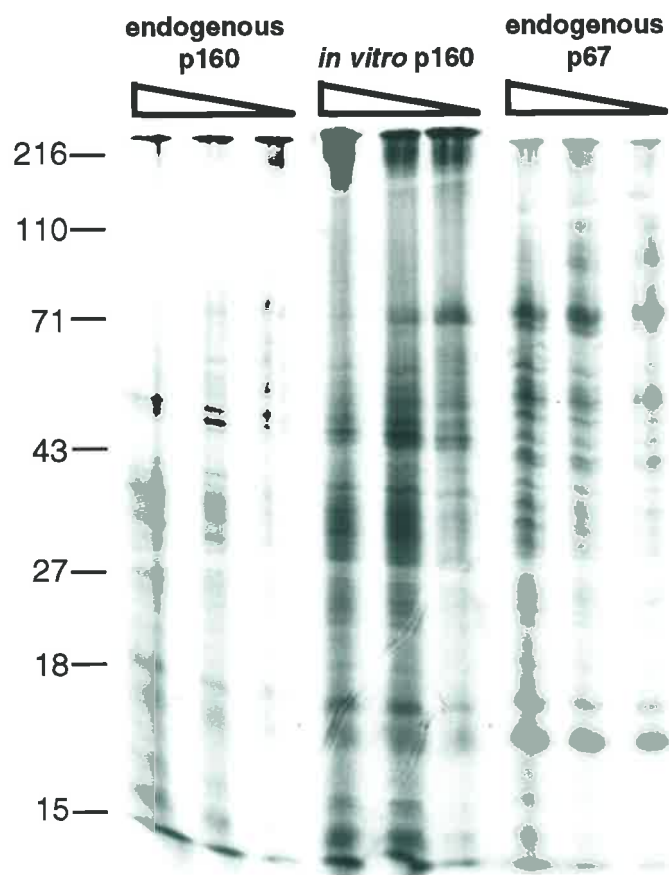
Figure 3.6 Schematic representation of murine p160.

Features of the p160 predicted polypeptide sequence (1344 residues) are represented schematically. Numbers below the schematic represent amino acid positions. The five p67-derived tryptic peptide sequences; 1. PDQETRLAT, 2. VILRSPK, 3. VLEEG, 4. VDHLHLEK, 5. ATPQIPETK, are represented by thick vertical lines, and the numbers 1-5 above the schematic. The p67-derived peptide sequences are clustered in the amino-terminal region of p160. 'L' denotes a putative leucine zipper motif. Residues in the *d* position of each putative leucine zipper motif are shown above the schematic. 'A' represents the acidic region. Core peptide motifs of putative leucine charged domains (LCDs) are represented by yellow, thin vertical lines. The basic repeats within the carboxyl-terminus are represented by grey vertical lines, and are embedded in a proline- and serine-rich region (represented by red rectangle). The p67* protein is derived from the p160 cDNA sequence (residues 1-580). The relationship of p67* with respect to the p160 protein is indicated by a black bar beneath the schematic. See text for details.
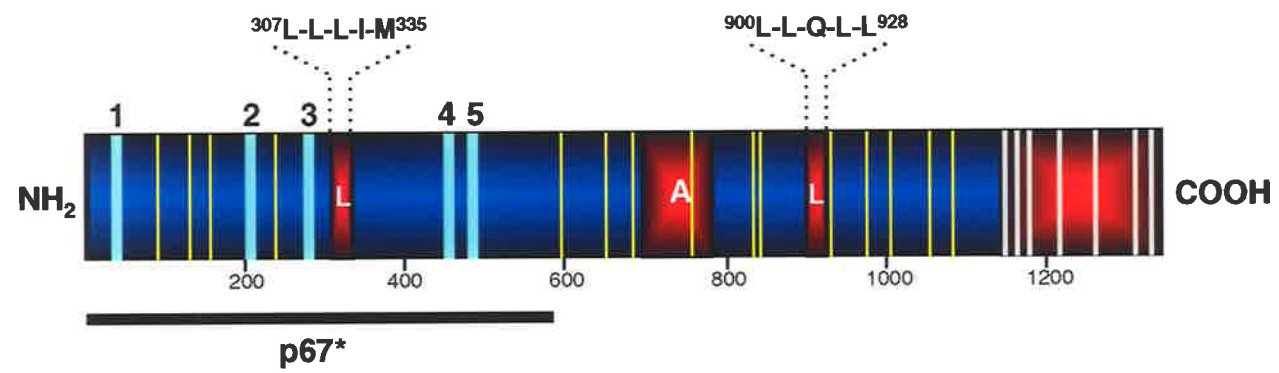
peptide motif (Figure 3.3). An additional ten short peptide sequences partially conform to the LCD core motif and alternatively contain an aliphatic amino acid other than leucine in one or more of the positions held by a leucine residue (Figure 3.3). A centrally-located region of p160 contains a cluster of predominantly acidic residues (Figure 3.3). The carboxyl terminus contains seven short repeats of basic residues, namely lysine and arginine, that are emedded in a proline- and serine-rich region (Figure 3.3). These repeats comprise the following consensus: $K^K/_R XX^K/_R$ (single letter amino acid code). Residues in the 'X' position of this consensus are often also lysine or arginine. An abundance of putative phosphorylation sites are present in the p160 sequence (not shown in Figure 3.3) that could be substrates for various kinases including MAP kinase (substrate recognition sequence: $PX^S/_T P$), $p34^{cdc2}$ ($^S/_T PX^K/_R$), casein kinase II ($^S/_T XX^E/_D$), and protein kinase C ($^S/_T X^R/_K$) (Figure 3.3) (Kennelly and Krebs, 1991; Gonzalez *et al.*, 1991). The features of the p160 sequence are summarised schematically in Figure 3.6.

Nucleic acid and protein databases were searched to identify sequence similarities and for the possibility of gaining insight into the potential functional role(s) of p160. These searches indicate that the p160 sequence is novel. A number of incomplete p160 cDNAs were identified and correspond to expressed sequence tags (ESTs) with no informative functional content. The *Drosophila* lethal (l) 1Bi protein (GenBank accession number U20542) (Voelker *et al.*, 1989), and a yeast hypothetical protein (Yelo55cp) of unknown function (GenPept accession number U18795), were identified as the highest scoring matches with p160. It is evident that the p160 acidic region shows varying degrees of sequence similarity to numerous proteins (data not shown). Other relatively high scoring matches, for example upstream binding factors (UBFs), have mostly been identified on the basis of sequence similarity with the p160 acidic region and do not display any similarity to other regions of p160, thus casting doubt on the significance of such matches. Recently, a rat peptide sequence, apparently homologous to p160, has been identified within the GenPept database (accession number U83590) and encodes a 160 kDa nucleolar protein termed 'PIP' that interacts with the PAR family of transcription factors (cited in GenPept entry). However, the 'PIP' peptide sequence lacks the first 67 amino acids of the p160 sequence. The p160 cDNA sequence has been submitted to GenBank (accession number U63648).

3.2.3 Detection of a single mRNA transcript with p160 and p67 sequences

Northern analysis of polyadenylated RNA from mostly murine cell lines with the partial p160 cDNA clone (3.1 kb) detected a single mRNA species of ~4.5 kb in each murine cell line examined (D. Favier; data not shown). Subsequent analysis of this same RNA blot with the ~1.1 kb p160 5'RACE PCR product also detected a single ~4.5 kb mRNA species in each murine cell line examined, including FDC-P1 and WEHI-3B, in which both p160 and p67 proteins were detected (Figure 3.7) (Favier and Gonda, 1994). The size of this transcript is comparable with the full-length 4.1 kb p160 cDNA clone and could accomodate untranslated regions (UTRs), including a short 3' UTR, and a poly A tail. It is apparent that a single transcript encodes polypeptides corresponding to both p160 and p67, as no additional transcripts were detected from FDC-P1 or WEHI-3B RNA. No transcript was detected with the p160 sequence in the human HL-60 cell line (Figure 3.7). This finding is consistent with the inability to previously detect either p160 or p67 proteins with the c-Myb leucine zipper fusion protein (GST-NRD2) in the HL-60 cell line, or any of the other human cell lines examined (Favier and Gonda, 1994). The inability to detect a human p160 mRNA species is not due to lack of RNA on this blot as the positive control, glyceraldehyde-3-phosphatedehydrogenase (GAPDH), transcript was detected (Tavner *et al.*, 1998; see appendix).

3.2.4 Binding properties of *in vitro* translated p160 and a derivative protein; p67*

Sequence information derived from p67 facilitated the isolation of a murine cDNA clone corresponding to p160 mRNA. The molecular basis underlying the relatedness of p160 and p67 was unknown. Several lines of evidence pointed to a precursor-product relationship for p160 and p67: (1) The p160 and p67 proteins have common peptides, as demonstrated by comparative peptide mapping (Figure 3.5) (Favier and Gonda, 1994); (2) Sequence analysis revealed that all five of the p67 tryptic peptide sequences are clustered in the amino-terminal region of p160 (Figure 3.3 and Figure 3.6); (3) A single ~4.5 kb mRNA species was detected with p160 and p67 sequences in cell lines that express both p160 and p67 proteins (Figure 3.7). These points support the notion that p67 is derived from p160 by proteolytic cleavage. Experimental evidence in support of a precursor-product relationship for p160 and p67 is presented in Chapter 4.

Figure 3.7 Northern analysis of murine and human cell lines.

Polyadenylated RNA (2 µg) from the indicated murine and human cell lines was probed with the p160 5' RACE PCR product (nucleotides 1-1087). An approximately 4.5 kb mRNA species was detected in each murine cell line, but not in the human HL-60 cell line. These cell lines correspond to the following cell types; HL-60: human promyelocytic leukaemia, NIH3T3: murine fibroblast, J774: murine macrophage, MTHC: abbreviation for *myb*-transformed haemopoietic cells, derived by retroviral infection of murine foetal liver cells (Gonda *et al.*, 1993), FDC-P1 and WEHI-3B: murine myelomonocytic progenitor-like. Cell type descriptions are cited from Favier and Gonda (1994). This blot has been stripped and reprobed with a glyceraldehyde-3-phosphate dehydrogenase (GAPDH) cDNA, as a control for RNA loading (Tavner *et al.*, 1998; see appendix).

It was therefore decided to construct a 'p67-encoding' cDNA sequence from the p160 cDNA in order to further study the properties of p67, which, for the reasons discussed in Chapter 4, is believed to represent the amino-terminal region of p160. On the basis of the location of the p67 peptide sequences within p160, a carboxyl-truncated p160 clone was generated by introducing a termination codon into the p160 cDNA sequence directly after nucleotide position 1755, that would yield a protein with a predicted molecular mass of 67 kDa (Chapter 2.1.12 (II)). This carboxyl-truncated form of p160 (580 amimo acids) is termed p67* to distinguish it from endogenous p67, and encompasses all of the p67-derived tryptic peptide sequences (Figure 3.6). Comparison of *in vitro* translated p67* with endogenous p67 showed that the apparent molecular weight of p67* is slightly smaller (~60 kDa) than that of endogenous p67 (Figure 3.8A). Peptide mapping of endogenous p67 and *in vitro* translated p67* demonstrated subtle differences in the peptide profiles of these proteins (Figure 3.8B). The p67* protein lacks a small number of peptides with respect to those generated from endogenous p67 (Figure 3.8B). Together these data imply that p67* lacks carboxyl-terminal peptides present in endogenous p67. Consequently, p67* has been used to approximate the properties of endogenous p67.

The binding properties of *in vitro* translated p160 and p67* were examined using various leucine zipper fusion proteins to determine if these properties correlated with those of endogenous p160 and p67. The GST fusion proteins used in this assay correspond to those described by Favier and Gonda (1994). The c-Myb fusion proteins comprise a region of the negative regulatory domain (residues 325-411 of murine c-Myb) encompassing the leucine zipper motif, and contain either a wild type or mutated leucine zipper motif (Figure 3.9A). The mutated c-Myb leucine zipper motifs contain either proline or alanine substitutions for leucine residues in the third and fourth *d* positions (heptad repeat nomenclature) (Figure 3.9A) (Favier and Gonda, 1994). The proline substitutions are expected to completely disrupt the α-helical structure of the leucine zipper. The alanine substitutions are expected to maintain an α-helical structure, but disrupt hydrophobic interactions. The c-Jun fusion protein encompasses the basic region and adjacent leucine zipper motif (bZIP) of human c-Jun (residues 193-316) (Figure 3.9A) (Favier and Gonda, 1994). The p160 and p67* proteins were translated *in vitro* and used in a binding assay with a specific affinity matrix composed of glutathione sepharose-bound leucine zipper fusion protein (ie GST-NRD2, GST-L3,4P, GST-L3,4A or GST-Jun) Both p160 and p67* bound directly to GST-NRD2,

Figure 3.8   Size comparison and comparative peptide mapping of endogenous p67 and *in vitro* translated p67*.

(A) Endogenous p160 and p67 were affinity purified from nuclear extracts of metabolically labelled FDC-P1 cells using the GST-NRD2 fusion protein, which contains the c-Myb leucine zipper motif. The p67* protein is a carboxyl-truncated form of p160 and has been generated by introducing a termination codon into the p160 cDNA sequence (see section 3.2.4). The p67* protein is slightly smaller (~60 kDa) than endogenous p67 (67 kDa). The positions of molecular weight markers (kDa) are shown on the left. Proteins are identified by labels shown on the right.

(B)   The endogenous p67 and *in vitro* translated p67* proteins shown in part (A) were partially digested with varying amounts of *Stapphylococcus aureus* V8 protease (300, 100, 50 ng). The resulting peptide profiles generated from these proteins are shown. The arrow on the right indicates the absence of peptides from p67*, that are present in endogenous p67. The positions of molecular weight markers (kDa) are shown on the left.
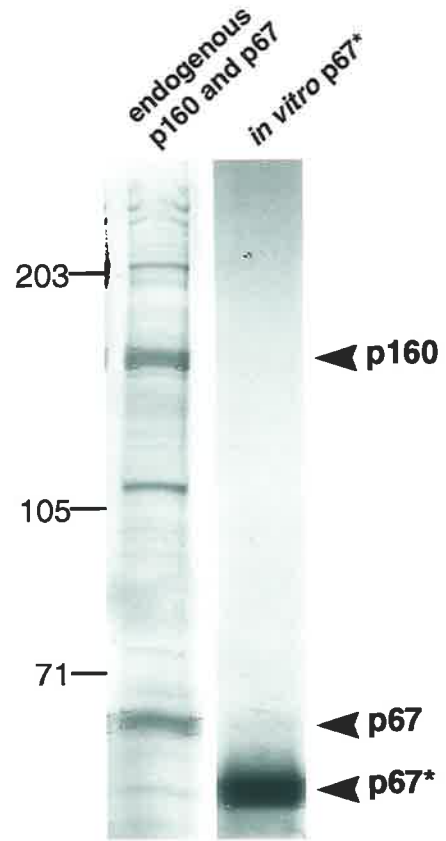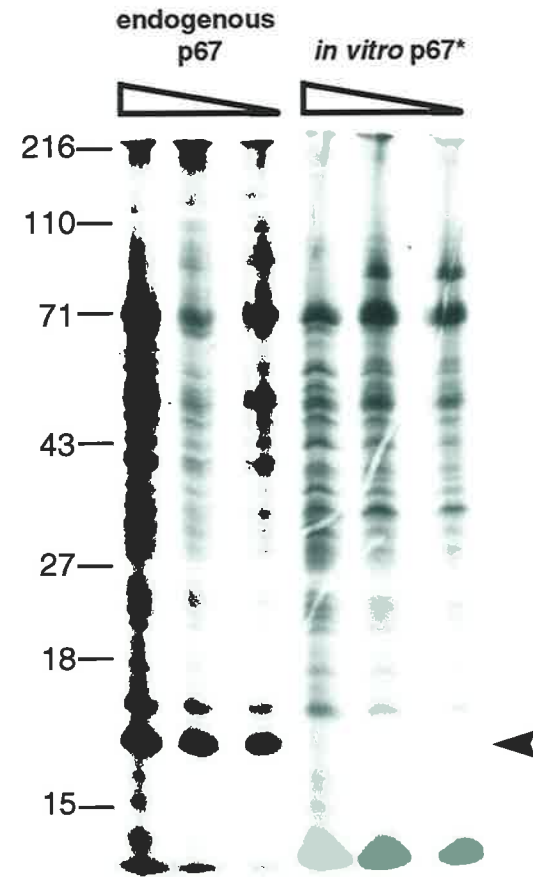
Figure 3.9   The binding properties of *in vitro* translated p160 and p67*.

(A)   A schematic representation of the GST fusion proteins used to analyse the binding properties of p160 and p67/p67*. These fusion proteins have been described by Favier and Gonda (1994). The GST-NRD2 fusion protein contains a region of the c-Myb negative regulatory domian (residues 325-411) that encompasses the leucine zipper motif (denoted as 'LZ'). The GST-L3,4P and GST-L3,4A fusion proteins each contain mutated c-Myb leucine zipper motifs that have been generated by substitution of leucine residues, in the third and fourth *d* positions, by proline or alanine, respectively. The GST-Jun fusion protein contains a region of human c-Jun (residues 193-316) that encompasses the basic-leucine zipper (bZIP) region (denoted as 'B' and 'LZ', respectively). This figure is modified from Favier and Gonda (1994).

(B)   The binding properties of *in vitro* translated p160 and p67*. Proteins were assessed by a binding assay using the GST fusion proteins; GST-NRD2, GST-L3,4P, GST-L3,4A and c-Jun (see above). The fractions of p160 and p67* bound to each GST fusion protein affinity matrix are shown. In addition, the GST-NRD2 unbound fraction is shown. Proteins are identified by labels shown on the right.

A

GST-NRD2

c-Myb NRD

GST | LZ

325   411

GST-L3,4P

GST | LZ

IIVHQGTILDNVKNLLEFAETLQFIDSFL

1   2   3   4   5

P   P

GST-L3,4A

GST | LZ

IIVHQGTILDNVKNLLEFAETLQFIDSFL

1   2   3   4   5

A   A

GST-Jun

c-Jun

GST | B | LZ

193   316

B

Bound-p67*   Bound-p160

NRD2   L3,4P   L3,4A   Jun   NRD2   L3,4P   L3,4A   Jun   NRD2 Unbound

◄ p160

◄ p67*

which contains the wild type c-Myb leucine zipper motif (Figure 3.9B). It was evident that a relatively small amount of p160 and p67* bound to both mutated leucine zipper fusion proteins (GST-L3,4P and GST-L3,4A) (Figure 3.9B). This binding reflects a non-specific interaction as binding of p160 and p67* to the mutant leucine zipper fusion proteins was diminished in the presence of BSA (data not shown).

In addition to its interaction with the c-Myb leucine zipper, *in vitro* translated p160 also bound to the c-Jun bZIP fusion protein (GST-Jun) (Figure 3.9B). In contrast, p67* did not bind significantly to the c-Jun bZIP fusion protein (Figure 3.9B). In summary, *in vitro* translated p160 and p67*, encoded by the p160 cDNA sequence, demonstrate the same binding specificities as their endogenous counterparts. That is, p160 binds to both the c-Myb leucine zipper motif and the c-Jun bZIP region (Favier and Gonda, 1994). In contrast, p67* (and endogenous p67) binds only to the c-Myb leucine zipper motif. The interaction of p160 and p67* with the c-Myb leucine zipper motif is specific as neither of these proteins bind significantly to fusion proteins containing mutated c-Myb leucine zipper motifs. The ability of p67* to bind to the c-Myb leucine zipper indicates that the 580 amino-terminal residues of p160 (ie p67*) contain a region that facilitates the interaction with the c-Myb leucine zipper (Figure 3.6). Similarly, the inability of p67* to bind to the bZIP region of c-Jun indicates that a region between p160 residues 580 and 1344 interacts with c-Jun (Figure 3.6).

In these binding assays, a high molecular weight smear was always observed with *in vitro* translated p160 (Figure 3.9) (Gardner, 1996). This smear was less pronounced in the absence of glutathione sepharose affinity matrix, and therefore suggests that p160 is covalently modified in the rabbit reticulocyte lysate and in the presence of the affinity matrix (glutathione sepharose-GST fusion protein). Treatment of the p160 lysate with phosphatase did not diminish the high molecular weight smear (data not shown). The p160 high molecular weight smear confounds quantitative analyses of binding assays. Identical binding analyses of *in vitro* translated p120 (encoded by the 3.1 kb cDNA clone) have been undertaken. Repeatedly, binding of p120 to GST-NRD2 was very inefficient, indicating that the amino terminus of p160 (absent from p120) may contain the Myb-interactive domain (data not shown). Interestingly, the p120 protein lacks part of the first putative leucine zipper motif (LZ1) of p160.

3.2.5 Binding analyses of epitope-tagged p160 and p67*

To study the properties of the cloned p160 and p67* proteins in a eukaryotic cellular context, it was necessary to be able to specifically detect and identify these proteins. In the absence of distinct p160 and p67 antibodies, the 'FLAG' epitope (IBI, USA) was introduced into the amino termini of both p160 and p67*. A PCR approach was initially attempted to introduce the FLAG epitope into the p160 cDNA sequence, but proved unsuccessful due to the large size of the oligonucleotide primers required. The FLAG octapeptide sequence (DYKDDDDK; single letter amino acid code) was introduced into p160 and p67* by site-directed mutagenesis (Chapter 2.2.1.11). The original translation initiation site of p160 and p67* (nucleotides 13-19) was replaced by an in-frame FLAG octapeptide sequence, downstream of an engineered translation initiation site (Kozak consensus) that also constitutes an *Nco*I site to facilitate cloning into the 'pact' mammalian expression vector (described by Nishina *et al.*, 1989) (Figure 2.1). Both p160 and p67* cDNA sequences containing the FLAG octapeptide sequence were cloned into the pact expression vector (Chapter 2.1.12 (III)).

Transient expression of FLAG-epitope-tagged p160 and p67* (termed p160FLAG and p67*FLAG) was examined in the human embryonic kidney 293T epithelial (HEK 293T) and murine NIH3T3 fibroblast cell lines. Transient transfection of HEK 293T cells with each expression construct and analysis of the corresponding total cell extracts with the anti-FLAG M2 monoclonal antibody (IBI, USA) showed that ectopically expressed p160FLAG and p67*FLAG were detected with observed molecular weights that correspond to those of endogenous p160 and *in vitro* translated p67*, respectively (Figure 3.10A).

The binding properties of ectopically expressed p160FLAG and p67*FLAG were examined using a similar assay to that previously described for *in vitro* translated p160 and p67* (section 3.2.4 and Chapter 2.2.3.6 (II)). Indeed, the binding properties of p160FLAG and p67*FLAG correlate with those of endogenous p160 and p67. In particular, p160FLAG and p67*FLAG bound specifically to the wild type c-Myb leucine zipper (GST-NRD2), but not to a mutated form (GST-L3,4P) (Figure 3.10B). In addition, p160FLAG, but not p67*FLAG, bound to a fusion protein containing the c-Jun bZIP region (Figure 3.10B). The binding specificities of the p160 carboxyl-truncated protein, p67*, are therefore identical to those observed for endogenous p67, indicating that this

Figure 3.10 Ectopic expression of epitope-tagged p160 and p67*, and the binding properties of these proteins.

(A) The p160 and p67* proteins containing the FLAG epitope at their amino termini, were expressed transiently from expression vectors transfected into HEK 293T cells. The corresponding proteins were detected in total cell extracts by Western analysis/ECL using the anti-FLAG M2 monoclonal antibody (1/500). Total cell extracts from mock-transfected HEK 293T cells were included as a control. Proteins are identified by labels shown on the right. The positions of molecular weight markers (kDa) are shown on the left.

(B) The binding properties of ectopically expressed p160FLAG and p67*FLAG. The binding propeties of these proteins were analysed by using the GST fusion proteins; GST-NRD2 (M), GST-L3,4P (M'), GST-Jun (J) (see Figure 3.9), and HEK 293T total cell extracts containing ectopically expressed p160FLAG and p67*FLAG. Bound and unbound fractions are shown. Note that only 50% of the unbound fraction is shown. Proteins were detected by Western analysis/ECL using the anti-FLAG M2 monoclonal antibody (1/500). Proteins are identified by labels shown on the right. The positions of molecular weight markers (kDa) are shown on the left.

A

p160FLAG p67*FLAG untransfected

250—

◀p160FLAG

98—

64—

◀p67*FLAG

50—

B

p160FLAG   p67*FLAG   p160FLAG   p67*FLAG

M M' J   M M' J   M M' J   M M' J

250—

◀p160FLAG

98—

64—

◀p67*FLAG

50—

BOUND   UNBOUND
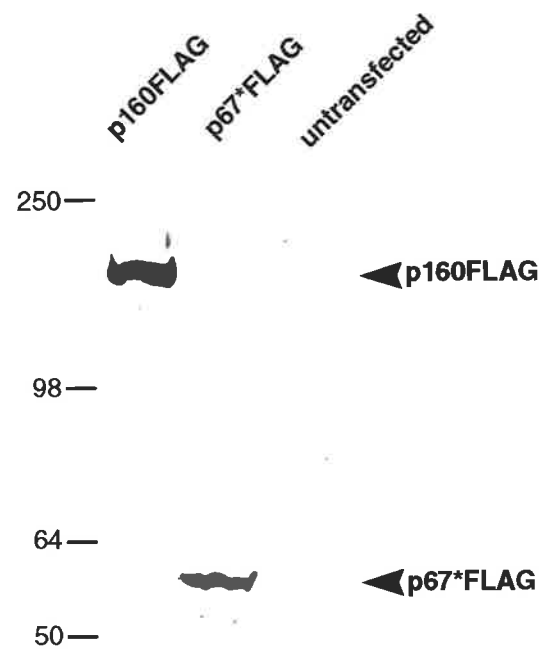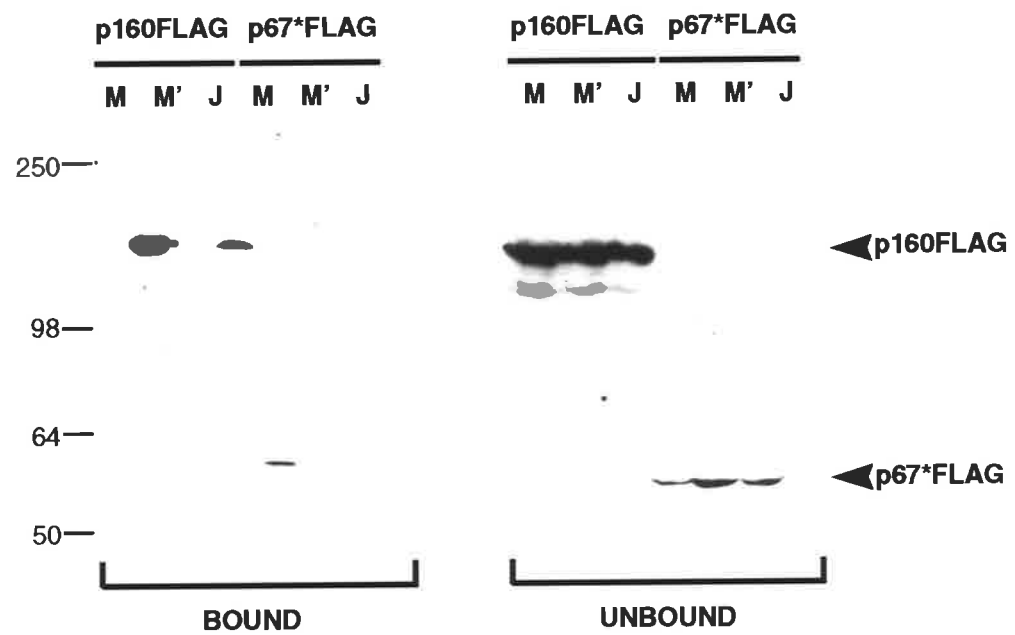
protein does approximate at least some of the properties of endogenous p67. In contrast to binding assays using *in vitro* translated p160, no high molecular weight smearing was observed with ectopically expressed p160FLAG, which demonstrates that the problems experienced with using *in vitro* translated proteins in binding assays can be alleviated by using proteins expressed in mammalian cells (Figure 3.10B). In addition, non-specific binding of p160 or p67* to the mutated leucine zipper fusion protein GST-L3,4P, (cf binding assays with *in vitro* translated proteins) was not seen with ectopically expressed proteins (Figure 3.10B). Comparison of the respective bound and unbound fractions suggests that p67*FLAG binds more efficiently to the c-Myb leucine zipper than p160FLAG (Figure 3.10B).

The aforementioned binding assays facilitated examination of the binding properties of cloned p160 and p67*. It was of particular importance to validate the interaction of p160 and p67* with the c-Myb leucine zipper motif by demonstrating that both p160 and p67* could interact specifically with full-length c-Myb within the cell. Coimmumoprecipitation of ectopically coexpressed p160FLAG or p67*FLAG with c-Myb was attempted to examine whether these proteins were capable of interacting within the intracellular environment. Analysis of repeated p67*FLAG transient expression in both HEK 293T and NIH3T3 cells resulted in two interesting observations. Firstly, transient expression of p67*FLAG in these cells was extremely erratic. Sometimes relatively good p67*FLAG expression was detected (eg see Figures 3.10 and 3.11), but on most occassions p67*FLAG expression was poor with respect to that observed for p160FLAG. Poor expression of p67*FLAG did not appear to be a consequence of 'cell death' as these cells did not differ morphologically to those expressing p160FLAG, or untransfected cells. It is unlikely that p67*FLAG represses transcription from the pact (mammalian expression vector; Figure 2.1) promoter as expression of c-Myb (from pact-c-*myb*) did not appear to be affected when cells were cotransfected with both p67*FLAG and c-Myb expression constructs (data not shown). Rapid degradation of p67*FLAG may account for the relatively low amount of protein detected. Alternatively, poor expression of p67*FLAG may be a consequence of an unstable mRNA transcript.

Secondly, analysis of nuclear and cytoplasmic extracts from HEK 293T cells transiently expressing p160FLAG or p67*FLAG showed that p67*FLAG could not be detected in the

nuclear fraction, as expected by analogy with endogenous p67, but rather was present only within the cytoplasmic fraction (Figure 3.11). The p67*FLAG protein was also detected entirely within the cytoplasmic fraction in the presence of coexpressed c-Myb, suggesting that if these proteins are able to interact *in vivo*, nuclear localisation of c-Myb is unable to assist translocation of p67*FLAG to the nucleus (Figure 3.11 and data not shown). The significance of the cytoplasmic localisation of p67*FLAG was examined further (see section 3.2.7). A small proportion of p160FLAG was also present within the cytoplasmic fraction, and is probably a consequence of protein over-expression (Figure 3.11).

Coimmunoprecipitation of transiently coexpressed p160FLAG and c-Myb was primarily pursued due to poor transient expression of p67*FLAG. This approach was undertaken by immunoprecipitating p160FLAG from total cell extracts with the anti-FLAG M2 monoclonal antibody and analysing the immunoprecipitates with the anti-Myb 1.1 and anti-FLAG M2 monoclonal antibodies. The anti-FLAG M2 monoclonal antibody is capable of immunoprecipitating p160FLAG from cell extracts (Figure 3.12B), but many problems were experienced whilst trying to demonstrate an interaction between p160FLAG and c-Myb *in vivo*. The most obvious problem was the non-specific interaction of c-Myb with Protein A sepharose. This was evident by the existence of c-Myb in anti-FLAG M2-immunoprecipitates lacking p160FLAG, and in control immunoprecipitates (Figure 3.12A). Interestingly, the control immunoprecipitate not only contains Myb (precipitated by non-specific interaction with Protein A sepharose), but also p160FLAG (Figure 3.12B). This observation may indirectly (as the anti-FLAG M2 antibody was not used) represent an interaction between c-Myb and p160FLAG because p160FLAG was not immunoprecipitated by the control monoclonal antibody (raised against the GM-CSF receptor common β chain) and did not bind to Protein A sepharose alone (data not shown). The c-Myb non-specific binding problem was eventually overcome by reducing the amount of Protein A sepharose and changing the salt concentration of the extraction and washing conditions. However, coimmunoprecipitation was not seen using these conditions (Chapter 2.2.3.9) (data not shown). A reciprocal coimmunoprecipitation was attempted using c-Myb rabbit polyclonal antibodies (provided by Dr L. Wolff) to immunoprecipitate ectopically expressed c-Myb. Immunoprecipitates were analysed using the anti-FLAG M2 and the anti-Myb 1.1 monoclonal antibodies. Optimal immunoprecipitation conditions were not achieved and thus this approach also remained unsuccessful. Due to time constraints, the

Figure 3.11 Subcellular localisation of ectopically expressed p160FLAG and p67*FLAG.

Nuclear and cytoplasmic extracts were prepared from transiently transfected HEK 293T expressing p160FLAG or p67*FLAG, or coexpressing either of these proteins with c-Myb. Extracts were fractionated by SDS-PAGE and proteins were detected by Western analysis/ECL using the anti-FLAG M2 monoclonal antibody (1/500). Cytoplasmic extracts prepared from untransfected HEK 293T cells were included as a control. Ectopically expressed proteins are identified by labels shown on the left and right.
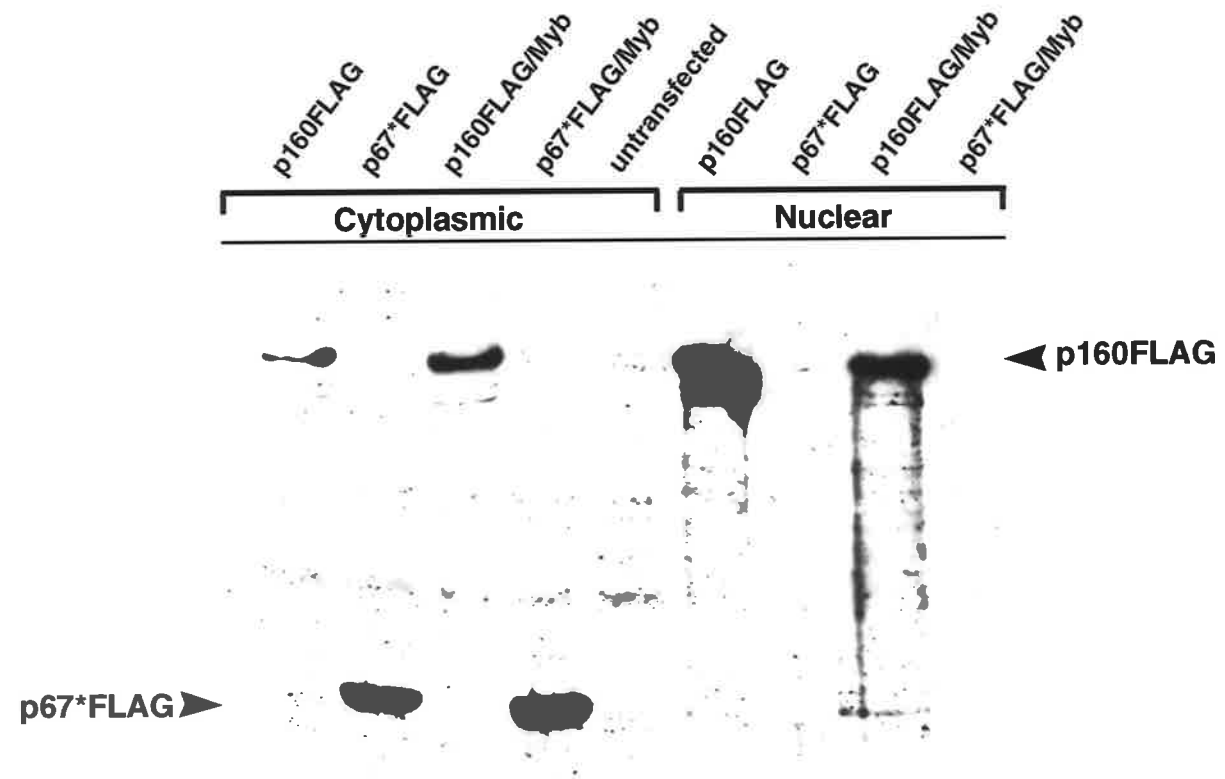
Figure 3.12 An attempt to coimmunoprecipitate ectopically expressed p160FLAG and c-Myb.

Total cell extracts were prepared from HEK 293T cells transiently expressing p160FLAG or c-Myb, or coexpressing these proteins. The anti-FLAG M2 monoclonal antibody was used as the immunoprecipitating antibody. Control immunoprecipitations were performed with an irrelevant monoclonal antibody (raised against the GM-CSF receptor common β-chain). Immunoprecipitates and a proportion (~1/8) of the corresponding anti-FLAG M2 supernatants are shown. The '+' or '-' symbols in each figure denote the presence or absence, respectively, of the p160FLAG- or c-Myb-encoding expression vectors in transfected cell extracts.

(A) Detection of c-Myb within immunoprecipitates and corresponding anti-FLAG M2 supernatants. The c-Myb protein (75 kDa) was detected by Western analysis/ECL using the anti-Myb 1.1 monoclonal antibody (1/500). The c-Myb protein is identified by the label shown on the right. The non-specific interaction of c-Myb with Protein A sepharose is illustrated by the presence of c-Myb in immunoprecipitates that lack p160FLAG (1st lane), and also by the control immunoprecipitate. Note the presence of c-Myb within extract prepared from cells not transfected with the pact-c-*myb* expression vector. The presence of c-Myb within this lane is probably due to overflow from adjacent wells during loading of the gel.

(B) Detection of p160FLAG within immunoprecipitates and corresponding anti-FLAG M2 supernatants. The protein blot shown in part (A) was analysed for the presence of p160FLAG, by stripping and reprobing with the anti-FLAG M2 monoclonal antibody (1/500). The p160FLAG protein is identified by the label shown on the right. Note the presence of p160FLAG within the control immunoprecipitate, which may be an indirect indication of coimmunoprecipitation (see text for details).

A

Anti-FLAG IP   control IP   unbound

| p160FLAG | – | + | + | + | – | + | + | + |
| Myb | + | – | + | + | + | – | + | + |

◄ Myb
75 kDa

B

Anti-FLAG IP   control IP   unbound

| p160FLAG | – | + | + | + | – | + | + | + |
| Myb | + | – | + | + | + | – | + | + |
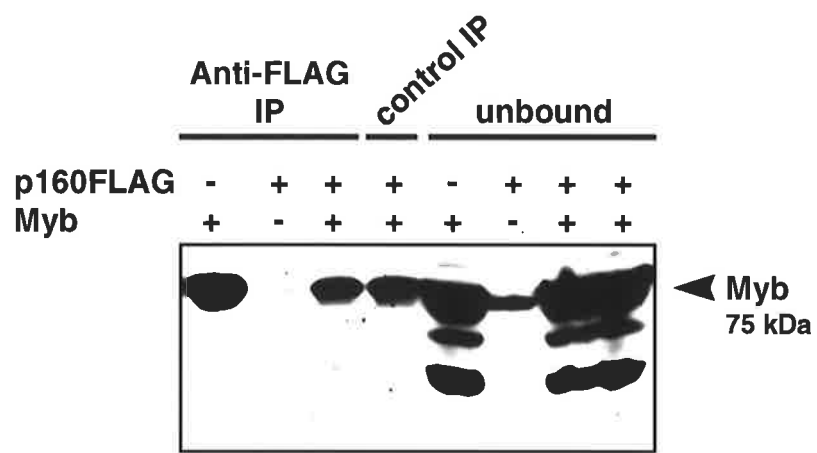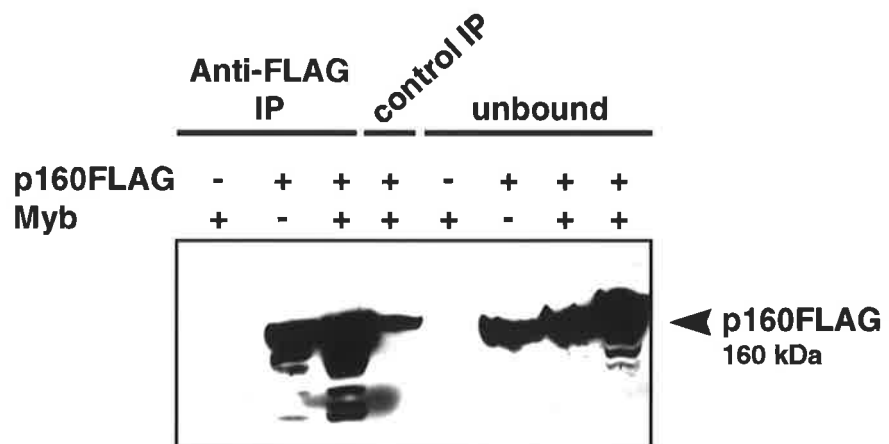
◄ p160FLAG
160 kDa

coimmunoprecipitation issue could not be resolved during the course of work described in this thesis. Subsequently though, a specific interaction between ectopically coexpressed p160FLAG and full-length c-Myb has recently been demonstrated by coimmunoprecipitation, using almost identical conditions to those described in Chapter 2.2.3.9 (Tavner *et al.*, 1998; see appendix). The only difference between these conditions is the presence of glycerol, which may have contributed to stabilization of the coimmunoprecipitated protein complex.

3.2.6 Generation of p160 polyclonal antisera that detect endogenous p160 and p67

Polyclonal antisera were raised against regions of the p160 amino and carboxyl termini in order to generate specific detection reagents for the endogenous proteins. Fusion proteins containing p160 amino-terminal residues 1-326 or carboxyl-terminal residues 1046-1344 were constructed as GST fusion proteins (Chapter 2.1.12 (V)). The amino-terminal 326 residues (1-326) comprise a region of p160 containing three p67-derived peptide sequences, four putative LCDs and part of the first putative leucine zipper motif (Figure 3.13A). The carboxyl-terminal 298 residues (1046-1344) comprise a region of p160 containing all seven basic repeats, the entire proline/serine-rich region and two putative LCDs (Figure 3.13A). The amino-terminal residues (corresponding to nucleotides 1-1083) were originally cloned into the pGex-3X vector in order to yield an in-frame GST fusion protein. However, bacterial expression of this clone produced only the 27 kDa GST moiety (data not shown). Sequencing of several of these clones revealed that an extra nucleotide (G) had been inserted at position 1 within the 5' UTR of the p160 sequence, and consequently was not in frame with the GST sequence (data not shown). Insertion of an extra nucleotide into this position of the p160 sequence (within the pGEM3Z-p160 construct) has probably occurred during subcloning, as the original p160 5' RACE PCR product does not contain this nucleotide. The p160 amino-terminal region was subsequently cloned into the pGex-4T-1 vector to produce an in-frame GST fusion protein (Chapter 2.1.12 (V)). The predicted molecular masses of GST-160N (p160 amino-terminal fusion protein) and GST-160C (p160 carboxyl-terminal fusion protein) are approximately 67 and 60 kDa, respectively. Expression of these constructs in *E.coli* produces fusion proteins that correspond approximately to the predicted molecular masses (Figure 3.13B). The GST-160C protein showed many degradation products upon expression, but also several proteins larger than the predicted molecular mass of 60 kDa (Figure 3.13B). The

Figure 3.13 The GST-p160 fusion proteins; a schematic representation, and bacterial expression.

(A) A schematic representation of the GST fusion proteins containing either amino- or carboxyl-terminal residues of p160. The GST-160N fusion protein contains p160 amino-terminal residues 1-326 and has a predicted molecular mass of 67 kDa. The GST-160C fusion protein contains p160 carboxyl-terminal residues 1046-1344 and has a predicted molecular mass of 60 kDa.

(B) Expression of the GST-160N and GST-160C fusion proteins in *E.coli*. Fusion proteins were purified from bacterial cell lysates with glutathione sepharose and analysed by SDS-PAGE. Proteins were detected by staining with Coomassie blue. The GST-160N (~67 kDa) and GST-160C (~60 kDa) fusion proteins are identified by the top and bottom arrows, respectively. Note the presence of higher molecular weight proteins in the GST-160C lane (see text for details). Molecular weight markers (kDa) are shown on the left and bovine serum albumin (BSA) standards (1.0, 0.5, 0.25 mg/ml) are shown on the right of the gel.

A

L-L-L-I-M   L-L-Q-L-L

1   2   3   4 5

NH₂   L   A   L   COOH

200   400   600   800   1000   1200

| GST | 1-326 |

| GST | 1046-1344 |

GST-160N (~ 67 kDa)   **GST-160C (~ 60 kDa)**

B

GST-160N   GST-160C   BSA

111—

71.5—

44—

28—
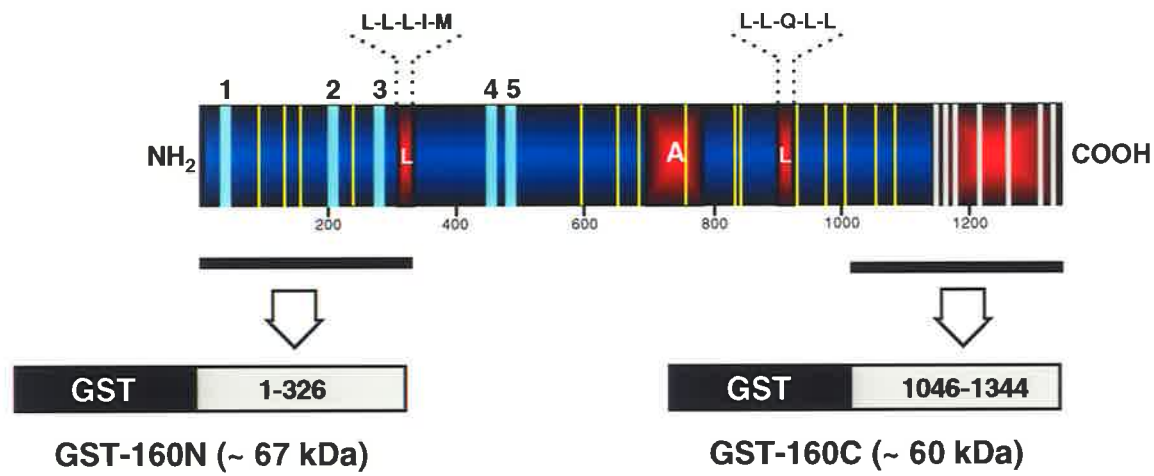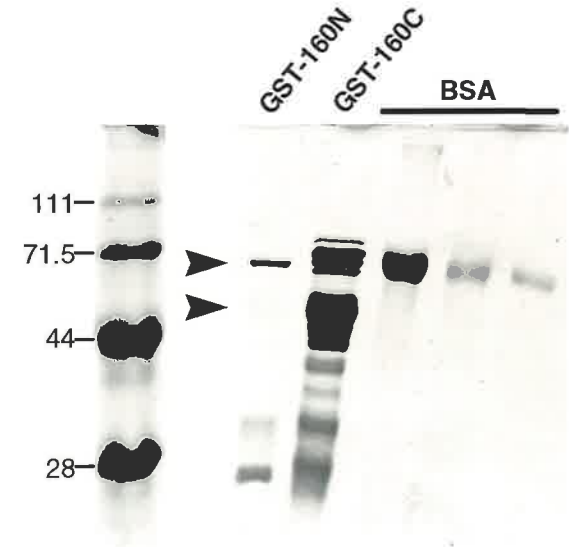
nature of these higher molecular weight proteins is unknown, however this observation is not uncommon for GST fusion proteins (Hengen, 1996).

The GST-160N and GST-160C fusion proteins were used to raise rabbit polyclonal antisera. The reactivity of these antisera with p160 and p67 was examined firstly by Western analysis using cell extracts prepared from HEK 293T cells in which p160FLAG and p67*FLAG were ectopically expressed. Antiserum raised against the amino-terminal region of p160 (termed 160N) recognised both p160 and p67* (Figure 3.14A). Antiserum raised against the carboxyl-terminal region of p160 (termed 160C) recognised only p160 (Figure 3.14B). Several unidentified endogenous proteins were also detected within HEK 293T cell extracts with the 160N antiserum (Figure 3.14A).

The ability of the 160N and 160C polyclonal antisera (unpurified) to detect the respective endogenous murine proteins was examined using nuclear and cytoplasmic extracts prepared from two murine cell lines; FDC-P1 and NIH3T3. The FDC-P1 cell line was chosen as both p160 and p67 were originally detected in radiolabelled nuclear extracts prepared from these cells (Favier and Gonda, 1994). In comparison, only p160 could be detected in NIH3T3 nuclear extracts (Favier and Gonda, 1994). The 160N antiserum detected two proteins of approximately 160 and 67 kDa within FDC-P1 nuclear extracts (Figure 3.15A). An approximately 160 kDa protein was also detected within NIH3T3 nuclear extracts (Figure 3.15A). The sizes and distribution of the proteins detected with the 160N antiserum are consistent with those of p160 and p67, as determined originally by binding analyses (Favier and Gonda, 1994). In addition, a protein of approximately 97 kDa was detected within FDC-P1 cytoplasmic extracts with the 160N antiserum (Figure 3.15A).

The 160C antiserum detected two proteins of approximately 160 and 110 kDa within FDC-P1 nuclear extracts (Figure 3.15B). A single protein of approximately 160 kDa was detected within NIH3T3 nuclear extracts (Figure 3.15B). The inability of the 160C antiserum to detect a 67 kDa protein within FDC-P1 nuclear extracts supports the notion that p67 corresponds to the amino-terminal region of p160. The detection of an ~160 kDa nuclear protein within both FDC-P1 and NIH3T3 nuclear extracts with the 160C antiserum correlates with the distribution seen previously with the 160N antiserum. The identities of the ~110 kDa protein detected within FDC-P1 nuclear extracts with the 160C antiserum,

**Figure 3.14 Detection of ectopically expressed p160FLAG and
p67*FLAG with antisera raised against murine p160.**

Nuclear and cytoplasmic extracts were prepared from HEK 293T cells transiently transfected with expression vectors encoding p160FLAG or p67*FLAG. Total cell extracts were prepared from untransfected HEK 293T cells and are included as a control. Cell extracts containing ectopically expressed p160FLAG or p67*FLAG were used to determine the reactivities of the 160N and 160C polyclonal antisera, raised against the amino- and carboxyl-terminal regions of murine p160, respectively. Proteins were detected by Western analysis/ECL using unpurified antisera at a dilution of 1/500. 'N' and 'C' denote nuclear and cytoplasmic extracts, respectively.

(A) The 160N antiserum detected both p160FLAG (160 kDa) and p67*FLAG (60 kDa) in the nuclear and cytoplasmic fractions, respectively. These proteins are identified by labels shown on the left and right. Note that the 160N antiserum cross-reacted with several endogenous (human) proteins, as shown by total extracts from untransfected cells.

(B) The protein blot shown in part (A) was stripped and reprobed with the 160C antiserum. The 160C antiserum detected p160FLAG (160 kDa), but not p67*FLAG. The p160FLAG protein is identified by a label shown on the left.
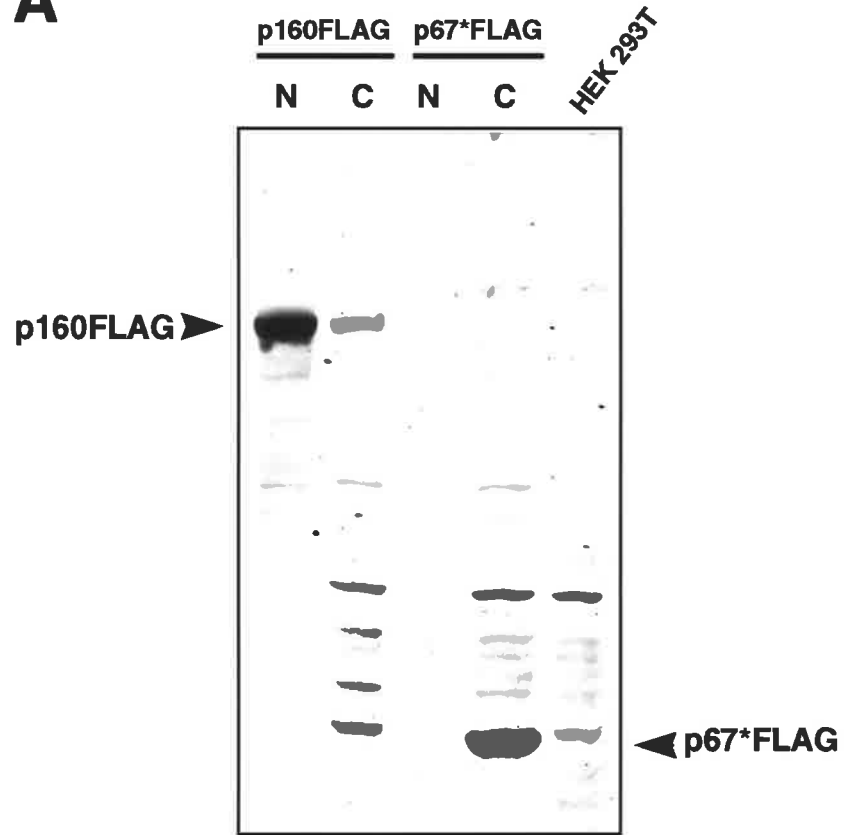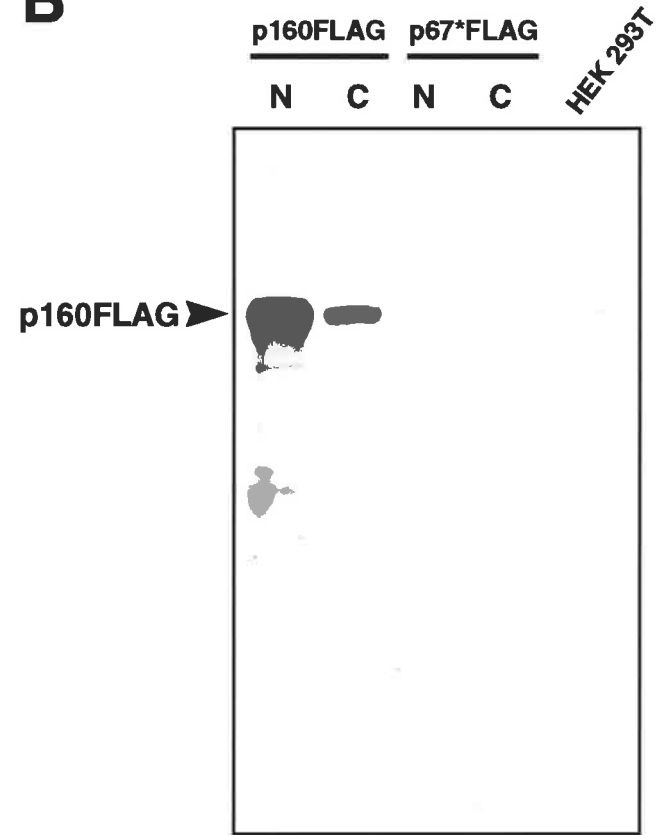
Figure 3.15 Detection of endogenous p160 and p67 with antisera raised against murine p160.

Nuclear and cytoplasmic extracts were prepared from FDC-P1 and NIH3T3 cells. Each lane contains 150 µg of extract, fractionated by SDS-PAGE. Proteins were detected by Western analysis/ECL using p160 polyclonal antisera raised against the amino-terminal region (160N) or the carboxyl-terminal region (160C). Unpurified 160N and 160C antisera were used at a dilution of 1/200 and 1/500, respectively. 'Nuc' and 'Cyt' denote nuclear and cytoplasmic extracts, respectively.

(A) Endogenous proteins detected with the 160N antiserum. Both endogenous p160 and p67 were detected within FDC-P1 nuclear extracts. Only p160 was detected within NIH3T3 nuclear extracts. An unidentified ~97 kDa protein was detected within FDC-P1 cytoplasmic extracts. Labels shown on the right identify p160 and p67. The positions of molecular weight markers (kDa) are shown on the left.

(B) Endogenous proteins detected with the 160C antiserum. Endogenous p160, but not p67, was detected within FDC-P1 nuclear extracts. An unidentified ~110 kDa was also detected within FDC-P1 nuclear extracts. Only p160 was detected within NIH3T3 nuclear extracts. The label shown on the right identifies p160. The positions of molecular weight markers (kDa) are shown on the left.
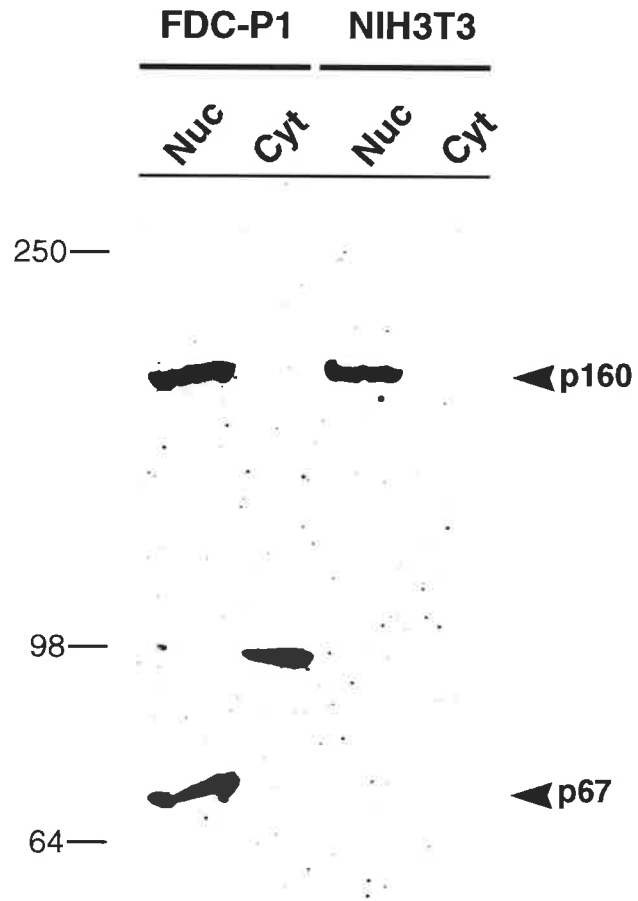
and the ~97 kDa protein detected within FDC-P1 cytoplasmic extracts with the 160N antiserum, remain unknown. However, it is possible that these proteins may be related to p160, as three murine chromosomal loci containing p160 sequences have been identified (see discussion and Tavner *et al.*, 1998; see appendix). Alternatively, these proteins may be non-specific reactivities.

To confirm the identity of the 160 and 67 kDa proteins detected with the p160 antisera, a binding assay was performed using nuclear extracts prepared from FDC-P1 cells, based on observations with radiolabelled endogenous p160 and p67 (Favier and Gonda, 1994). This binding analysis was performed with the GST fusion proteins containing either the wild type c-Myb leucine zipper (GST-NRD2) or a mutated form (GST-L3,4P) and GST-Jun, which contains the c-Jun bZIP region (Figure 3.9A). Bound and unbound fractions were analysed with both 160N and 160C antisera, which had been previously depleted of GST antibodies (see Chapter 2.2.3.11). The two endogenous proteins of ~160 and ~67 kDa, detected with the 160N antiserum, showed the same binding specificities as p160 and p67 (Figure 3.16A). That is, both p160 and p67 bound to fusion proteins containing the wild type c-Myb leucine zipper, whereas only p160 bound to a fusion protein containing the c-Jun bZIP region. Neither p160 or p67 bound to fusion proteins containing a mutated c-Myb leucine zipper. It is therefore confirmed that these proteins correspond to p160 and p67. Similarly, analysis of this binding assay with the 160C antiserum verifies the identity of the ~160 kDa protein as that of p160 (Figure 3.16B). Additionally, it was noticed that the 160C antiserum appears to detect p160 more efficiently than the 160N antiserum (Figure 3.16B). Interestingly, the unidentified ~110 kDa protein detected with the 160C antiserum showed the same binding specificities as p160, suggesting that this protein may represent either an amino-truncated form of p160 (as it is detected only with the 160C antiserum), or a highly-related protein (Figure 3.16B). Collectively, these data confirm that the 160N and 160C antisera specifically detected endogenous p160, and that only the 160N antiserum detected endogenous p67. Furthermore the specific reactivities of the 160N antiserum provide support for the notion that p67 corresponds to the amino-terminal region of p160. Unexpectedly, it was noticed that the 160N antiserum detected p160, but not p67 in FDC-P1 whole cell lysates, prepared by SDS lysis. This observation was further investigated and is discussed in more detail in Chapter 4.

Figure 3.16 The binding specificities of FDC-P1 nuclear proteins detected with antisera raised against murine p160.

Nuclear extracts were prepared from FDC-P1 cells and used in a binding assay with the GST fusion proteins; GST-NRD2 (M), GST-L,34P (M') and GST-Jun (J) (see Figure 3.9). Bound fractions are shown. Nuclear and cytoplasmic FDC-P1 extracts, and FDC-P1 whole cell lysates prepared by SDS lysis of cells, were included as controls. Proteins were detected by Western analysis/ECL using the 160N (1/500) or 160C (1/500) antisera depleted of anti-GST antibodies. 'N' and 'C' denote nuclear and cytoplasmic extracts, respectively.

(A) Detection of endogenous proteins with the 160N antiserum. The p160 and p67 proteins bound to GST-NRD2 (M), which contains the c-Myb leucine zipper motif. Only p160 bound to GST-Jun (J), which contains the c-Jun bZIP region. Neither protein bound to the GST-L3,4P, which contains a mutated c-Myb leucine zipper motif. Note that the p67 protein was not detected within the FDC-P1 whole cell lysate. This absence of p67 is denoted by an asterisk and an arrow. The positions of molecular weight markers (kDa) are shown on the left.

(B) Detection of endogenous proteins with the 160C antiserum. The protein blot shown in part (A) was stripped and reprobed with the 160C antiserum. The p67 protein was not detected with the 160C antiserum. Both the p160 protein and the unidentified ~110 kDa protein bound to GST-NRD2 and GST-Jun. Neither of these proteins bound to GST-L3,4P. The positions of molecular weight markers (kDa) are shown on the left.

A

FDC-P1
Nuclear

M   M'   J

250 —

p160 ▶

98 —

p67 ▶

64 —

BOUND

FDC-P1
whole cell

p160 ▶

* ▶

FDC-P1

N   C

◀ p160

◀ p67

B

FDC-P1
Nuclear

M   M'   J

250 —

p160 ◀

98 —

64 —

BOUND

FDC-P1
whole cell

p160 ▶

FDC-P1

N   C
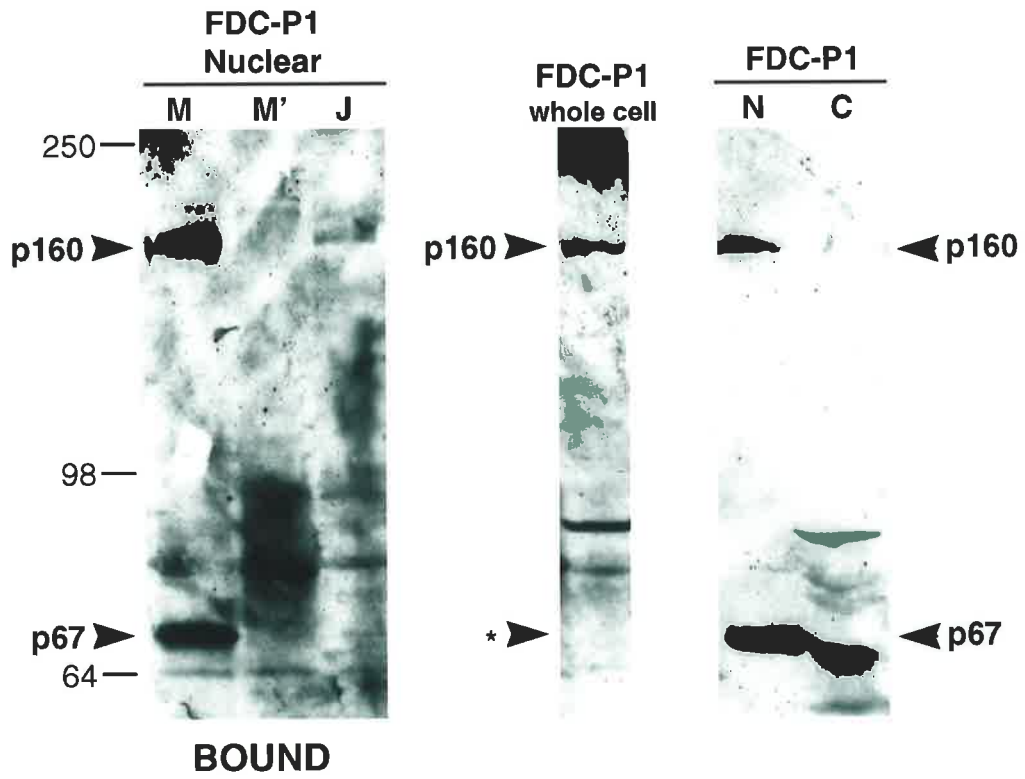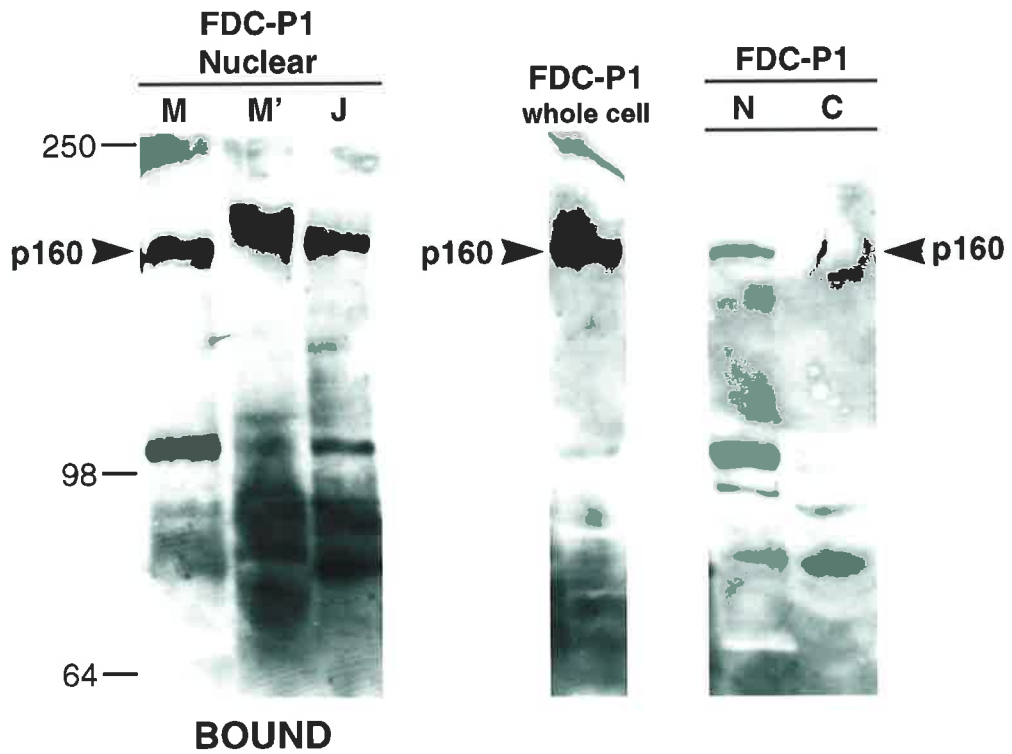
◀ p160

3.2.7 Analysis of p160 carboxyl-truncated proteins

The finding that the p160 carboxyl-truncated protein, p67*FLAG, was located within the cytoplasmic fraction of cell extracts, and not within the nuclear fraction, like endogenous p67, prompted a further investigation into this phenomenon. The p67* protein is comprised of the 580 amino-terminal residues of p160 (Figure 3.6). The presence of p160FLAG within HEK 293T nuclear extracts (Figure 3.11) and NIH3T3 nuclear extracts (data not shown) indicates that nuclear import is functional with respect to p160FLAG in these cells. Furthermore, the absence of p67*FLAG within nuclear fractions suggests that the nuclear localisation signal(s) of p160 is absent from p67*, and therefore is presumably located between residues 580 and 1344. The cytoplasmic localisation of p67* may be due to the fact that p67* (~60 kDa) lacks carboxyl-terminal residues present in endogenous p67, which may contain a nuclear localisation signal. Two approaches were undertaken to investigate the basis of the cytoplasmic localisation of p67*, and to gain insight into putative subcellular localisation signals that reside within p160. Firstly, the carboxyl terminus of 67*FLAG was extended by 59 and 121 residues to create two additional p160 carboxyl-truncated proteins; termed p73 and p80, respectively, that may potentially encompass the carboxyl terminus of endogenous p67 (Figure 3.17). The p73FLAG and p80FLAG constructs were generated by introducing termination codons into the p160FLAG cDNA sequence by PCR, to yield proteins with predicted molecular masses of 73 and 80 kDa, respectively (Chapter 2.1.12 (III)). Secondly, a heterologous nuclear localisation signal (NLS), the SV40 large tumour antigen (T-ag) NLS, was introduced into the carboxyl terminus of p67*FLAG in order to target this protein to the nucleus. The SV40 large T-ag NLS is well characterised and been used extensively to direct heterologous proteins into the nucleus (reviewed in Jans and Hubner, 1996). The SV40 large T-ag minimal NLS (^{126}PKKKRKV^{132}) was introduced into the carboxyl terminus of p67*FLAG by PCR (termed p67*FLAG NLS (m); see Chapter 2.1.12 (IV)) (Figure 3.17).

The subcellular localisation of p67*FLAG, p73FLAG, p80FLAG and the targeted p67*FLAG NLS (m) was assessed by examination of nuclear and cytoplasmic extracts from HEK 293T cells transiently expressing these proteins. Both p67*FLAG and p73FLAG were detected only within the cytoplasmic fraction, indicating that the additional 59 amino acids were apparently unable to confer translocation to the nucleus (Figure 3.18A). In contrast, a proportion of p80FLAG was present within the nuclear fraction as

Figure 3.17 Schematic representation of p160 carboxyl-truncated proteins, and fusion proteins containing the SV40 large T-antigen nuclear localisation signal.

Carboxyl-truncated proteins (represented by black bars) are shown with respect to p160. All proteins contain the FLAG epitope at the amino terminus. The position of the carboxyl-terminal residue of each truncated protein is shown. The SV40 large T-antigen (T-ag) minimal nuclear localisation signal (NLS) has been introduced into the carboxyl terminus of p67* (termed p67*FLAG NLS (m)), in addition to a dipeptide spacer (glycine-alanine). The CcN motif (Jans and Hubner, 1996), which contains the SV40 large T-ag NLS and regulatory phoshorylation sites, has been introduced into the carboxyl terminus of p67* (termed p67*FLAG NLS (e)), in addition to a single amino acid spacer (glycine). The inhibitory cdc2 phosphorylation site has been mutated by substitution of critical residues (denoted by arrows) with alanine.

L-L-L-I-M

L-L-Q-L-L

NH₂ — 1 — 2 3 — 4 5 — L — A — L — COOH

200   400   600   800   1000   1200

FLAG ——————————| p67*FLAG
                580

FLAG ———————————| p73FLAG
                 639

FLAG ————————————| p80FLAG
                  701

FLAG ——————————GA-¹²⁶PKKKRKV¹³²
                SV40 large T-ag
p67*FLAG NLS (m)   NLS

AA

FLAG ——————————G-¹¹⁰S SSDDE ATADSQH STP PKKKRKV¹³²
                    casein   cdc2
p67*FLAG NLS (e)   kinase II
                              A
                    SV40 large T-ag CcN motif
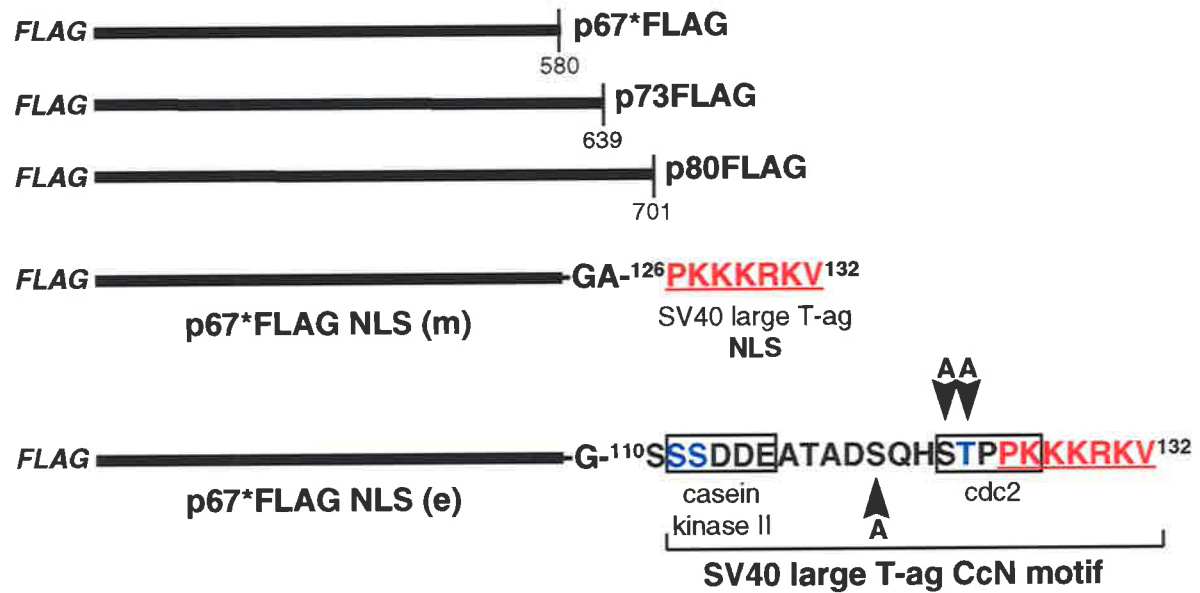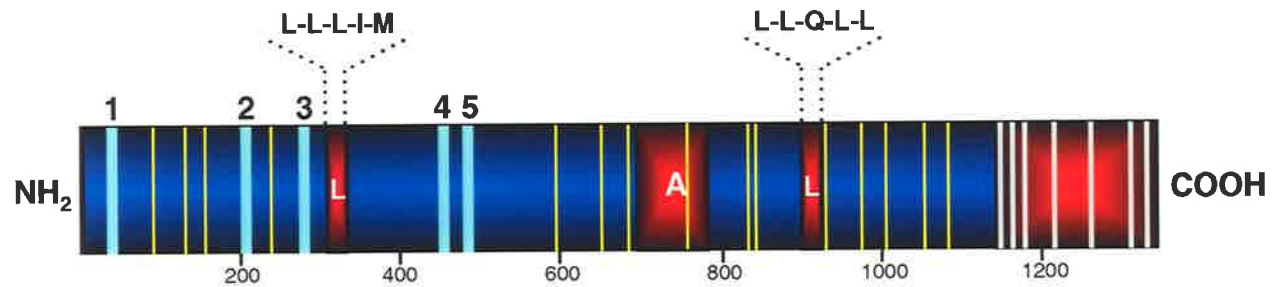
well as the cytoplasmic fraction, suggesting the presence of a signal between residues 640 and 701 that contributes to nuclear import (Figure 3.18A). Comparison of the apparent molecular weights of p73FLAG and p80FLAG with that of endogenous p67, detected with the 160N antiserum, indicates that the carboxyl terminus of endogenous p67 is likely to reside between p160 residues 580 (the carboxyl-terminal residue of p67*FLAG) and 639 (the carboxyl-terminal residue of p73FLAG) (Figure 3.18B). Furthermore, the p73FLAG and endogenous p67 proteins migrated to almost the same position in the SDS polyacrylamide gel, indicating that the carboxyl terminus of endogenous p67 lies within close proximity to p160 residue 639 (Figure 3.18B).

The p67*FLAG fusion protein containing the SV40 large T-ag minimal NLS could not be detected within the nuclear fraction, but rather, was present only within the cytoplasmic fraction. (Figure 3.18A). Subsequently, a modified nuclear-targeted p67*FLAG fusion protein (termed p67*FLAG NLS (e)) was constructed using the SV40 large T-ag 'CcN motif' which contains the NLS and adjacent casein kinase II and $p34^{cdc2}$ cyclin-dependent kinase regulatory phosphorylation sites that influence the kinetics of nuclear transport (Jans et al., 1991) (Figure 3.17). Phosphorylation of threonine 124 in the CcN motif by $p34^{cdc2}$ negatively regulates nuclear transport and therefore inhibits maximal nuclear accumulation (Jans et al., 1991). This inhibitory site had been mutated in the CcN motif introduced into the carboxyl terminus of p67*FLAG (Figure 3.17) (Chapter 2.1.12 (IV)). Mutation of the $p34^{cdc2}$ phosphorylation site, by substitution of critical serine/threonine residues with alanine, increases the rate of nuclear accumulation approximately fifty fold with respect to the minimal NLS (D. Jans; pers. comm.). The p67*FLAG NLS (e) fusion protein was constructed using a PCR approach (Chapter 2.1.12 (IV)). The subcellular localisation of p67*FLAG NLS (e) was examined in an identical manner to the previously discussed p160 carboxyl truncations (see above). Surprisingly, incorporation of the mutated CcN motif into p67*FLAG was apparently unable to direct this protein into the nucleus, as assessed by the presence of this protein only within the cytoplasmic fraction (Figure 3.18A). However, it was noticed that two higher molecular weight species were present, which may correspond to dimeric and tetrameric forms (Figure 3.18A). This phenomenon is attributed to the presence of the mutated CcN motif within p67*FLAG as no higher molecular weight species were seen with the p67*FLAG fusion protein containing the minimal NLS (Figure 3.18A). Interestingly a small proportion of these complexes were present within the nuclear

Figure 3.18 Subcellular localisation of p160 carboxyl-truncated proteins, and targeted fusion proteins.

Nuclear and cytoplasmic extracts were prepared from HEK 293T cells transiently transfected with expression vectors encoding epitope-tagged, full-length p160, p160 carboxyl-truncated proteins or targeted-p67* fusion proteins. Two separate isolates of the p67*FLAG NLS (e) clone were used for transfections. Proteins were detected by Western analysis/ECL. 'N' and 'C' denote nuclear and cytoplasmic extracts, respectively.

(A) Subcellular localisation of ectopically expressed proteins, detected with the anti-FLAG M2 monoclonal antibody (1/500). Note the poor expression of p67*FLAG (denoted by vertical arrow) with respect to other ectopically expressed proteins. The presence of p80FLAG in the lane containing nuclear extract prepared from cells expressing p73FLAG is probably due to overflow from the adjacent lane during loading of the gel. The asterisks denote the higher molecular weight species seen in the presence of p67*FLAG NLS (e) proteins. The positions of molecular weight markers (kDa) are shown on the left of the figure.

(B) Comparison of endogenous p67 with the p160 carboxyl-trucated proteins, p80FLAG and p73FLAG. The protein blot shown in part (A) was stripped and reprobed with the 160N antiserum (1/500) (raised against the amino-terminal region of p160). Note that the lanes containing the FDC-P1 nuclear and cytoplasmic extracts have been manipulated, such that they are adjacent to the lanes containing the carboxyl-truncated proteins, for ease of comparison. Labels on the left correspond to the position of endogenous p67 (and p160), while those on the right correspond to the positions of ectopically expressed proteins.
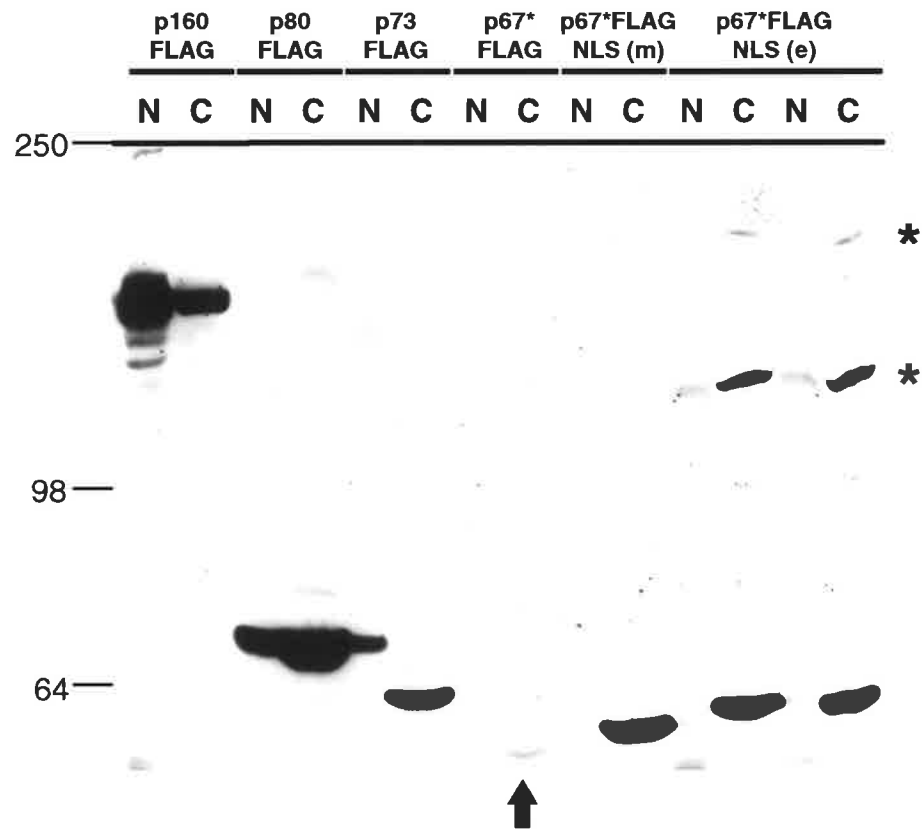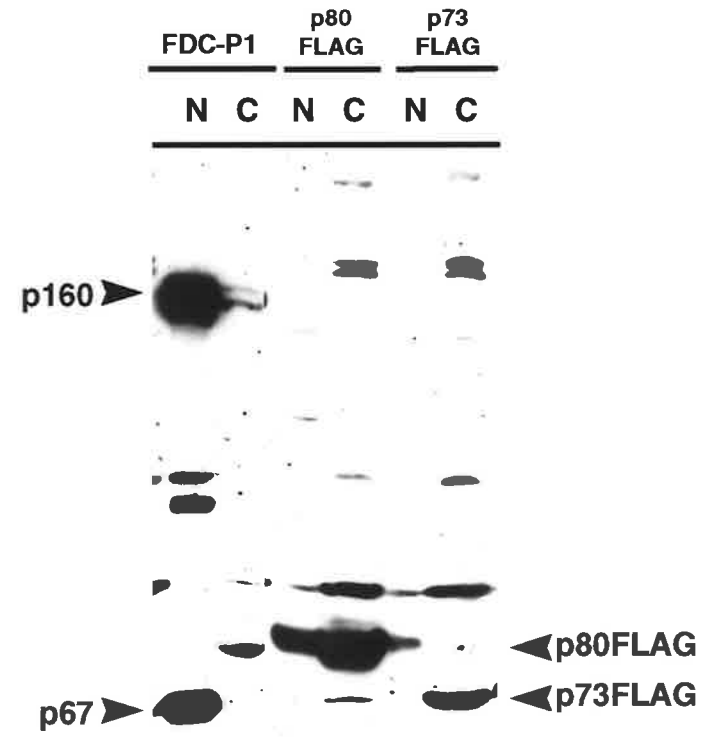
fraction (Figure 3.18A). The inability to detect p67* fusion proteins containing the SV40 large T-ag NLS or CcN motif within nuclear extracts suggests that there is a region(s) within p67* (p160 amino-terminal residues 1-580) that confers cytoplasmic localisation, and which acts dominantly with respect to the heterologous SV40 large T-ag NLS and CcN motif. This region(s) may constitute a nuclear export signal or cytoplasmic retention domain. Expression of the p160 carboxyl-truncated proteins and the targeted p67* fusion proteins has also been examined in NIH3T3 cells, and are consistent with the observations of expression in HEK 293T cells (data not shown).

The transiently expressed p160FLAG, p73FLAG, p80FLAG, p67*FLAG NLS (m) and p67*FLAG NLS (e) proteins were consistently more readily detected within cell extracts than p67*FLAG (Figure 3.18A). The differing ability to detect these proteins may be an indication of relative protein stability *in vivo*, and suggests that the addition of residues to the carboxyl terminus of p67*FLAG may result in increased protein stability with respect to p67*FLAG.

3.3   Discussion

3.3.1   The predicted polypeptide sequence of p160 provides clues to its function

The related nuclear proteins, p160 and p67, interact specifically with the c-Myb leucine zipper motif, a component of the c-Myb negative regulatory domain. A 4.1 kb murine cDNA clone has been obtained using sequence information derived from p67 (Tavner *et al.*, 1998; see appendix and section 3.2.1). Evidence presented in this chapter demonstrates that this clone corresponds to p160, and contains all five tryptic peptide sequences derived from p67. The relationship between p160 and p67 is believed to be one of precursor and product, whereby p67 is derived from the amino-terminal region of p160. Experimental evidence for this relationship is presented and discussed in Chapter 4.

The p160 cDNA clone encodes a polypeptide comprising 1344 amino acids, with a predicted molecular mass of 152 kDa. As p160 and p67 were both detected as a consequence of their interaction with the c-Myb leucine zipper, the p160 cDNA sequence was searched specifically for putative leucine zipper motifs. Two putative leucine zipper motifs have been identified in the p160 cDNA sequence and contain the following residues in the repeated *d* position (heptad repeat nomenclature): ^{307}L-L-L-I-M^{335} (LZ1) and ^{900}L-L-

Q-L-L^{928} (LZ2). These putative leucine zipper motifs partially conform to the typical features of leucine zippers, and in particular, contain a predominance of hydrophobic residues in the *a* and *d* positions. However, the noticeable lack of charged residues in the *e* and *g* positions suggests that p160 LZ1 and LZ2 may not act in the classical context of a leucine zipper. Each p160 putative leucine zipper motif may function as an amphipathic, α-helical hydrophobic interaction face, but may not necessarily interact specifically with other leucine zippers. Mutagenesis studies of each putative leucine zipper will be required to ascertain functionality (see section 3.3.3 for further discussion). The presence of a proline residue within LZ1 may indicate that this region is not entirely α-helical in structure.

The presence of other potential protein-protein interaction motifs in p160 is highlighted by the identification of putative leucine charged domains (LCDs), a recently described motif that mediates the interaction between nuclear hormone receptors and their transcriptional coactivators (Torchia *et al.*, 1997; Heery *et al.*, 1997). The LCDs of the p/CIP, NCoA-1 and CBP/p300 transcriptional coactivators are believed to provide the functional interaction specificity with distinct protein partners. Leucine charged domains share a leucine-rich core peptide motif (LXXLL), flanked by a small region of charged residues, and are predicted to form an amphipathic α-helical domain (Torchia *et al.*, 1997; Heery *et al.*, 1997). Five putative LCDs have been identified in the p160 sequence, in addition to a further ten that partially conform to the core peptide motif and contain an alternative aliphatic amino acid (valine or isoleucine) in one or more of the positions held by a leucine residue. Mutational analyses of selected LCDs of CBP, NCoA-1 and RIP-140 (a nuclear receptor-binding protein) indicate that substitution of residues in the core motif disrupt interactions with specific binding partners (Torchia *et al.*, 1997; Heery *et al.*, 1997). In particular, substitution of leucine with valine in one LCD of RIP-140 suggests that aliphatic residues other than leucine are tolerated in some positions of the core motif, but not others (Heery *et al.*, 1997). Furthermore, a requirement for specific hydrophobic residues may be necessary to constitute LCD interaction faces, as substitution of conserved leucine residues in positions 4 and 5 of a RIP-140 'LXXLL' motif, with valine, significantly disrupted interaction with the ligand-binding domain of the oestrogen receptor (Heery *et al.*, 1997). Further characterisation of the critical determinants that mediate LCD interaction specificity is awaited. It is therefore difficult to predict whether the putative

LCDs identified in the p160 sequence could conceivably function as protein-protein interaction domains, and if so, which ones.

The basic repeats located within the carboxyl-terminus of p160 show similarities to nuclear localisation signals (NLSs). A NLS consensus sequence has not yet emerged, however NLSs generally comprise short peptide sequences that predominantly contain hydrophilic residues (reviewed in Jans and Hubner, 1996; Garcia-Bustos *et al.*, 1991). Furthermore, the basic amino acids, lysine and arginine, have been shown to be critical determinants of NLSs (reviewed in Jans and Hubner, 1996; Garcia-Bustos *et al.*, 1991; Dingwall and Laskey, 1991). Basic residues that comprise a NLS are often found in single or multiple clusters (eg bipartite NLSs) (reviewed in Dingwall and Laskey, 1991). The potential functional role of the p160 basic repeats as multiple nuclear localisation signals is suggested by the cytoplasmic localisation of p160 carboxyl-truncated proteins lacking these repeats (see section 3.2.7). This role is further discussed in section 3.3.4. The p160 carboxyl terminus additionally contains a large abundance of potential phosphorylation sites that flank the basic repeats, which may be substrates for several kinases, namely $p34^{cdc2}$, protein kinase C and MAP kinase. If the basic repeats of p160 constitute multiple NLSs, it is tempting to speculate that at least some of these phosphorylation sites may regulate nuclear transport of p160, by analogy with several examples of NLSs that are regulated by phosphorylation (reviewed in Jans and Hubner, 1996).

The centrally-located acidic region of p160 bears a general resemblance to acidic transcriptional activation domains and contains four potential casein kinase II phosphorylation sites. Acidic activation domains are commonly found in numerous transcription factors and function by associating with components of the basal transcription machinery, either directly or indirectly (reviewed in Tjian and Maniatis, 1994). However, it is uncertain how acidic transactivation domains achieve the specificity for particular protein interactions. It is proposed that the conformation of the activation domain may contribute towards the determination of protein interaction specificity (Tjian and Maniatis, 1994).

Database searches with the p160 cDNA sequence identified the *Drosophila lethal (1) 1Bi* gene product as the highest scoring match, and also indicate that the p160 sequence is

novel. In particular, this sequence similarity encompasses the p160 acidic region and other tracts of the p160 cDNA sequence. Unfortunately, the function of the *Drosophila* lethal (1) 1Bi protein remains unknown. The *Drosophila melanogaster lethal (1) 1Bi* gene was initially identified as a genetic complementation group encoded by the *suppressor of sable su(s)* region (Voelker *et al.*, 1989). Although a molecular characterisation of *lethal (1) 1Bi* has not yet been reported, the protein encoded by this gene is apparently nucleolar (cited in GenBank; accession number U20542).

Examination of endogenous and epitope-tagged p160 in NIH3T3 fibroblasts by immunofluorescence studies revealed that p160 is predominantly associated with the nucleolus, with a proportion also detected within the nucleoplasm (Tavner *et al.*, 1998; see appendix). This finding was unexpected for a protein that interacts with c-Myb in a potential regulatory context. This reasoning is primarily based on the functions of c-Myb, which constitute transcriptional regulation of genes transcribed by RNA polymerase II. Furthermore, immunofluorescence studies demonstrate that c-Myb is not detected in the nucleolus (Bading *et al.*, 1989b). The 160 kDa PIP protein, identified as interacting with the PAR family of transcription factors, and which is believed to be the rat homologue of murine p160, is described as a nucleolar protein with similarity to the *Drosophila* lethal (1) 1Bi protein (cited in GenPept database; accession number U83590). A molecular characterisation of PIP has not yet been reported.

The nucleolus is a distinct sub-nuclear site where RNA polymerase I-mediated transcription of ribosomal genes (other than 5S rRNA genes) is active, and predominantly where ribosome subunit biogenesis occurs (reviewed in Shaw and Jordan, 1995). Ribosomal genes exist in multiple copies (100-5000/haploid genome) within eukaryotic genomes, and are genetically organised in tandem arrays that constitute nucleolar organising regions (reviewed in Shaw and Jordan, 1995; Sollner-Webb and Mougey, 1991). Nucleolar organising regions form the focal point of nucleolar activity and require the functions of a plethora of proteins that contribute to ribosomal gene transcription and ribosome biogenesis (reviewed in Shaw and Jordan, 1995).

It is apparent that p160 shows similarities to other nucleolar proteins. Upstream binding factor (UBF), an RNA polymerase I (nucleolar) transcription factor, was found to be a

relatively high scoring match with p160. This similarity encompasses the acidic regions of both proteins (data not shown). Although database searches revealed that the p160 acidic region displays similarity to numerous proteins, the similarity between the p160 and UBF acidic regions may be of particular significance in light of the fact that these proteins are both nucleolar (Jantzen *et al.*, 1990; Tavner *et al.*, 1998; see appendix). The highly acidic region of UBF lies within the carboxyl terminus and comprises a substrate for casein kinase II phosphorylation (Voit *et al.*, 1995; Voit *et al.*, 1992; Jantzen *et al.*, 1990). Furthermore the acidic region of UBF is required for transactivation activity and can negate histone H1-mediated transcriptional repression by displacing histone H1 from the nucleosome (Kermekchiev *et al.*, 1997; Kuhn *et al.*, 1994; Voit *et al.*, 1992; Kuhn and Grummt, 1992). Phosphorylation of the UBF acidic region by casein kinase II contributes to transactivation activity (Voit *et al.*, 1995). The potential functional role of the p160 acidic region is yet to be investigated, however, this region may function as a transcriptional activation domain, possibly in the context of both RNA polymerase I and II-mediated transcription, by interacting with distinct proteins.

More generally, p160 has features in common with Nopp140, a 140 kDa phosphoprotein that is predominantly located within the nucleolus, but which is also detected within the nucleoplasm (Meier and Blobel, 1992). These similarities are not significant with respect to actual sequence information. The Nopp140 protein was originally identified by its interaction with the SV40 large T-antigen NLS and has been shown to shuttle between the nucleolus and cytoplasm along curvilinear tracks between the nucleolus and nuclear pore complexes (Meier and Blobel, 1992). Similar to p160, the Nopp140 amino acid sequence contains an acidic region, and seven basic (lysine and arginine) repeats primarily located within the carboxyl-terminus (Meier and Blobel, 1992). The acidic region of Nopp140 is comprised of ten acidic repeats which contain clusters of serine residues, and consensus sites for casein kinase II. Similar acidic regions are found in several other nucleolar proteins that also interact with the SV40 large T-antigen NLS *in vitro* (reviewed in Xue and Melese, 1994). The functional significance of these interactions are not understood, however it has been proposed that these nucleolar proteins may play a role in ribosome assembly, rather than transport of proteins into the nucleus (Xue and Melese, 1994). The Nopp140 protein has recently been characterised as a transcriptional coactivator of alpha-1 acid glycoprotein (*agp*) gene expression (Miau *et al.*, 1997). Interestingly, another

nucleolar protein, nucleolin, has also been shown to transcriptionally regulate *agp* gene expression, but unlike Nopp140, nucleolin acts as a repressor of *agp* expression (Yang *et al.*, 1994). Activation of *agp* gene expression is achieved by synergistic interaction of Nopp140 with the C/EBPβ bZIP transcription factor and TFIIB. Unlike C/EBPβ, Nopp140 was unable to directly bind to the *agp* promoter region. This report demonstrates that Nopp140, a predominantly nucleolar protein, can interact with transcription factors that are associated with RNA polymerase II-mediated gene expression, and may possibly explain the significance of the proportion of Nopp140 located within the nucleoplasm. It is therefore conceivable that the proportion of p160 detected within the nucleoplasm could functionally interact with c-Myb, and possibly additional RNA polymerase II transcription factors (ie c-Jun).

The features of p160 (eg LCDs) and similarities with UBF and Nopp140, suggest that p160 may function as a transcriptional coactivator, possibly of both RNA polymerase I and II-mediated transcription. This suggested function may seem contradictory with respect to the association of p160 with the c-Myb negative regulatory domain. If p160 does indeed function, in part, as a coactivator of c-Myb transactivation, then this function may not be entirely essential for the transcriptional activities of c-Myb, as disruption or deletion of the leucine zipper motif enhances the transactivation and transformation abilities of c-Myb (Kanei-Ishii *et al.*, 1992; Hu *et al.*, 1991; Sakura *et al.*, 1989). Therefore, p160 may be required for transcriptional activation of certain Myb target genes, rather than all. For example, p160 may function as a coactivator of Myb target genes whose products contribute to a relatively mature cell phenotype, rather than those associated with cell proliferation and/or an immature cell phenotype. In this scenario, disruption or deletion of the c-Myb leucine zipper motif would favour transcriptional activation of target genes associated with cell proliferation and/or an immature phenotype. Transcriptional activation of particular Myb target genes may require an association of p160 with the CBP and/or p100 transcriptional coactivators, known to interact with c-Myb, in addition to other DNA-binding transcription factors that act cooperatively (eg C/EBPβ and Ets) (Oelgeschlager *et al.*, 1996; Dash *et al.*, 1996; Dai *et al.*, 1996). Transcription factors that cooperate with Myb may also potentially interact with p160 to form a promoter-specific activator complex.

3.3.2 The expression patterns of p160 and p67

Initial examination of murine and human cell lines by binding analyses with GST fusion proteins containing the c-Myb leucine zipper motif, indicated that p160 is expressed ubiquitously among murine cell lines, whereas p67 was only detected in a subset of murine immature myeloid cell lines (Favier and Gonda, 1994). However, neither p160 or p67 were detected in any of the human cell lines examined (Favier and Gonda, 1994). Examination of several cell lines by Northern analysis with the p160 cDNA clone indicated that a single ~4.5 kb transcript was detected in all murine cell lines examined (Figure 3.7), and is consistent with the ubiquitous distribution of p160. No additional transcript was detected in murine cell lines containing both p160 and p67, suggesting that the p160 transcript encodes both p160 and p67. A single p160 mRNA transcript of 4.5 kb has been detected in all murine tissues and organs examined (Tavner *et al.*, 1998; see appendix). The murine p160 cDNA sequence failed to detect a human HL-60 transcript. This is somewhat surprising on the basis that the gene encoding the human p160 homologue, which has been isolated using murine p160 sequence information, shows approximately 74.4% sequence (nucleotide) identity with that of murine p160, including untranslated regions (R. Keough; unpublished data). Moreover, a p160 transcript has been detected by RT-PCR (reverse transcriptase-PCR) in all human cell lines examined to date (R. Keough; unpublished data).

Antisera were raised against the amino and carboxyl termini of murine p160, and were shown to be capable of detecting endogenous p160 within nuclear extracts of the murine cell lines examined (ie FDC-P1 and NIH3T3). Furthermore, endogenous p67 was detected with antiserum raised against the amino-terminal region of p160, but not with antiserum raised against the p160 carboxyl-terminal region, thus providing support for the notion that p67 corresponds to the amino-terminal region of p160. Moreover, the distribution of endogenous p160 and p67, detected with the p160 antisera, amongst the FDC-P1 and NIH3T3 murine cell lines is consistent with that previously determined by binding analyses (Favier and Gonda, 1994). The ubiquitous expression of p160 suggests that this protein may function in a diverse spectrum of cell types, which is consistent with a potential nucleolar role.

It remains unresolved as to whether the antisera raised against murine p160 can cross-react with human p160 (or p67). Nuclear and cytoplasmic extracts from HEK 293T cells have

been examined for the presence of p160. A nuclear protein slightly smaller than murine p160 was detected with antiserum raised against the carboxyl-terminal region of p160 (data not shown). Verification of this protein as p160 awaits examination with antiserum raised against the human p160 protein.

Two unidentified murine proteins were detected with the p160 antisera. These proteins may be related to p160 as three murine chromosomal loci contain sequences that hybridize with p160 cDNA (Tavner *et al.*, 1998; see appendix). The *P160* locus resides on chromosome 11. Additionally, chromosomes 1 and 12 contain p160-related loci, designated *P160-rs1* and *P160-rs2*. Alternatively, the *P160-rs* loci may not encode p160-related proteins, and may actually be pseudogenes. The latter situation is probably more likely as no additional transcripts have been detected with the p160 cDNA, unless *P160-rs* transcripts are of the same size as the *P160* transcript.

3.3.3   The binding properties of p160 and p67, and their functional significance

Like the endogenous p160 and p67 proteins, cloned p160, and a carboxyl-truncated form of p160 (p67*) used to approximate the properties of endogenous p67, were able to interact with the c-Myb leucine zipper motif, but not to mutated forms. Binding assays using *in vitro* translated p160 and p67* indicate that the interaction with c-Myb is direct. The p160 protein is additionally able to interact with the human c-Jun bZIP region. Conversely, p67 is unable to bind to this region of c-Jun. The ability of p67 to interact with the c-Myb leucine zipper motif, but not the c-Jun bZIP region indicates that the 580 amino-terminal residues of p160 encompass the Myb-binding region. The region carboxyl-terminal to p160 residue 580 must therefore encompass the Jun-binding region. Interestingly, mutation of each p160 leucine zipper motif (LZ1 and LZ2) by substitution of certain leucine residues with proline did not appear to alter the ability of *in vitro* translated p160 to bind to GST fusion proteins containing either the c-Myb leucine zipper or the c-Jun bZIP region (R. Keough, unpublished data). Mapping of the respective binding domains within p160 is currently being examined by deletion studies (J. Lutwyche, R. Keough, T. Gonda; pers. comm.). Mutational studies of the c-Jun bZIP region suggest that p160 may not interact with the leucine zipper motif *per se* (Gardner, 1996). Further binding analyses using proteins expressed within mammalian cells are required to resolve this issue.

Although an *in vivo* interaction between p160 and full-length c-Myb could not be demonstrated during the course of work described in this thesis, due to technical difficulties, this interaction has recently been shown by coimmunoprecipitation of ectopically coexpressed proteins (Tavner *et al.*, 1998; see appendix). As expected, by analogy with binding assays using GST fusion proteins containing the c-Myb leucine zipper motif, p160 interacts with full-length, wild type c-Myb, but not with full-length c-Myb containing a mutated leucine zipper motif (L3,4PMyb). Similarly, an *in vivo* interaction between p67* and full-length c-Myb has also recently been demonstrated (J. Lutwyche; unpublished data). Ultimately, a demonstration of the interaction between endogenous p160, p67 and c-Myb would be desirable. To this end, immunoprecipitations from FDC-P1 cell extracts were attempted using the p160 antisera, however the resulting immunoprecipitates contained an excessive number of 'background' proteins (data not shown). It was therefore evident that affinity purified p160 antisera would be required for immunoprecipitation of endogenous proteins.

The c-Myb negative regulatory domain (NRD) down-regulates DNA-binding, transactivation and transformation ability of this protein. The association of p160 and p67 with the NRD via an interaction with the leucine zipper motif, initially suggested that p160 and p67 may function as repressors of c-Myb activity, since disruption of the c-Myb leucine zipper motif enhances transactivation and transformation ability (Kanei-Ishii *et al.*, 1992). Gel mobility shift assays using recombinant Myb proteins and *in vitro* translated p160 and p67* were attempted to investigate the effect of p160 and p67* on the ability of Myb to bind to oligonucleotides containing a Myb-binding site (data not shown). It was envisaged that p160 and/or p67* may inhibit c-Myb activity by disrupting its ability to interact with DNA. It became apparent that purified p160 and p67* proteins would be required for such analyses, due to technical difficulties associated with *in vitro* translated proteins (eg co-translation of p160 or p67* with Myb) (data not shown). Consequently further assays were not attempted.

To examine whether p160 could alter the ability of c-Myb to transcriptionally activate gene expression, transactivation analyses were performed with ectopically expressed wild type or leucine zipper-mutated (L3,4P) Myb and epitope-tagged wild type p160 or carboxyl-truncated p160 (p67*), and a reporter construct containing the human c-*myc* promoter

(Tavner *et al.*, 1998; see appendix). The c-*myc* promoter contains several Myb-binding sites that are required for transcriptional activation by c-Myb (Nakagoshi *et al.*, 1992). The presence of p160 had a minor effect on the ability of c-Myb to transactivate the c-*myc* promoter. A small stimulatory effect was seen in the presence of a relatively small amount (1 μg) of p160 effector plasmid, whereas a modest inhibition of c-Myb transactivation (~25%) was observed in the presence of a larger amount (5 μg) of p160 effector plasmid. In contrast, increasing amounts of the p67* effector plasmid (3-5 μg) significantly inhibited c-Myb transactivation (~88%) of the c-*myc* promoter, and suggests that p67* is acting as a dominant negative form of p160 (Tavner *et al.*, 1998; see appendix). Unexpectedly, inhibition (~50%) was seen in the presence of leucine zipper-mutated Myb (L3,4PMyb) and p67*, indicating that p67* may additionally interact with another region of c-Myb and/or other transcription factors that potentially cooperate with c-Myb to transactivate the c-*myc* promoter.

Taken together, these data support the notion that p160 is a coactivator of c-Myb-dependent transactivation. The inability of ectopically expressed p160 to significantly affect c-Myb transactivation of the c-*myc* promoter may be explained by one or more of the following points: (1) The endogenous level of p160 may be sufficient for coactivation of c-Myb-dependent transcription, and thus additional p160 would probably have a minimal effect. (2) Transcriptional activation of the c-*myc* promoter by c-Myb may require more than one coactivator. This additional coactivator may be rate-limiting and therefore the endogenous amount of this factor may be insufficient with respect to the amount of ectopically coexpressed p160 and c-Myb. A good candidate for such an additional coactivator is CBP, which is rate-limiting within cells and functions in the context of c-Myb-dependent transcription (Oelgeschlager *et al.*, 1996; Dai *et al.*, 1996; Petrij *et al.*, 1995). Binding studies with p160 and p67* indicate that p160 can directly interact with CBP (S. Ishii; pers. comm.). The interaction of known coactivators with CBP is highlighted by the nuclear receptor coactivator (NCoA) family members; NCoA-1/SRC-1, NCoA-2 and p/CIP (Torchia *et al.*, 1997; Smith *et al.*, 1996; Kamei *et al.*, 1996). The interaction of p160 with CBP and their association with c-Myb, and possible additional cooperating factors, may constitute a promoter-specific activator complex. (3) Finally, c-Myb-dependent transactivation of the c-*myc* promoter may not require the function of p160, in which case, promoters of other known Myb target genes need to be investigated.

The small level of inhibition (~25%) of c-Myb transactivation observed in the presence of a larger amount of p160 effector plasmid may suggest that p160 can homodimerize. Homodimerization may be mediated by one of the putative leucine zipper motifs that have been identified in the predicted polypeptide sequence of p160, however, this suggestion is purely speculative. Homodimerization may disrupt the interaction of p160 with c-Myb or impart allosteric changes that interfere with the interactions of the protein complex associated with the promoter. This could initially be tested, both *in vitro* and *in vivo*, by performing the following experiments: (1) Gel mobility shift assays with c-Myb, oligonucleotides containing a Myb-binding site (wild type and mutated), and several varying amounts (titration) of p160 protein. (2) Transactivation assays similar to those previously described (Tavner *et al.*, 1998; see appendix), but using a broader range of the amount of p160 effector plasmid transfected into cells. Additionally, it may be necessary to perform such transactivation analyses in cells that lack, or contain a low level of p160 (for reasons discussed above). If these experiments are suggestive of p160 homodimerization, then the region(s) that potentially mediate this interaction will need to be identified and verified (eg by mutation), and the experiments repeated with mutated and wild type p160.

The mechanism of p67*-mediated inhibition of c-Myb transactivation is currently unknown. The cytoplasmic localisation of ectopically expressed p67* may account for the observed inhibition by sequestering c-Myb within the cytoplasm and thus inhibiting the translocation of c-Myb to the nucleus. Alternatively, a small, undetectable amount of p67* may enter the nucleus and exert its effects on the activity of c-Myb. It is probable that p67* lacks functional domains required for p160 coactivation. Therefore, although p67* can bind to c-Myb, it may be unable to interact with other necessary proteins of the putative activator complex on the c-*myc* promoter, for example CBP, and thus act as a dominant negative inhibitor.

3.3.4 Truncation of p160 reveals properties that affect subcellular localisation

Nuclear localisation signals confer translocation into the nucleus by associating with specific shuttling transport proteins, namely importin-α (NLS 'receptor'), which interacts directly with NLSs, and importin-β, which cooperates with importin-α to constitute a stable NLS protein-importin carrier complex and moreover, mediates recognition of the nuclear pore complex (reviewed in Nigg, 1997).

The inability of p67*, and other p160 carboxyl-truncated proteins, to be detected within nuclear extracts, and the cytoplasmic localisation of p67* fusion proteins containing the SV40 large T antigen (T-ag) NLS or CcN motif, indicates that carboxyl truncation has unmasked regions of p160 that influence subcellular localisation. Investigation of the subcellular localisation of carboxyl-truncated forms of p160 described in this thesis and elsewhere (Gardner, 1996; J. Lutwyche, R. Keough; pers. comm.) indicate that the carboxyl terminus of p160 contains peptide sequences that direct this protein into the nucleus. It is hypothesised that the basic repeats located within the carboxyl-terminal region of p160 (see Figure 3.6), which bear a resemblance to known NLSs (discussed previously in section 3.3.1), constitute multiple nuclear localisation signals that act cooperatively to transport p160 into the nucleus. A series of carboxyl-truncated mutants of p160 that have increasing numbers of these basic repeats deleted, demonstrate a progressive cytoplasmic localisation (Gardner, 1996). However, these results await verification. Several proteins are known to contain multiple NLSs that are cumulatively required for nuclear localisation (reviewed in Jans and Hubner, 1996).

Interestingly, a proportion of the carboxyl-truncated protein p80, which lacks all of the basic repeats, was found in the nuclear fraction of cell extracts. A similarly truncated protein (p81; residues 1-739) also demonstrates the same subcellular distribution as p80 (Gardner, 1996). It is concluded that p160 residues 640-739 constitute a region that either contains a NLS that is distinct from the carboxyl-terminal basic repeats, or that this region interacts with a nuclear/nucleolar protein that is able to cotransport p80/p81 into the nucleus. No obvious NLS can be found in the predicted peptide sequence between residues 640 and 739. In order to demonstrate whether this region contains a NLS, it will be necessary to incorporate this region into a heterologous cytoplasmic protein and examine its subcellular loclisation. The alternative proposed function of this region is equally valid. Proteins lacking NLSs can be cotransported ('piggy-backed') into the nucleus by associating with NLS-containing proteins. For example, the cyclin-dependent kinase p34^{cdc2} does not contain a NLS, but enters the nucleus via an association with the NLS-containing cyclin B protein (Ookata et al., 1992). Residues 640-739 impinge on the acidic region of p160 and may therefore mediate interaction with a nuclear/nucleolar transcription factor(s).

Some nucleolar proteins associate with the nucleolus via short peptide sequences termed nucleolar 'localisation signals' (NOSs) that primarily consist of basic residues, and which may in fact be related to nuclear transport (reviewed in Jans and Hubner, 1996). Other known NOSs are structurally diverse and distinct (reviewed in Jans and Hubner, 1996). Identification of the regions in p160 that determine association with the nucleolus may provide insight into the potential nucleolar roles of this protein, as it is believed that the predominant mechanism which mediates assocoiation with the nucleolus, a sub-nuclear structure that is not membrane-bound, is one of interaction with other nucleolar proteins (reviewed in Jans and Hubner, 1996). It will be interesting to examine the sub-nuclear localisation of truncated and internally deleted forms of p160 by immunofluorescence microscopy, in order to map the region(s) that determine association with the nucleolus. The basic repeats of p160 may not only constitute multiple NLSs, but may also direct association with the nucleolus.

The SV40 large T-ag NLS and CcN motif were apparently unable to confer translocation of p67* to the nucleus, indicating that nuclear transport may be hindered by the context of the NLS within the protein or that a competing subcellular localisation signal(s) has been unmasked by carboxyl-truncation. Protein context is known to affect the function of NLSs (eg Roberts *et al.*, 1987). In particular, regions within the NLS-carrying protein that are largely composed of hydrophobic residues can impair nuclear transport, presumably by altering the accessibility of proteins, such as importin-α/-β, to the NLS. Introduction of the SV40 large T-ag NLS or CcN motif into another location within p67*, for example the amino terminus, may be required to assess whether protein context accounts for the observed inability of this signal to confer nuclear transport when located within the carboxyl terminus of p67*.

Alternatively, the SV40 large T-ag NLS and CcN motif may be functional in the p67* fusion proteins, but are competing with a dominant nuclear export signal or cytoplasmic retention domain. The existence of a nuclear export signal within p160 is favoured on the basis that endogenous p160 has not yet been detected within the cytoplasmic fraction of cell extracts. Nuclear export is a fundamental feature of eukaryotic cell biology because of the partitioning of nuclear activities (ie DNA replication, gene expression and ribosome biogenesis) from those within the cytoplasm (ie protein synthesis). Integration of these

activities requires regulated trafficking of molecules across the nuclear envelope. Macromolecules that are exported from the nucleus therefore include protein-bound RNA complexes and regulatory proteins not directly associated with RNA (eg protein kinases and transcription factors). In addition, several proteins are known to shuttle continuously between the nucleus/nucleolus and cytoplasm. Shuttling proteins include nucleolar proteins (eg Nopp140 and nucleolin), RNA-binding proteins (eg components of ribonucleoprotein (RNP) particles and the HIV-1 Rev protein) and transport receptors themselves (eg importin-α and exportin 1) (Borer *et al.*, 1989; reviewed in Ullman *et al.*, 1997; Jans and Hubner, 1996; Shaw and Jordan, 1995; Gerace, 1995). The export of macromolecules from the nucleus is not as well understood as the mechanisms that govern nuclear import. However, recent studies reveal that the mechanisms of nuclear export bear similarities to those of nuclear import. In particular, these similarities include inherent peptide signals of transported proteins, transport receptors/carriers (exportins) and energy dependence (reviewed in Ullman *et al.*, 1997; Nigg, 1997).

Very few nuclear export signals (NESs) have been identified and characterised. Known NESs, of PKI; an inhibitor of cAMP-dependent protein kinase (PKA), the HIV-1 Rev protein, TFIIIA (an RNA polymerase III transcription factor), and Ran-binding protein (RanBP1), constitute short, predominantly hydrophobic peptide sequences that are leucine-rich (Richards *et al.*, 1996; Fridell *et al.*, 1996; Wen *et al.*, 1995; Fischer *et al.*, 1995). Examination of the p67* peptide sequence (p160 residues 1-580) for putative NESs indicates a small region between residues 543 and 551 (^{543}LVQLADMLL^{551}) that is predominantly hydrophobic and leucine-rich. In order to determine whether this region does indeed constitute a NES, the subcellular localisation of proteins (ie p67* and p160) lacking this region, and heterologous proteins containing this region in addition to a NLS, need to be examined.

The endogenous p160 protein has not yet been detected in the cytoplasmic fraction of cell extracts. This may be an indication that the p160 NLS, which is believed to be composed of multiple carboxyl-terminal basic repeats, is a dominant competing signal with respect to the putative NES(s), in the context of full-length p160. Furthermore, this proposition may explain the apparent inability of the SV40 large T-ag NLS and CcN motif to direct p67* into the nucleus. The targeted p67* fusion protein may in fact have entered the nucleus by a

functioning signal, but was rapidly exported by the putative NES(s), which must be acting in a dominant manner with respect to the SV40 large T-ag NLS and CcN motif. The small proportion of targeted p67* fusion protein that may have existed within the nucleus was not detected within nuclear extracts by Western analysis.

The putative presence of both nuclear import and export signals raises the possibility that p160 might be a shuttling protein (cf Nopp140). The potential shuttling aspect of p160 may be related to a nucleolar function, for example ribosome biogenesis, which requires bidirectional nuclear transport. Alternatively (or additionally), nucleocytoplasmic shuttling of p160 may potentially be required to regulate its putative transcriptional activity. The 'strength' of the putative NES(s) within p160 may explain why p160 contains a putative NLS that is composed of multiple basic repeats. The cumulative effect of the basic repeats within p160 may be required to overcome the competing putative NES(s). It will be interesting to determine the outcome of these potentially competing signals in the context of p67*, by introducing varying numbers of the basic repeats of p160 into the carboxyl terminus of p67* (by internal deletion of p160). This approach may also reveal whether additional regions of p160 contribute to nuclear import, as highlighted by the subcellular distribution of the p160 carboxyl-truncated proteins p80 and p81.

CHAPTER FOUR

Determination of the Relationship Between p160 and p67

4.1 Introduction

The p160 and p67 proteins are related both functionally and structurally, but have distinct properties, namely their expression patterns among murine cell lines, and the ability to interact with the c-Jun bZIP region (Favier and Gonda, 1994). A functional relatedness was evident by the ability of both p160 and p67 to interact with the c-Myb leucine zipper motif. A structural relatedness was initially indicated by comparative peptide mapping, and more recently, when a cDNA clone corresponding to p160 was obtained using sequence information derived from p67 (Favier and Gonda, 1994; Tavner *et al.*, 1998; see appendix, and also Chapter 3). Sequence analysis of the full-length cDNA clone encoding p160 revealed that the five tryptic peptide sequences derived from p67 are found clustered in the amino-terminal region of p160 (see Figure 3.6).

There are at least three possible explanations that may account for the structural relatedness of p160 and p67; (1) p160 and p67 may be separately encoded by two highly related genes. (2) p67 may be derived from an alternatively spliced transcript of the p160 gene. (3) p160 and p67 may have a precursor-product relationship, whereby p67 is derived from p160 by proteolytic cleavage.

As discussed previously in Chapter 3, a potential precursor-product relationship for p160 and p67 was investigated on the basis of the following observations; (1) p160 and p67 share many peptides, as indicated by partial digestion of these proteins with V8 protease (Favier and Gonda, 1994). (2) The amimo-terminal region of p160 contains all five of the tryptic peptide sequences derived from purified p67 (Tavner *et al.*, 1998; see appendix). (3) A single mRNA species of approximately 4.5 kb was detected with p160 cDNA in all murine cell lines examined, including those in which both p160 and p67 proteins were detected (Tavner *et al.*, 1998; see appendix). If p160 and p67 are encoded by separate genes or alternatively spliced transcripts, it was expected that an additional mRNA species may be detected with the p160 cDNA (unless the additional species is of the same size as that

91

already detected). Later studies showed that antiserum raised against the amino-terminal region of p160 detected endogenous p160, as well as p67 (see Figure 3.15). Conversely, antiserum raised against the carboxyl-terminal region of p160 was unable to detect endogenous p67 (see Figure 3.15). This chapter describes an investigation of the precursor-product relationship of p160 and p67, and the finding that the distribution of p67 among murine cell lines correlates with the presence of a cell type-specific proteolytic activity that can cleave p160.

4.2 Results

4.2.1 Pulse-chase analysis of p160 in FDC-P1 cells

In order to assess whether p67 is generated from p160 *in vivo*, a pulse-chase analysis of p160 was undertaken using FDC-P1 cells, in which both p160 and p67 were detected. Cells were metabolically labelled with [^{35}S]-methionine for 1hr (the 'pulse'), after which a 1000-fold molar excess of non-radiolabelled methionine was added (the 'chase'). Nuclear extracts were prepared from cells at various time points during the chase (0 to 4 hours). Endogenous p160 and p67 proteins were affinity purified from radiolabelled nuclear extracts using the GST-NRD2 fusion protein, which contains the c-Myb leucine zipper motif, and analysed by SDS-PAGE. If the relationship between p160 and p67 is one of precursor and product, then it was expected that the pulse-chase analysis would ideally show diminishment of the amount of radioactive label in p160 and subsequent accumulation of label in p67 during the time course. The outcome of such a pulse-chase experiment in FDC-P1 cells is shown in Figure 4.1A, and a quantitative analysis of p160 and p67 at each time point is shown in Figure 4.1B. The relative amounts of p160 and p67 at each time point were quantitated with respect to the total amount of labelled protein extracted from cells at each specified time point. The analysis demonstrates that after 1 hour of metabolic labelling, p67 was present (0 hour time point), indicating that putative proteolytic cleavage of p160 may be a rapid process (Figure 4.1A). Up to thirty minutes after the start of the chase, incorporation of radiolabelled methionine still appeared to be occurring (Figure 4.1A,B). The amount of p160 rapidly diminished after the 0.5 hour time point (Figure 4.1A). Moreover, an accumulation of p67 was seen between the start of the chase (0 hour) and the 1 hour time point, after which the presence of p67 gradually diminished (Figure 4.1A,B). Taken together, these data are suggestive of a precursor-product relationship for p160 and p67, because the amount of radioactive label in p160

Figure 4.1 Pulse-chase analysis of p160 in FDC-P1 cells.

(A) FDC-P1 cells were metabolically labelled with [^{35}S]-methionine for 1 hour (the 'pulse'), after which a 1000-fold molar excess of methionine was added (the 'chase'), and cells incubated for up to 4 hours. Nuclear extracts were prepared from cells at the specified time points. The '0 hour' time point corresponds to the start of the 'chase'. Endogenous, radiolabelled p160 and p67 proteins were affinity purified using the GST-NRD2 fusion protein, which contains the c-Myb leucine zipper motif. Affinity purified nuclear extracts were fractionated by SDS-PAGE. Proteins were detected by Phosphorimager. The p160 and p67 proteins are identified by labels shown on the right. The positions of molecular weight markers (kDa) are shown on the left.

(B) Quantitative analysis of the pulse-chase shown in part (A). The relative amounts (expressed as arbitrary units) of p160 and p67 at each time point are represented graphically. Relative amounts were quantitated (using ImageQuant software; Molecular Dynamics, USA) with respect to the total amount of labelled protein extracted from cells at each time point.

A

0 hr 0.5 hr 1 hr 2 hrs 4 hrs

216 —

◄ p160

110 —

71 —

◄ p67

43 —
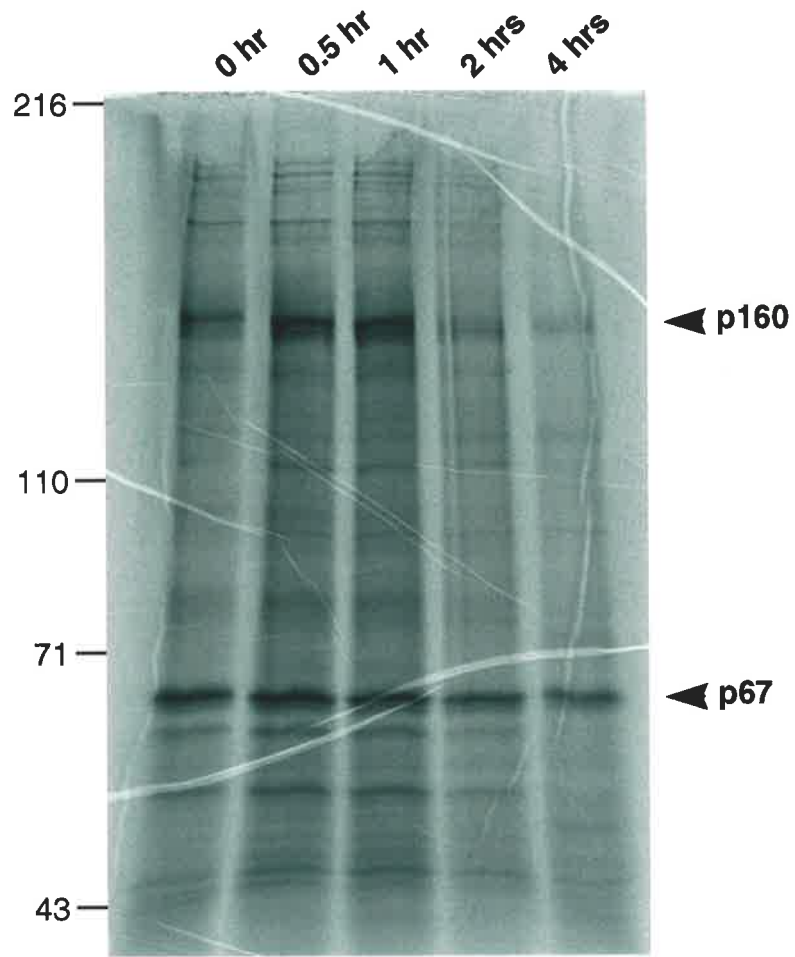
B
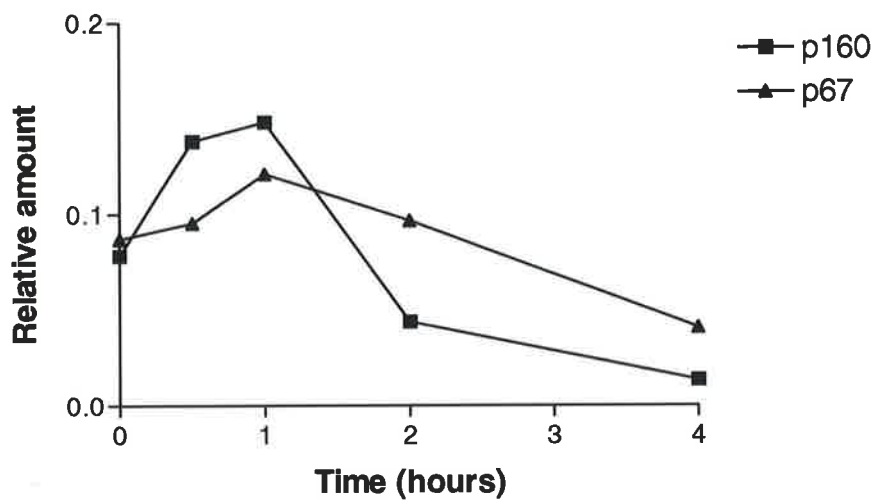
diminished greatly (after maximum incorporation) during the time course, in which the amount of label in p67 accumulated, then gradually diminished. The gradual diminishment of p67 after the one hour time point may indicate that this protein is in turn being degraded. Alternatively, the diminishing presence of both p160 and p67 after the one hour time point, albeit at differing rates, may indicate that these processes are independent events. That is, p160 and p67 may be independently synthesised and are subsequently degraded at differing rates.

4.2.2   Proteolysis of *in vitro* translated p160 by cell extracts

If p67 is derived by proteolytic cleavage of p160, then cells in which p67 was detected should contain a proteolytic activity that is able to cleave p160. This was examined by incubating *in vitro* translated, radiolabelled p160 with total cell extracts prepared from FDC-P1 or WEHI-3B cells, in which both p160 and p67 were detected. Proteolytic cleavage of *in vitro* translated p160 by FDC-P1 or WEHI-3B total cell extracts generated four major cleavage products of approximately 97, 87, 67 and 60 kDa (Figure 4.2A). The apparent molecular weights of these cleavage products were estimated from molecular weight markers. Comparison of the p160 proteolysis profiles, generated by FDC-P1 or WEHI-3B total cell extracts during a 1 to 90 minute time course, indicates that additional cleavage products were generated by the WEHI-3B proteolytic activity (Figure 4.2A). Both profiles indicate that proteolysis of p160 is rapid. This finding is consistent with the pulse-chase analysis of p160 in FDC-P1 cells. The cleavage profile generated by WEHI-3B total cell extracts demonstrates that proteolysis of p160 is even more rapid than that seen with FDC-P1 cell extracts (eg compare 20 minute time points) (Figure 4.2A). This rapid proteolysis may indicate that the proteolytic activity within WEHI-3B cells is relatively more abundant and/or efficient, or entirely distinct, with respect to that within FDC-P1 cells.

Furthermore, consistent with the distribution of p67 among murine cell lines, total cell extracts prepared from NIH3T3 cells, in which p67 was not detected, were unable to proteolytically cleave *in vitro* translated p160, even after prolonged periods of incubation (Figure 4.2B). The inability of NIH3T3 total cell extracts to cleave p160 *in vitro* indicates an inhibition or absence of the corresponding proteolytic activity within these cells, with respect to the proteolytic activities of FDC-P1 and WEHI-3B cells.

Figure 4.2 *In vitro* proteolysis of p160 by total cell extracts.

Proteolysis assays were performed with *in vitro* translated, radiolabelled p160 and 10 µg of total cell extract, prepared from either FDC-P1, WEHI-3B or NIH3T3 cells. Proteolysis reactions were incubated for varying times (shown in minutes) at 37°C, for up to 90 minutes. Reactions without any total cell extract added were included as a control. Proteolysis reactions were analysed by SDS-PAGE. Proteins were detected by autoradiography.

(A) Proteolysis of p160 by FDC-P1 and WEHI-3B total cell extracts. The four major p160 proteolytic cleavage products are identified by labels shown on the left and right. Additional proteolytic cleavage products generated by the WEHI-3B proteolytic activity are denoted by a black star.

(B) Proteolysis of p160 by FDC-P1 and NIH3T3 total cell extracts. The four major p160 proteolytic cleavage products are identified by labels shown on the left. The positions of molecular weight markers (kDa) are shown on the right.
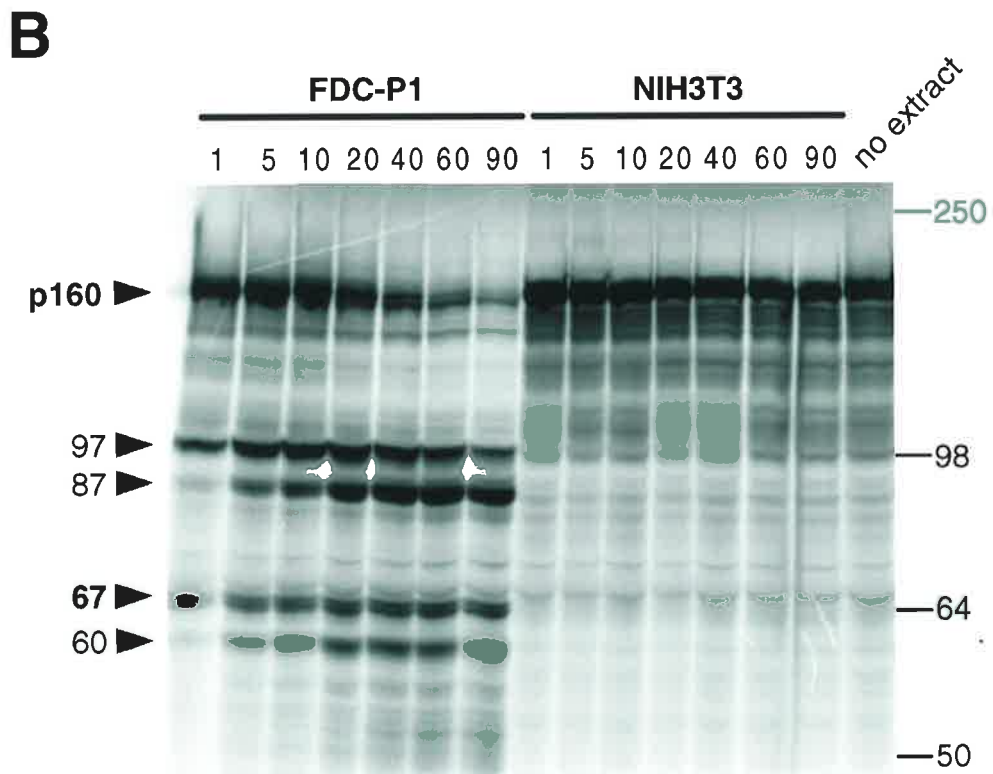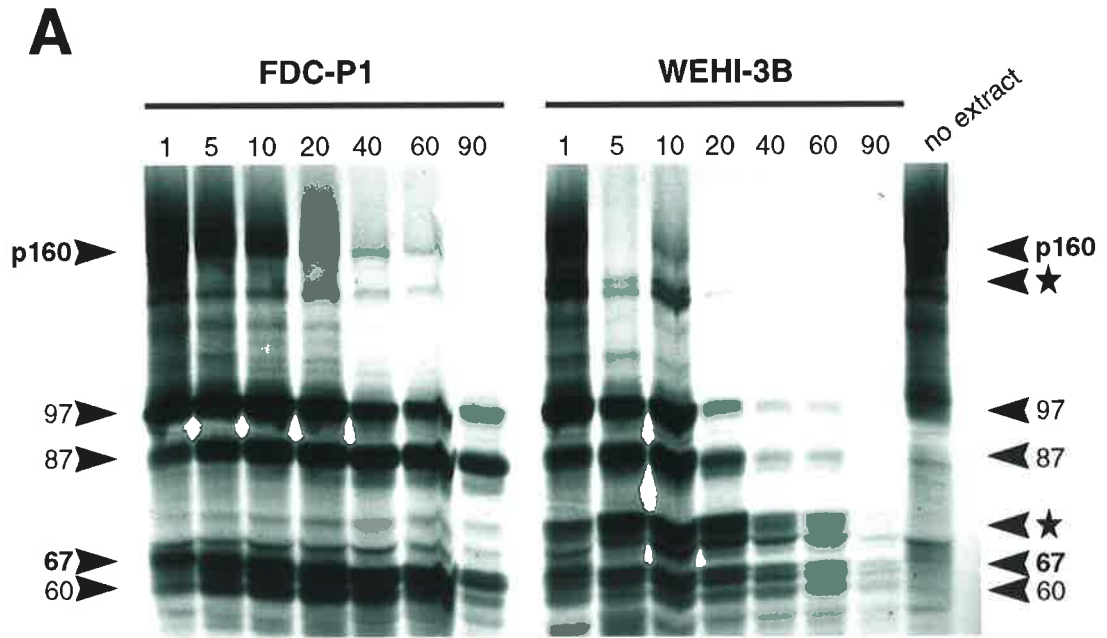
The four major cleavage products generated by the proteolysis of p160 by FDC-P1 total cell extracts include a protein of approximately 67 kDa. It was of particular interest to determine if this 67 kDa product mapped to the amino-terminal region of p160, which contains peptide sequences derived from endogenous p67. The p160 proteolytic cleavage products were mapped with respect to the p160 precursor protein by immunoprecipitation with either the 160N or 160C antisera, raised against the amino- and carboxyl-terminal regions of p160, respectively. Proteolysis reactions containing *in vitro* translated p160 and FDC-P1 total cell extract were incubated for 1 or 10 minutes, and subsequently immunoprecipitated.

Each antiserum was tested for the ability to immunoprecipitate *in vitro* translated proteins. The 160N antiserum immunoprecipitated *in vitro* translated p160, and p67*, a carboxyl-truncated form of p160 which behaves similarly to endogenous p67 in binding assays (Figure 4.3A). A high molecular weight smear was observed with immunoprecipitated p160, and is reminiscent of that seen in binding assays in which *in vitro* translated p160 was used (see Figure 3.9). As expected, the 160C antiserum only immunoprecipitated *in vitro* translated p160 (seen as a high molecular weight smear), but not p67*, which lacks carboxyl-terminal residues of p160 (Figure 4.3B). Neither p160 or p67 were precipitated by an irrelevant antibody or Protein A sepharose alone (data not shown).

Immunoprecipitation of proteolysis reactions with the 160N antiserum showed that the 97, 87 and 67 kDa products were immunoprecipitated by this antiserum, and therefore map to the amino-terminal region of p160 (Figure 4.3C). No cleavage product was immunoprecipitated by the 160C antiserum (Figure 4.3C). The 60 kDa product was not immunoprecipitated by either the 160N or 160C antisera and must presumably map to a central region of p160 (Figure 4.3C). It was noticed that an approximately 63 kDa protein was precipitated by the 160C antiserum (Figure 4.3C). This protein does not correspond to the 60 kDa cleavage product, and its presence remains obscure.

In summary, the FDC-P1 proteolytic activity cleaved *in vitro* translated p160, firstly within the carboxyl-terminal region, and subsequently generated products of 97, 87 and 67 kDa that contain the amino-terminal region of p160, and a 60 kDa product that does not contain

Figure 4.3 **Immunoprecipitation of *in vitro* translated p160 and p67*, and p160 proteolytic cleavage products, with antisera raised against p160.**

(A) Immunoprecipitation of *in vitro* translated p160 (160 kDa) and p67* (60 kDa) proteins with the 160N antiserum (raised against the amino-terminal region of p160). The *in vitro* translated proteins are shown in the left panel. Immunoprecipitates (denoted by 'IP') are shown in the right panel. Proteins are identified by labels shown on the right. Proteins were detected by autoradiography.

(B) Immunoprecipitation of *in vitro* translated p160 (160 kDa), but not p67* (60 kDa) with the 160C antiserum (raised against the carboxyl-terminal region of p160). The *in vitro* translated proteins are shown on the left of the gel. Immunoprecipitates (denoted by 'IP') are shown on the right of the gel. Note the high molecular weight smear of the p160 protein. Proteins are identified by labels shown on the left. Proteins were detected by Phosphorimager.

(C) Immunoprecipitation of the proteolytic cleavage products generated by proteolysis of *in vitro* translated p160 by FDC-P1 total cell extracts. Proteolysis reactions contained 50 µg FDC-P1 total cell extract and *in vitro* translated p160, and were incubated for 1 or 10 minutes at 37°C. Prior to immunoprecipitation with either the 160N or 160C antisera, an aliquot was taken from each reaction after incubation to show the resulting proteolysis profile (shown in the left panel; note that only one profile is shown, as all profiles were identical). Immunoprecipitates (denoted by 'IP') of the 160N and 160C antisera are shown in the middle and right panels, repectively. The p160 protein and resultant proteolytic cleavage products (sizes shown in kDa) are identified by labels shown on the left. Proteins were detected by autoradiography.
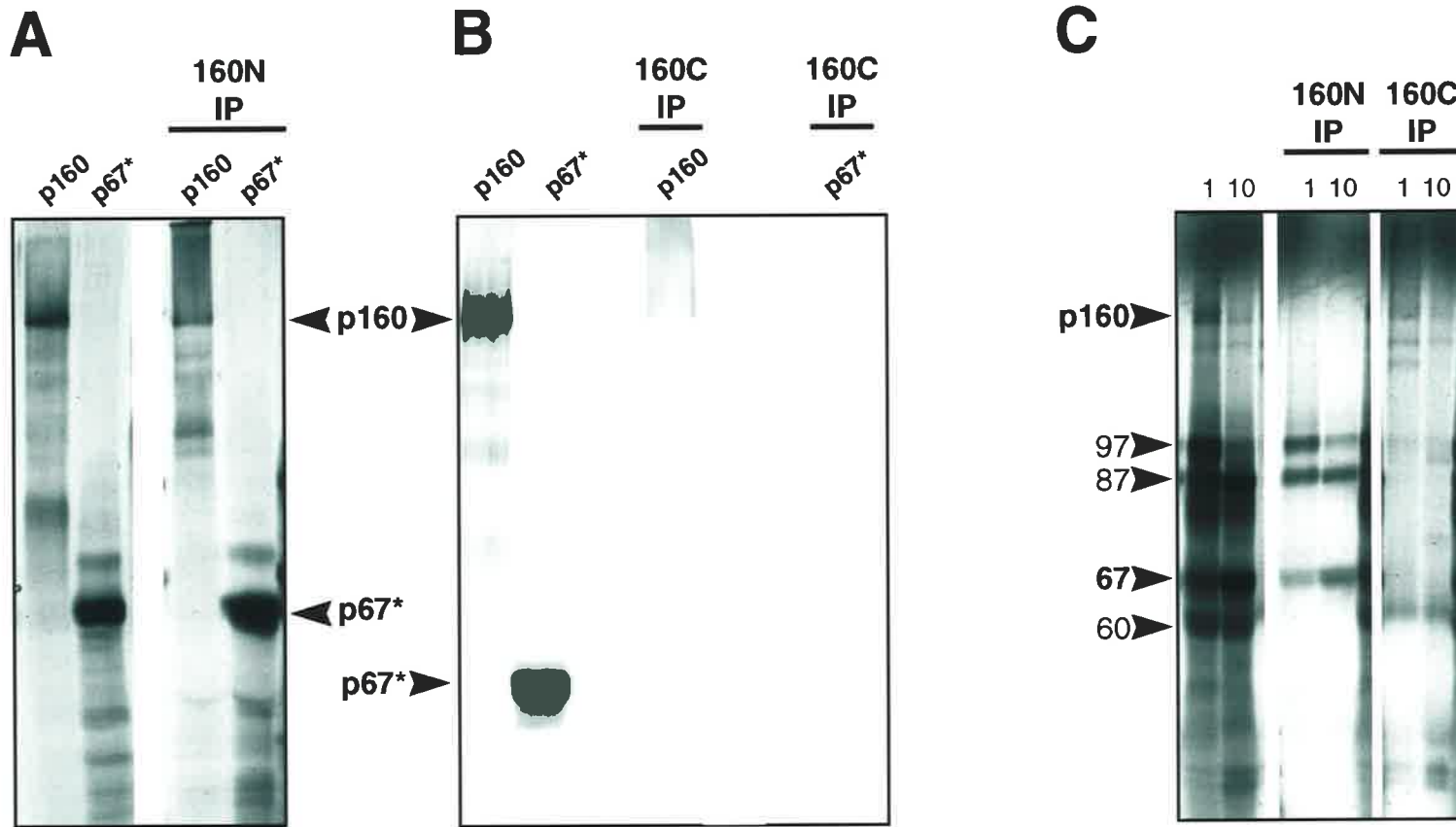
either amino- or carboxyl-terminal regions. Importantly, the 67 kDa cleavage product maps to the region of p160 in which the p67-derived peptide sequences are located.

### 4.2.3 Inhibition of proteolytic cleavage of p160

It was of interest to identify a specific inhibitor of the proteolytic activity that cleaved p160. This purpose was two-fold. Firstly, identification of an inhibitor(s) may reveal the nature of the proteolytic activity (eg class of protease). The second reason for identifying an inhibitor of p160 proteolysis pertains to the observation that endogenous p67 could not be detected within FDC-P1 whole cell lysates, prepared by SDS lysis of these cells (see Figure 3.16A and also Figure 4.5). The ability to detect p67 within nuclear extracts, but not within whole cell lysates of the same cells, may be due to the relative amount of p67 within each extract. That is, nuclear fractionation may enrich for p67. However, this explanation is unlikely as p160 can be readily detected within whole cell lysates and is routinely detected in a roughly equal proportion (if not slightly less) to that of p67 within nuclear extracts. Alternatively, p67 may be an artefact of the nuclear extraction procedure, whereby it is generated by proteolysis of p160 during the process of cell fractionation or protein extraction. If this were the case, during the preparation of whole cell lysates, proteolysis would presumably not occur due to the presence of SDS, a strong denaturant of proteins, and thus no p67 would be detected. Identification of an effective inhibitor of p160 proteolysis, which could be added during cell fractionation, would aid in the resolution of this issue.

Seventeen protease inhibitors (see Table 2.1) were assayed for their ability to inhibit the proteolysis of p160 by FDC-P1 total cell extracts. Together, these protease inhibitors inhibit a broad spectrum of proteolytic activities, which include common classes of proteases (eg serine, cysteine and metallo proteases), ICE-like proteases (induced in apoptotic cells) and 26S proteasome function, which mediates proteolysis of ubiquitinated proteins. Proteolysis reactions containing *in vitro* translated p160 and FDC-P1 total cell extract were performed essentially as described in section 4.2.2, except that each specific protease inhibitor was pre-incubated (for 20 minutes) with the FDC-P1 cell extract, after which p160 was added and the reactions incubated for 1 hour. One protease inhibitor, pefabloc (4-[2-aminoethyl]-benzenesulfonyl fluoride hydrochloride), was able to completely inhibit the proteolysis of p160 by FDC-P1 total cell extracts (Figure 4.4). In addition, TPCK (*N*-tosyl-L-phenylalanine chloromethyl ketone) partially inhibited p160

Figure 4.4 Inhibition of p160 proteolysis by FDC-P1 total cell extracts.

Proteolysis reactions were performed with 10 μg FDC-P1 total cell extract, *in vitro* translated p160 and a single protease inhibitor (see Table 2.1). Prior to addition of radiolabelled p160, each inhibitor was pre-incubated with total cell extract. Proteolysis reactions containing an absence of inhibitor ('+ extract') or total cell extract ('- extract'), were included as controls. The p160 protein and resultant proteolytic cleavage products (sizes shown in kDa), including p67, are identified by labels shown on the left. Vertical arrows beneath the gel identify protease inhibitors that blocked the production of the 67 kDa cleavage product. The numbers 1-18, above the gel, correspond to the following protease inhibitors;

1. Antipain dihydrochloride
2. Aprotinin
3. Bestatin
4. Chymostatin
5. E-64
6. EDTA
7. Leupeptin
8. Pefabloc
9. Pepstatin
10. Phosphoramidon
11. Ac-DEVD-CHO
12. Ac-YVAD-CMK
13. PMSF
14. Iodoacetamide
15. TLCK
16. TPCK
17. MG132 (Al)
18. MG132 (OH)
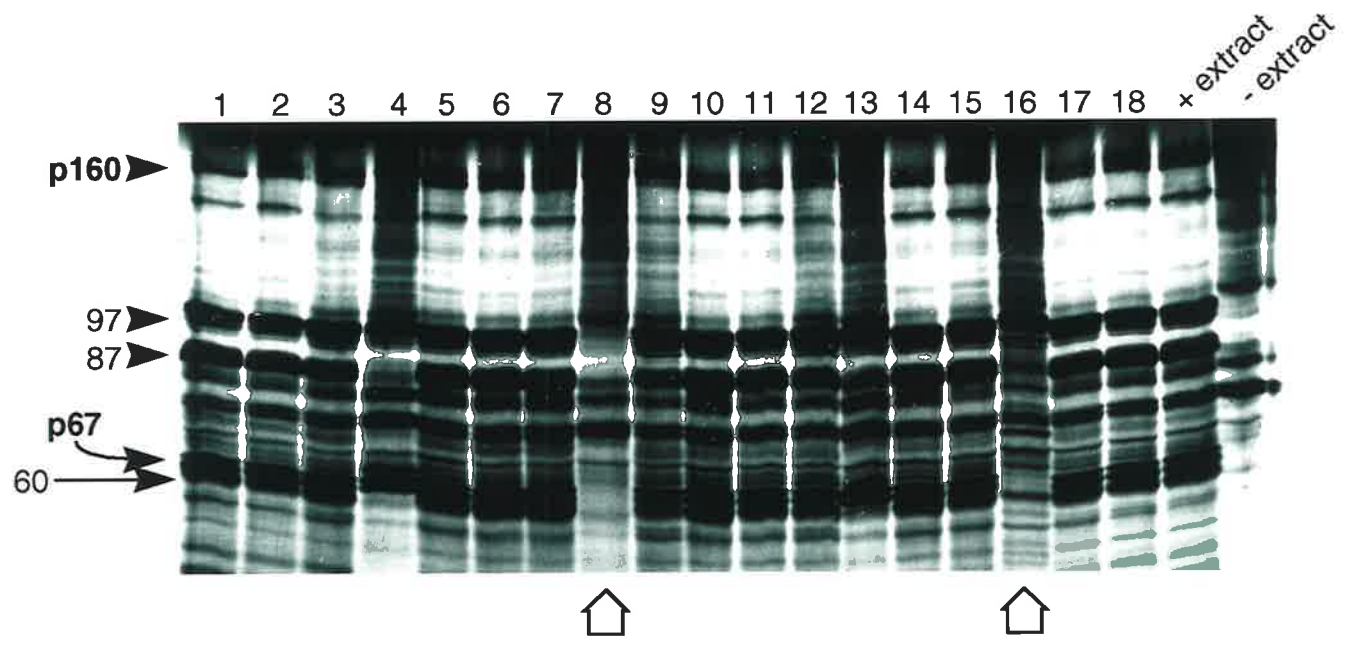
proteolysis (Figure 4.4). Two other inhibitors, PMSF (phenyl-methyl sulfonyl fluoride) and chymostatin, also partially inhibited p160 proteolysis, but not production of the 67 kDa cleavage product (Figure 4.4). Pefabloc specifically inhibits many serine proteases. The TPCK inhibitor also inhibits many serine proteases, as well as cysteine proteases, and is consistent with the protease-inhibiting properties of pefabloc. It is therefore concluded that the FDC-P1 proteolytic activity that cleaved p160 is composed of a serine protease(s). Furthermore, the partial inhibition of p160 by chymostatin may indicate that the proteolytic activity is chymotrypsin (or a chymotrypsin-like protease), a serine protease that can be inhibited by chymostatin, pefabloc, TPCK and PMSF. However, chymotrypsin is a proteolytic enzyme associated with the digestive system, and therefore is not consistent with the cell type used in this inihibitor assay (ie the FDC-P1 murine haemopoietic cell line).

Identification of a specific protease inhibitor that is capable of blocking the proteolysis of p160 facilitated a further investigation of p67 within cell extracts prepared by different procedures. Both p160 and p67 were detected within nuclear extracts of FDC-P1 and WEHI-3B cells by Western analysis with antiserum raised against the amino-terminal region of p160 (160N) (Figure 4.5A). In contrast, p160, but not p67, was detected within FDC-P1 and WEHI-3B whole cell (SDS) lysates (Figure 4.5A). Comparison of WEHI-3B nuclear extracts and whole cell lysates indicates that very little full-length p160 was detected within nuclear extracts, which may be indicative of proteolysis during the nuclear extraction procedure (Figure 4.5A). In order to determine whether p67 was generated during the nuclear extraction procedure, the pefabloc protease inhibitor was added to each extraction solution used to prepare nuclear extracts. Analysis of FDC-P1 and WEHI-3B nuclear extracts prepared in the presence of pefabloc showed that p67 was not detected within FDC-P1 nuclear extracts, and was barely detected within WEHI-3B nuclear extracts (Figure 4.5B). The detection of a relatively small amount of p67 within WEHI-3B extracts probably indicates that the proteolysis of p160 was not entirely inhibited by pefabloc, as a number of additional cleavage products were also detected (Figure 4.5B). This indicates that a greater quantity of pefabloc may be required to completely inhibit the WEHI-3B proteolytic activity, compared to the corresponding FDC-P1 activity. The differential proteolytic activities of WEHI-3B and FDC-P1 cells are reflected in previous p160 proteolysis assays (Figure 4.2A). The inability to detect significant amounts of p67 within

Figure 4.5   Detection of endogenous p160 and p67 proteins within nuclear extracts or whole cell lysates, with the 160N antiserum.

(A) Detection of endogenous p67 within nuclear extracts, but not whole cell lysates, prepared from FDC-P1 or WEHI-3B cells. Both p160 (160 kDa) and p67 (67 kDa) were detected in FDC-P1 and WEHI-3B nuclear extracts (shown in the left panel), with the 160N antiserum (1/500). 'N' and 'C' denote nuclear and cytoplasmic fractions (70 µg), respectively. Endogenous proteins are identified by labels shown on the left. Note that very little full-length p160 was detected within the WEHI-3B nuclear extract. In addition, several proteolytic cleavage products were also detected in the FDC-P1 and WEHI-3B nuclear extracts. The p160 protein, but not p67, was detected within FDC-P1 or WEHI-3B whole cell lysates (200 µg), with the 160N antiserum (1/1000) (shown in right panel). Whole cell lysate prepared from NIH3T3 cells was included as a control, as p67 was not detected in nuclear extracts prepared from these cells. The endogenous p160 protein (160 kDa) is identified by a label shown on the right. The absence of p67 is denoted by an asterisk. Proteins were detected by Western analysis/ECL.

(B) Detection of endogenous p67 within FDC-P1 or WEHI-3B nuclear extracts (70 µg) prepared in the prescence of pefabloc, an inhibitor of p160 proteolysis. Only p160 (160 kDa) was detected within FDC-P1 nuclear extracts prepared in the presence of pefabloc (1 mg/ml). Very little endogenous p67 (67 kDa) was detected within WEHI-3B nuclear extracts prepared with pefabloc. Note that additional p160 proteolytic cleavage products were also identified within the WEHI-3B nuclear extract. Endogenous proteins are identified by labels shown on the left and right. Proteins were detected with the 160N antiserum (1/1000) by Western analysis/ECL. 'N' and 'C' denote nuclear and cytoplasmic fractions, respectively.

A

FDC-P1   WEHI-3B
N   C   N   C

p160 ▶

p67 ▶

FDC-P1 WEHI-3B NIH3T3

◀ p160

◀ *

WHOLE
CELL LYSATE

B

FDC-P1 WEHI-3B
N   C   N   C

p160 ▶

◀ p67
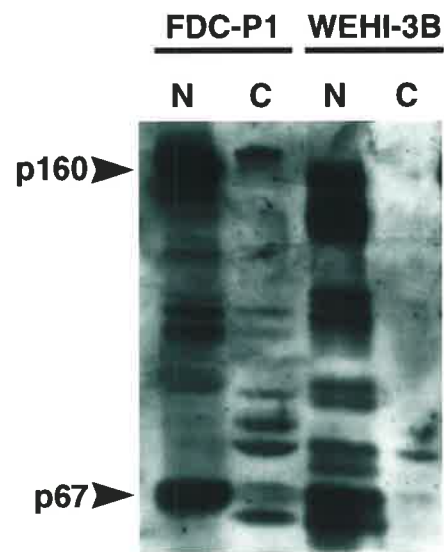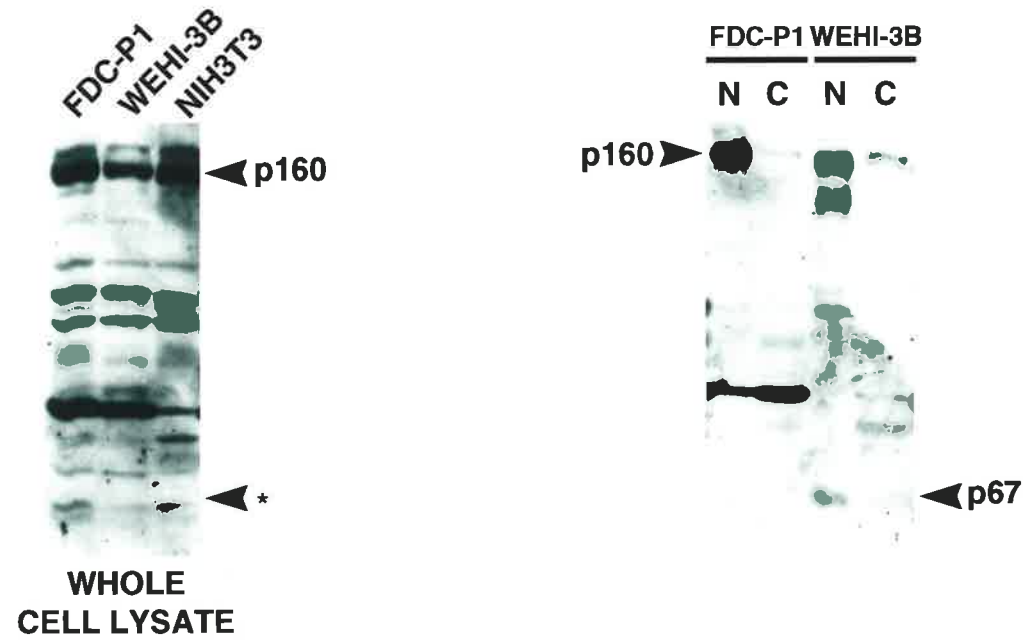
FDC-P1 and WEHI-3B nuclear extracts prepared in the presence of pefabloc leads to the conclusion that p67 is an artefact of the nuclear extraction procedure. However, this conclusion does not immediately dismiss the existence of p67 *in vivo* (see below).

4.2.4 Can p67 be generated *in vivo*?

In view of the fact that p67 can be generated artefactually during the nuclear extraction procedure, it was necessary to repeat the pulse-chase analysis of p160 in FDC-P1 cells. This analysis was repeated in an identical manner to the previous pulse-chase (section 4.2.1), except that nuclear extracts were prepared in the presence of pefabloc, a protease inhibitor that was shown (see above) to block the proteolysis of p160 by an FDC-P1 proteolytic activity. Figure 4.6 clearly shows that p67 was not detected in the nuclear fractions of FDC-P1 cells during the specified time course, in which proteolysis of p160 was expected to occur if the corresponding proteolytic activity was active. Extracts prepared from cells in this pulse-chase experiment should have been prepared in both the presence or absence of the pefabloc protease inhibitor. Therefore this pulse-chase experiment lacked a suitable control, and consequently remains inconclusive.

A cell type-specific proteolytic activity was able to cleave p160 *in vitro*. Presumably, proteolytic cleavage of p160 also occurs *in vivo*. Proteolytic cleavage of p160 *in vivo* could not be demonstrated by pulse-chase analysis (Figure 4.6). However, it is conceivable that the proteolytic activity is transient, and activated in response to certain physiological conditions (eg induction of cellular differentiation or growth factor deprivation). Without knowing the specific functional activities of p160 and p67, it is difficult to predict the cellular conditions that may induce the generation of p67 *in vivo*.

Preliminary experiments were undertaken to determine whether differing cellular conditions could induce proteolysis of p160 and subsequent generation of p67. Growth factor (granulocyte-macrophage colony-stimulating factor (GM-CSF)) deprivation and serum starvation of FDC-P1 cells were investigated with respect to p160 proteolysis, but analyses of the corresponding whole cell lysates remain inconclusive (data not shown). A WEHI-3B cell line that can be induced to differentiate with human GM-CSF (Smith *et al.*, 1997) was used to examine the expression of p160, and presence of p67 during cellular differentiation. Cells were induced to terminally differentiate over a period of 5 days,

Figure 4.6   Pulse-chase analysis of p160 in FDC-P1 cells, after nuclear extracts were prepared in the presence of pefabloc, an inhibitor of p160 proteolysis.

(A) FDC-P1 cells were metabolically labelled with [^{35}S]-methionine for 1 hour (the 'pulse'), after which a 1000-fold molar excess of methionine was added (the 'chase'), and cells incubated for up to 4 hours. Nuclear extracts were prepared in the presence of pefabloc (1 mg/ml) from cells at the specified time points. The '0 hour' time point corresponds to the start of the 'chase'. Endogenous, radiolabelled p160 was affinity purified using the GST-NRD2 fusion protein, which contains the c-Myb leucine zipper motif. Affinity purified nuclear extracts were analysed by SDS-PAGE. The p160 protein is identified by a label shown on the right. Note the absence of p67 (67 kDa). The positions of molecular weight markers (kDa) are shown on the left.
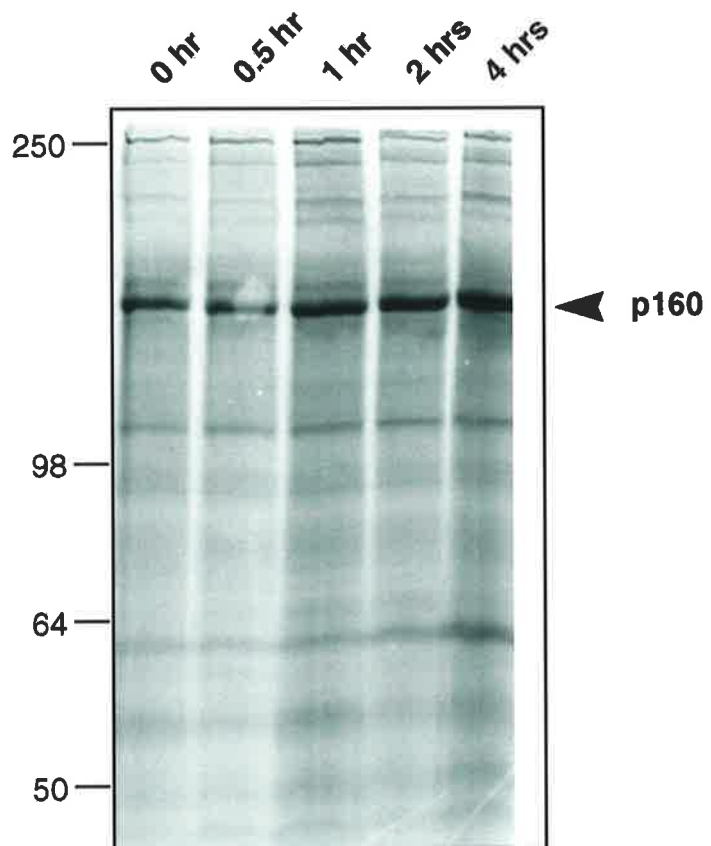
during which time morphological changes (as assessed by staining cells with giemsa) were observed as cells differentiated from myelomonocytic precursor-like cells into macrophages. Whole cell lysates were prepared from differentiating cells at 24 hour intervals, for up to 5 days, and analysed by SDS-PAGE. Unexpectedly, an approximately 67 kDa protein was detected within whole cell lysates prepared from undifferentiated and differentiated cells, with the 160N antiserum (raised against the amino-terminal region of p160) (Figure 4.7). This protein has was not detected previously within WEHI-3B whole cell lysates, and its identity remains obscure. However, like p67, this protein was not detected by the 160C antiserum (raised against the carboxyl-terminal region of p160) (data not shown). Additionally, it was noticed that there was not much change in the level of this 67 kDa protein during the course of differentiation. No conclusions can be drawn from this experiment with respect to the generation of p67 *in vivo*. This experiment needs to be repeated to verify the presence of the ~67 kDa protein within WEHI-3B whole cell lysates. Examination of the level of p160 within whole cell lysates prepared from cells at various stages of induced differentiation suggest that expression of p160 may decline as cells terminally differentiate (Figure 4.7). It will be interesting to examine the p160 mRNA transcript during cellular differentiation in order to determine whether the apparent decrease in the amount p160 (Figure 4.7) is a consequence of a reduction in the level of the corresponding transcript.

4.3   Discussion

4.3.1   The precursor-product relationship of p160 and p67

A structural relationship between p160 and p67 was initially indicated by comparative peptide mapping, which demonstrated that these proteins share numerous peptides (Favier and Gonda, 1994). This structural relationship was further indicated when a cDNA clone corresponding to p160 was obtained using sequence information derived from p67 (Tavner *et al.*, 1998; see appendix, and Chapter 3). Evidence presented in this chapter demonstrates that the structural relationship of p160 and p67 is one of precursor and product, respectively. Furthermore, the restricted distribution of p67, within murine immature myeloid cell lines, correlates with the presence of a proteolytic activity that can cleave p160.

Figure 4.7 Detection of p160 and p67 within whole cell lysates prepared from WEHI-3B cells induced to differentiate.

A WEHI-3B cell line, expressing the human GM-CSF receptor α-chain and common β-chain (Smith *et al.*, 1997), was induced to differentiate with human GM-CSF over a period of 5 days. Whole cell lysates were prepared from cells at 24 hour intervals after induction (shown on the left of the gel). 'D0' denotes cells grown in the absence of GM-CSF (uninduced). 'D1-4' denotes cells grown in the presence of GM-CSF for 1, 2, 3, or 4 days, respectively. Note that the 'day 5' whole cell lysate is not shown, as a large proportion of cells at this time point were dead (due to overgrowth). Proteins were detected by Western analysis/ECL using the 160N antiserum (1/500). Endogenous p160 is identified by a label shown on the left. The 67 kDa protein, which was not previously detected within WEHI-3B whole cell lysates, is identified by a question mark (?). The positions of molecular weight markers (kDa) are shown on the left. Nuclear ('N') and cytoplasmic ('C') extracts prepared from FDC-P1 and WEHI-3B (uninduced) cells are shown on the right of the gel, and were included to show the position of endogenous p67. The positions of endogenous p160 and p67 are shown on the right.

WEHI-3B DIFFERENTIATION

D0 D1 D2 D3 D4

WEHI-3B   FDC-P1

N C   N C

250 —

p160 ▶

98 —

? ➤

64 —

◀ p160

◀ p67

WHOLE CELL LYSATES

The FDC-P1 and WEHI-3B cell lines, in which both p160 and p67 were detected (Favier and Gonda, 1994), contain proteolytic activities that cleave p160. *In vitro* proteolysis assays showed that the FDC-P1 and WEHI-3B proteolytic activities rapidly cleave *in vitro* translated p160 and subsequently generate four major products, including a 67 kDa cleavage product that maps to the amino-terminal region of p160, in which the peptide sequences derived from p67 are located. Two other cleavage products, of 97 and 87 kDa, also map to the amino-terminal region of p160. The 67 kDa protein is the only proteolytic cleavage product that is consistently detected (either by binding to the c-Myb leucine zipper-containing GST-NRD2 fusion protein or by antiserum raised against the amino-terminal region of p160) within nuclear extracts prepared from FDC-P1 and WEHI-3B cells. If proteolysis of p160 occurs similarly *in vivo*, as it does *in vitro*, the 97 and 87 kDa products may be subsequently degraded and therefore not detected within nuclear extracts. In particular, proteolysis assays suggest that the 97 kDa product may be an intermediate precursor for the 67 kDa product, therefore suggesting that endogenous p67 may be derived by at least two proteolytic cleavage steps from the endogenous p160 precursor protein. Furthermore, it is proposed that the first proteolytic cleavage step occurs within the carboxyl-terminal region of p160, as none of the major cleavage products contain sequences derived from the carboxyl-terminal region of p160.

The FDC-P1 and WEHI-3B proteolytic activities similarly cleave p160 *in vitro*, however, there are a number of differences between these activities. Although each proteolytic activity rapidly cleaved p160 , they did so with differing kinetics. Proteolysis of p160 by the WEHI-3B activity was extremely rapid. This was evident not only in proteolysis assays performed *in vitro*, but also by examination of WEHI-3B nuclear extracts, in which very little full-length p160 protein was detected. The WEHI-3B proteolytic activity also generated additional cleavage products from the p160 precursor protein *in vitro*, indicating that this activity may be composed of more than one protease. Alternatively, the WEHI-3B proteolytic activity may be composed of a protease(s) that is distinct to that composing the FDC-P1 activity. If the latter situation is true, it is presumed that this protease has similar cleavage properties to the FDC-P1 proteolytic activity, as both of these activities generated similar p160 proteolysis profiles *in vitro*. Both the FDC-P1 and WEHI-3B proteolytic activities could be specifically inhibited by pefabloc, an inhibitor of various serine proteases. In particular, examination of a broad range of protease inhibitors suggests that

the FDC-P1 proteolytic activity may be composed of a chymotrypsin-like protease. It remains to be determined whether the WEHI-3B proteolytic activity can be inhibited by the same group of protease inhibitors that were shown to differentially inhibit the FDC-P1 proteolytic activity. Therefore, although the FDC-P1 and WEHI-3B proteolytic activities were both inhibited by pefabloc, they may be composed of different serine proteases, as suggested by *in vitro* proteolysis assays.

4.3.2 Generation of p67 *in vivo*

Pulse-chase analysis of p160 in FDC-P1 cells initially suggested that p67 can be generated from p160 *in vivo*. This analysis was repeated after it was shown that p67 can be generated artefactually during the nuclear extraction procedure. Analysis of this subsequent pulse-chase experiment, in which the pefabloc protease inhibitor had been added to the protein extraction solutions to stop the artefactual generation of p67 from p160, revealed that p67 could not be detected within the nuclear fractions. There are several possible explanations that may account for the inability to detect p67 in this latter pulse-chase experiment (assuming that the FDC-P1 proetolytic activity was indeed active); (1) If p67 can be generated *in vivo* and proteolytic cleavage of p160 occurs within the nucleus, p67 may be exported to the cytoplasm (by a putative nuclear export signal(s) within the amino-terminal region of p160; see Chapter 3.3.4) following proteolytic cleavage of p160. However, endogenous p67 has not yet been detected within cytoplasmic fractions by Western analysis, although in this context, p67 was presumably generated during the protein extraction procedure, in which cells obviously do not remain intact. Unfortunately, the cytoplasmic fractions of the repeated FDC-P1 pulse-chase experiment were not analysed. The corresponding cytoplasmic fractions of the FDC-P1 pulse-chase experiment will have to be investigated by repeating the pulse-chase analysis. The subcellular localisation of the FDC-P1 proteolytic activity currently remains unknown. (2) The proteolytic activity that cleaves p160 may only be active during certain physiological conditions. In this scenario, p67 would be generated transiently by activation of the proteolytic activity. This possibility was investigated by a preliminary examination of particular cell conditions, namely growth factor (GM-CSF) deprivation and serum starvation of FDC-P1 cells, and induced differentiation of WEHI-3B cells. However, these experiments were inconclusive and need to be repeated. It is difficult to predict what conditions may induce proteolytic cleavage of p160, without knowing the specific functions of p160, and possibly p67. (3) The p67

protein may have been generated *in vivo*, but was subsequently rapidly degraded. *In vitro* proteolysis assays using FDC-P1 total cell extracts indicate that the proteolysis of p160 is a rapid process, and that the 67 kDa cleavage product was degraded after prolonged incubation (> 60 minutes) with the cell extract. If the proteolysis of p160 and subsequent cleavage products occurs with similar kinetics *in vivo*, then the pulse-chase experiment may have been undertaken with an inappropriate time window. Although, an early time point (0.5 hour) was taken in this pulse-chase experiment, but no p67 was seen. In order to determine whether an inappropriate time course was taken, the pulse-chase experiment will have to be repeated, but using a much shorter time course, up to 90 minutes, and taking several time points within the first 60 minutes.

If p67 is generated *in vivo*, by either a constitutive or transient proteolytic activity, then the following question remains the most pertinent at this point: Is p67 generated specifically for a functional requirement, or does p67 represent a transitional degradation product of p160? There are numerous examples of proteolysis-mediated activation of protein functionality. For example, NF-κB and sterol regulatory element binding proteins (SREBPs) are both transcription factors whose activities are activated by proteolytic cleavage events. The NF-κB protein is a heterodimeric protein complex, composed of RelA (p65) and NF-κB1/ NF-κB2 (p50/p52), that becomes transcriptionally active in response to inflammation and other immunological stimuli (reviewed in Finco and Baldwin, 1995). NF-κB becomes active by dissociation from IκB, an inhibitory molecule that sequesters NF-κB within the cytoplasm and therefore prevents translocation of NF-κB into the nucleus. Dissociation of IκB from NF-κB is mediated by phosphorylation, and proteolysis via the ubiquitin-26S proteasome pathway (reviewed in Finco and Baldwin, 1995). In addition, the p50/p52 subunits of NF-κB are proteolytic cleavage products of the p105/p100 precursor proteins. These precursors act similarly to the IκB inhibitor, as they bind to, and sequester the RelA (p65) subunit of NF-κB within the cytoplasm (reviewed in Finco and Baldwin, 1995).

Similarly, the SREBPs are translocated to the nucleus following a cascade of proteolytic cleavage events. However, unlike NF-κB, SREBPs are located within membranes of the endoplasmic reticulum and nuclear envelope (reviewed in Brown and Goldstein, 1997). In response to low sterol (eg cholesterol) levels, an active transcription factor (of the basic-

helix-loop-helix-leucine zipper family) is generated by proteolytic cleavage of the membrane-bound SREBP precursor protein. The SREBP proteolytic cleavage products are subsequently transported into the nucleus where they transcriptionally regulate genes required for sterol and fatty acid metabolism (reviewed in Brown and Goldstein, 1997).

It is conceivable that p67 may be generated *in vivo* for a specific requirement. However, as the functions of p160 are only beginning to be investigated, it is difficult to speculate on what the significance of p67 may be. It is possible that p67 may act as an inhibitor of Myb-dependent transcriptional activation, by analogy to the p160 carboxyl-truncated protein, p67*, which has several properties in common with endogenous p67 and can inhibit transactivation of the c-*myc* promoter by c-Myb in reporter assays (Tavner *et al.*, 1998; see appendix, and also Chapter 3). It is also possible that p67 may play a nucleolar role, although it is not known whether p67 is associated with the nucleolus. In addition, the potential nucleolar function(s) of p160 have not yet been determined.

CHAPTER FIVE
CONCLUSIONS AND PERSPECTIVES

The p160 and p67 nuclear proteins were originally detected as a consequence of their ability to interact with the c-Myb leucine zipper motif, a component of the negative regulatory domain (Favier and Gonda, 1994). In addition to their functional relatedness, p160 and p67 are also structurally related. Despite the relatedness of these proteins, each possess distinct properties that pertain to their distribution among murine cell lines, and the ability to interact with the bZIP region of the c-Jun transcription factor (Favier and Gonda, 1994).

Following purification of p67 from the FDC-P1 murine cell line, a 3.1 kb cDNA clone was isolated using sequence information obtained from tryptic peptide sequences derived from p67 (Tavner *et al.*, 1998; see appendix). The studies presented in this thesis describe the complete cloning a cDNA encoding both p160 and p67, and characterisation of the properties of these proteins. Furthermore, the structural relationship of p160 and p67 has been determined, and is one of precursor and product, respectively.

The full-length cDNA clone (4.1 kb) corresponding to murine p160 mRNA was assembled from several clones, isolated by PCR, library screening and 5' RACE PCR (Tavner *et al.*, 1998; see appendix). Analysis of this cDNA clone revealed that the p160 sequence is novel and shows similarity to regions of the *Drosophila lethal (1) 1Bi* gene product, a nucleolar protein of unknown function, and a yeast hypothetical protein (Yelo55cp), also of unknown function. The putative rat homologue of p160, a 160 kDa protein termed 'PIP', has recently been identified by database searches. The function of PIP currently remains unknown, however, this protein is described as being nucleolar, and interacts with the PAR family of transcription factors (cited in GenPept database; accession number U83590). Several features have been recognised in the p160 predicted polypeptide sequence, and include many potential protein-protein interaction domains and phosphorylation sites, an acidic region that shows similarities to numerous proteins, seven clustered, basic repeats that may constitute multiple nuclear localisation signals, and a relatively large proportion of leucine residues. Furthermore the amino-terminal region of p160 harbours all five tryptic peptide sequences derived from p67.

The binding specificities of p160, encoded by the full-length 4.1 kb cDNA clone, and an engineered protein, p67*, comprising the amino-terminal 580 residues of p160, are identical to those of endogenous p160 and p67. Thus, these studies indicate that the Myb-binding region of p160 is located within the amino-terminal region (residues 1-580) of this protein.

Polyclonal antisera were rasied against p160 and provide a useful reagent for future studies. Specifically, antiserum raised against the amino-terminal region of p160 detected both endogenous p160 and p67, whereas antiserum raised against the carboxyl-terminal region detected endogenous p160, but not p67. The differential specificities of the p160 antisera is a reflection of the relationship between p160 and p67. That is, p67 is derived from the amino-terminal region of p160 by proteolytic cleavage. Furthermore, the restricted distribution of p67, within a subset of murine immature myeloid cell lines, correlates with the presence of a cell type-specific proteolytic activity that cleaves p160. Specific inhibition of this activity, which is comprised of a serine protease(s), demonstrated that p67 can be generated artefactually during the nuclear extraction procedure. It is currently unknown if p67 does in fact exist *in vivo*. This is an important issue that needs to be resolved because if p67 is generated *in vivo*, by either a transient or constitutive proteolytic activity, the function of this protein may interfere with those of its precursor, p160, as suggested by inhibition of c-Myb transactivation by p67* (Tavner *et al.*, 1998; see appendix).

The association of p160 and p67 with the c-Myb negative regulatory domain (via interaction with the leucine zipper), a region that down-regulates several functional aspects of c-Myb, including DNA-binding, transactivation and transformation ability, would initially suggest that p160 and p67 may act as repressors of Myb activity. Characterisation of p160, by sequence analysis and transactivation assays, suggests that p160 may function as a coactivator of c-Myb-dependent transcriptional activation (Tavner *et al.*, 1998; see appendix). This proposed function of p160 is not necessarily contradictory with respect to the properties of the c-Myb negative regulatory domain, as p160 may be required for the transcriptional coregulation of certain Myb target genes, possibly those that contribute to a relatively mature cell phenotype. Direct transcriptional coregulation of gene expression by c-Myb and p160 has not yet been demonstrated. Further transactivation analyses of p160

and c-Myb, in the first instance, may need to be performed in a cell line in which p160 is absent, or present at a relatively low level, in order to avoid 'masking' of any effect by endogenous p160, and thus assess the potential requirement for p160 in a transcriptional context. Identification of a cell line in which p160 is absent may prove difficult in view of the ubiquitous expression of this protein. Alternatively, the level of p160 in a particular cell line in which transactivation analyses can be suitably studied (eg NIH3T3, CV1), could potentially be reduced by immunodepletion, as demonstrated for the p/CIP nuclear-receptor transcriptional coactivator (Torchia *et al.*, 1997), or by antisense RNA.

The selection of Myb target gene promoters to be used for transactivation analyses of p160 and c-Myb represents a difficult decision. It is possible that p160 may be required for transcriptional coactivation of only a subset of Myb target genes. It is currently unresolved as to whether p160 is required for c-Myb-dependent transactivation of the human c-*myc* promoter (Tavner *et al.*, 1998; see appendix). However, it has been shown that the p160 carboxyl-truncated protein, p67*, inhibited c-Myb-dependent transactivation of the c-*myc* promoter, and thus may be acting as a dominant negative inhibitor (Tavner *et al.*, 1998; see appendix). As discussed in Chapter 1.6.1, the *hsp70* promoter was shown to be transactivated by cooperation of heat shock transcription factor 3 (HSF3) and c-Myb (Kanei-Ishii *et al.*, 1997; Kanei-Ishii *et al.*, 1994). However, mutation of the c-Myb leucine zipper motif consequently abolished transactivation of the *hsp70* promoter, indicating that the c-Myb leucine zipper may recruit an additional protein(s) that is required for transcriptional activation of this promoter (Kanei-Ishii *et al.*, 1997; Kanei-Ishii *et al.*, 1994). These studies suggest that the *hsp70* promoter may be an appropriate choice for examining the potential transcriptional properties of p160, as p160 could be a suitable candidate for this putative c-Myb leucine zipper-interactive protein.

The ubiquitous distribution of p160 among murine cell lines, and the ability of this protein to interact with the leucine zipper motif of c-Myb, and the bZIP region of c-Jun, suggest that this protein may have numerous roles (Favier and Gonda, 1994). This notion is further supported by the finding that p160 is also associated with the nucleolus (Tavner *et al.*, 1998; see appendix). This finding is consistent with the information obtained from database searches with the p160 nucleotide and predicted polypeptide sequences. Identification of possible additional proteins that may interact with p160 will aid in

elucidating the functions of p160, which may be required for transcriptional regulation of specific genes transcribed by RNA polymerase II, and possibly ribosomal RNA genes, transcribed by RNA polymerase I. The postulated transcriptional properties of p160 have only begun to be investigated at this stage, and therefore remain largely unknown. Use of the yeast two-hybrid system would be a suitable approach for identifying additional p160 interaction partners, and would also facilitate the delineation of putative protein-protein interactive regions within p160.

Another fascinating aspect of p160 is the presence of putative nuclear import and export signals. Functionality of these putative transport signals firstly needs to be shown, by introduction into heterologous proteins, and by mutation within p160. These necessary studies could be performed by microinjection of fusion proteins containing either the putative import or export signal(s) of p160, and mutated (transport signal) p160 proteins, into subcellular compartments (ie nucleus or cytoplasm) and subsequent determination of their movement within the cell. Deletion analyses of p160, and nuclear-targeting of p67*, suggest that the putative transport signals of p160 may compete with each other. This raises the interesting possibility of a potential nucleocytoplasmic shuttling function(s) of p160. The implications of such an attribute, with respect to the activities of p160, may become clearer when the functions of p160 are further elucidated. It is speculated that the potential nucleocytoplasmic shuttling of p160 may be required to transport other proteins that may associate with p160 and/or to efficiently regulate the activity of this protein.

The mechanisms underlying the functions of the c-Myb negative regulatory domain (NRD) are only beginning to be characterised. Insight into these mechanisms is being gained by the study of cellular proteins that interact with this domain. It is envisaged that study of the the c-Myb leucine zipper-binding proteins, p160 and p67, should contribute to a further understanding of how gene-specific transcriptional regulation by c-Myb is achieved. Furthermore, the activities of p160 probably extend beyond those associated with c-Myb, to include other transcription factors. These exciting prospects await investigation.

APPENDIX

This appendix contains a manuscript entitled 'Molecular cloning reveals that the p160 Myb-binding protein is a novel, predominantly nucleolar protein which may play a role in transactivation by Myb', by Tavner *et al*. This manuscript is currently in press in the journal 'Molecular and Cellular Biology', and will be published in the February 1998 issue; Vol. 18. No. 2.

MOLECULAR CLONING REVEALS THAT THE p160 MYB-BINDING PROTEIN IS A NOVEL, PREDOMINANTLY NUCLEOLAR PROTEIN WHICH MAY PLAY A ROLE IN TRANSACTIVATION BY MYB

Fiona J Tavner[1], Richard Simpson[2], Shigeki Tashiro[3], Diane Favier[1], Nancy A Jenkins[4], Debra J Gilbert[4], Neal G Copeland[4], Elizabeth M Macmillan[1], Jodi Lutwyche[1], Rebecca A Keough[1], Shunsuke Ishii[3] and Thomas J Gonda[1]*

[1]Hanson Centre for Cancer Research, Institute of Medical and Veterinary Science, Frome Road, Adelaide, South Australia 5000, Australia.
[2]Joint Protein Structure Laboratory of the Ludwig Institute for Cancer Research and the Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria 3052, Australia.
[3]Laboratory of Molecular Genetics, Tsukuba Life Science Center, The Institute of Physical and Chemical Research (RIKEN) Tsukuba, Ibaraki 305, Japan.
[4]Mammalian Genetics Laboratory, ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, Frederick, Maryland 21702, USA.

* To whom correspondence should be addressed at:
        Telephone: +61-8-8222-3305
        FAX: +61-8-8232-4092
        E-mail: Tom.GONDA@imvs.sa.gov.au

ABSTRACT

We have previously detected two related murine nuclear proteins, p160 and p67, that can bind to the leucine zipper motif within the negative regulatory domain of the Myb transcription factor. We now describe the molecular cloning of cDNA corresponding to murine p160. The *P160* gene is located on mouse chromosome 11, and related sequences are found on chromosomes 1 and 12. The predicted p160 protein is novel, and in agreement with previous studies, we find the corresponding 4.5kb mRNA is ubiquitously expressed. We show that p67 is an N-terminal fragment of p160 which is generated by proteolytic cleavage in certain cell types. The protein encoded by the cloned p160 cDNA, and an engineered protein (p67*) comprising the amino-terminal region of p160, exhibit binding specificities for the Myb and Jun leucine zipper regions identical to those of endogenous p160 and p67, respectively. This implies that the Myb binding site of p160 lies within the N-terminal 580 residues and that the Jun binding site is C-terminal of this position. Moreover, we show that p67*, but not p160, can inhibit transactivation by Myb. Unexpectedly, immunofluorescence studies show that p160 is predominantly localized in the nucleolus. The implications of these results for possible functions of p160 are discussed.

INTRODUCTION

It is becoming increasingly clear that c-*myb* plays an essential role in controlling the proliferation and differentiation of hemopoietic cells. This was first suggested on the basis of its preferential expression in immature hemopoietic cells and the subsequent decrease in expression on differentiation (25,70). Confirmation has been provided by more recent loss-of-function studies employing targeted disruption of c-*myb*, which results in the abrupt failure of development of the fetal hemopoietic system (50), and by the ability of anti-sense oligonucleotides to inhibit proliferation of both normal and transformed hemopoietic cells (1,21). Furthermore, enforced expression of activated (23,28) and normal (18) forms of c-*myb* can transform hemopoietic cells *in vitro* (but not, in general, other cell types), and inhibit the induced differentiation of certain leukemic cell lines (8,11,72). Taken together these data suggest that one major function of c-*myb* is to maintain the proliferative state and immature characteristics of early hemopoietic cells.

The proteins encoded by normal and oncogenically activated *myb* genes (Myb) are transcription factors, ie they bind to specific DNA sequences (7) and can enhance

109

transcription of genes and reporter constructs carrying Myb binding sites (53,54,71). These functions are also essential for the ability of *myb* oncogenes to transform hemopoietic cells(34,43). Oncogenically activated forms of Myb differ from normal c-Myb in that they are truncated either at their amino termini, their carboxyl termini or both (22,64). Carboxyl truncation activates c-Myb by disrupting or deleting a region - termed the negative regulatory domain (NRD) - which appears to down-modulate transactivation, DNA binding and transformation (15,34,58,62).

One significant clue as to how the NRD exerts its effects on Myb function comes from the observation that the NRD contains a leucine zipper-like motif (7) and that disruption of this motif by point mutations enhances transactivation and transformation (38). Because leucine zippers generally mediate protein-protein interactions, it seems likely that the Myb leucine zipper promotes association between c-Myb and another protein which inhibits Myb function. There is ample precedent for the existence of protein inhibitors of transcription factors, including Id (6), and IκB (5), which antagonize the function of MyoD and NF-κB, respectively. Alternatively, the Myb inhibitory protein could be c-Myb itself, since the leucine zipper is capable of mediating homodimerization and since Myb homodimers are ineffective in DNA binding or transactivation (55). In this scenario, dimerization may be modulated by competition with another protein capable of forming heterodimers; this latter protein would then function as an activator of Myb. In either case, understanding the regulation of Myb activity would clearly be aided by the identification and characterisation of proteins which interact with the c-Myb leucine zipper.

We have previously described two murine proteins, termed p67 and p160 (17), that can bind to the c-Myb leucine zipper. These were identified by using a bacterially-expressed fusion protein containing the Myb leucine zipper region as an affinity reagent to capture proteins from radiolabelled nuclear extracts. Their specificity was demonstrated by the observation that they do not bind to similar fusion proteins in which two of the critical leucine residues were replaced with proline or alanine residues. Although peptide mapping revealed that p67 and p160 are closely related, there are (at least) two important differences. Firstly, p160, but not p67, can also bind to the c-Jun basic/leucine zipper (bZip) region, suggesting that it may be involved in the regulation of other transcription factors in addition to c-Myb. Secondly, p160 is expressed in all the murine cell lines we have studied to date, whereas p67 was found only in a subset of early myeloid lines (17).

In the present work, we report the molecular cloning of cDNA sequences corresponding to murine p160. The predicted amino acid sequence indicates that p160 appears to be a novel protein. We have used the p160 cDNA to examine the relationship between p160 and p67; we find that p67 represents the amino (N)-terminal region of p160 and is generated by proteolytic cleavage. We further demonstrate that the cloned p160 can specifically associate with Myb, and that a truncated form of p160, which retains its only N-terminal region, can inhibit transactivation by Myb. Immunofluorescence studies of the p160 protein show, surprisingly, that most of the protein is present in the nucleolus. This finding draws attention to some interesting parallels with other nucleolar proteins and, together with our other findings, suggests possible functions for p160 and p67.

MATERIALS AND METHODS

Purification of p67

Nuclear extracts were prepared from approximately 1×10^{10} FDC-P1 cells using a scaled-up version of the protocol described in (17); see also (24). FDC-P1 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS) and 50 units/ml mouse GM-CSF (kindly supplied by Dr Tracy Wilson, Walter and Eliza Hall Institute, Melbourne), to a density of $1-2 \times 10^6$ cells/ml. The nuclear extract was divided into two 50ml aliquots and precleared successively with 2ml glutathione sepharose (Pharmacia), 1.25ml GST-sepharose (~8mg protein) and 1.25ml GST-L34P Myb-sepharose (17)(~2mg protein), each rotating for 1 hour at 4°C. The precleared supernatant was incubated with 1.5ml GST-NRD2-sepharose (17)(~2mg protein) as before in order to isolate p67. The sepharose matrix and bound proteins were washed five times with 10ml cold NETN buffer (20mM TRIS-HCl pH 8.0, 100mM NaCl, 1mM EDTA, 0.5% NP-40) and twice with 10ml cold phosphate buffered saline (PBS) and 0.5% Triton X-100. Bound proteins were eluted with two washes of 1.5ml DOC buffer (100mM NaCl, 50mM TRIS-HCl pH 8.0, 1% Deoxycholic acid, 1% NP-40, 5mM DTT) and concentrated (~100μl) using a centricon-30 column (Amicon). The concentrated protein was separated by electrophoresis in an 8% SDS-polyacrylamide gel, and stained with CBR-250 to visualise proteins. The p67 protein band was excised from the gel and subjected to in-gel proteolysis.

In-gel protein digestion and amino acid sequence analysis

In-gel proteolysis of acrylamide gel resolved proteins was performed essentially as described (49,61). Tryptic peptides extracted from acrylamide gel pieces by trifluoroacetic acid/acetonitrile were fractionated by rapid RP-HPLC on a Brownlee RP-300 cartridge (2.1mm x 100mm, Applied Biosystems, Foster City, CA) employing a Hewlett-Packard 1090A liquid chromatograph fitted with a model 1040A diode-array detector as described previously (49). N-terminal amino acid sequence analysis of peptides was performed by automated Edman degradation using a Hewlett-Packard (model G1005A) protein sequencer operating with the routine 1.3 sequencer program. A Hewlett-Packard model HP1090M liquid chromatograph was used for phenylthiolhydantoin (PTH) amino acid analysis (49). The following five peptide sequences, listed using the single-letter code, were obtained: 1. PDQETRLAT(E); 2. VILRSPK; 3. VLEEG(STGR); 4. VDHLHLEK; 5. ATPQIPETK; the residues shown in parentheses were ambiguous, and in fact are not found in the predicted p67/p160 sequence (see Results and Fig.1).

Degenerate oligonucleotides and library screening

Degenerate oligonucleotides were derived from four of the p67 partial tryptic peptide sequences (see Fig 1B) and used in various combinations to amplify FDC-P1 cDNA by PCR. First strand cDNA was synthesized from FDC-P1 poly A^+ RNA (4μg) using an oligo (dT$_{18}$) primer and SuperScript Reverse Transcriptase (Gibco BRL) as described (60). Degenerate oligonucleotides corresponding to peptide 4: 5'-ATAGGATCCGA(C/T)CA(C/T)(C/T)T(C/T/G)CA(C/T)(C/T)T(C/T/G)GA(A/G)AA-3', incorporating a *Bam*HI site and peptide 5: 5'-AGGAATTCCTT(T/G/A)GT(T/C)TC(T/G/A)GG(G/A/T)AT(T/C)TG(T/G/A)GG-3', incorporating an *Eco*RI site, yielded an amplification product of 114bp using the following PCR conditions: 94°C, 1 minute; 56°C, 2 minutes; 72°C, 2 minutes x 35 cycles. The 114bp product was digested with *Bam*HI/*Eco*RI, cloned into pGEM-4Z (Promega) and sequenced. DNA from a WEHI-3B cDNA Lambda Zap library (Stratagene) was used to obtain a 459bp PCR product using the T3 RNA polymerase promoter sequence and a primer derived from the initial 114bp cDNA sequence: 5'-ATAGGATCCTCGTCTTAAAGAAGGCATGG-3' (94°C, 30 seconds; 58°C, 2 minutes; 72°C, 3 minutes x 35 cycles). The 459 bp product was digested with *Bam*HI, cloned into pBluescript SK (Stratagene) and verified by DNA sequencing.

The same WEHI-3B cDNA Lambda Zap library (Stratagene) was screened by plaque hybridisation with the 459bp cDNA sequence according to the manufacturer's instructions. A positive plaque was identified and a 3.1 kbp cDNA clone obtained by excision of pBluescript SK from the UNI-ZAP XR vector according to the manufacturer's instructions. This cDNA clone was sequenced by first generating a unidirectional deletion series using the Erase-a-Base system (Promega) according to the manufacturer's instructions, then sequencing several of the deleted clones.

5' RACE PCR

5' RACE PCR (19) was performed using the 5'-AmpliFINDER kit protocol (Clontech). First strand cDNA was synthesized from FDC-P1 poly A$^+$ RNA (2μg) using a p160 cDNA-derived oligonucleotide:

(5'-TTGTAAAGGCCACCATCATAGTCAACTGCC-3') and SuperScript reverse transcriptase (Gibco BRL) as described (60). 5' RACE PCR was undertaken with a p160 gene-specific primer, incorporating a *Bam*H1 site:

(5'-CGGATCCGGAAGGTACCCACGTATG-3'), and the AmpliFINDER anchor primer (Clontech) using the following conditions: 94°C, 45 seconds; 60°C, 1 minute; 72°C, 2 minutes x 30 cycles. The resulting PCR product was digested with *Eco*R1/*Bam*H1, cloned into pBluescript SK (Stratagene), and verified by DNA sequencing.

Construction of plasmids

A full-length p160 cDNA clone was constructed in pGEM-3Z (Promega) by utilising the unique *Kpn*I site present in both the 5' RACE PCR product and the 3.1 kbp partial p160 cDNA clone (nucleotides 992-4126). Partial p160 cDNA in pBluescript (Stratagene) was digested with *Eco*RI/*Xho*I and subcloned into the *Eco*RI/*Sal*I sites of pGEM-3Z. This construct was subsequently digested with *Eco*RI/*Kpn*I and ligated with the 5' RACE PCR *Eco*RI-*Kpn*I fragment to yield a full-length p160 cDNA clone (pGEM3Z-160).

The p67* cDNA construct was derived from p160 cDNA by introducing a termination codon after nucleotide position 1755 by PCR, in order to yield a protein with a predicted Mr of approximately 67kD. The following PCR primers were used to generate the p67* cDNA clone: 5'-TATCGAATTCGTGACGTGTTTGGCTCAGC-3' and

5'-ACGTGGATCCTCATTCCTTCAGAGTACTC-3', incorporating a *Bam*HI site, (94°C, 45 seconds; 60°C, 1 minute; 72°C, 90 seconds x 30 cycles). The resulting PCR product was digested with *Eco*RI/*Bam*HI and cloned into pGEM-3Z to yield pGEM3Z-67*.

Two GST fusion constructs containing 5' or 3' p160 sequences were made as follows:

The 5' p160 fragment (nucleotides 1-1083) was generated by digestion of pGEM3Z-160 with *Eco*RI/*Kpn*I, blunt-ended, and ligated into the *Sma*I site of pGEX-4T-1 (Pharmacia). The 3' p160 fragment (nucleotides 3152-4101) was generated by digestion of partial p160 cDNA in pBluescript (Stratagene) with *Bam*HI/*Ssp*I, blunt-ended, and ligated into the *Sma*I site of pGEX-3X (Pharmacia). The in-frame GST-p160 sequences were verified by DNA sequencing.

Site directed mutagenesis and construction of expression vectors

The FLAG epitope (DYKDDDDK) was introduced into the amino-termini of p160 and p67* by *in vitro* site directed mutagenesis using the pALTER-1 system (Promega) according to the manufacturer's instructions. The mutagenic oligonucleotides for both p160 and p67* cDNAs incorporated the FLAG octapeptide sequence downstream of a Kozak consensus initiation site, with the original translation initiation site replaced. The presence of the mutagenized sequences was verified by DNA sequencing.

The FLAG epitope tagged p160 and p67* expression plasmids were constructed from the pact-c-*myb* expression vector (described in (54))by first excising the c-*myb* cDNA by digestion with *Nco*I/*Xba*I. This was replaced by an *Nco*I-*Xba*I fragment containing FLAG p67* cDNA to yield pact-67*FLAG. pact-160FLAG was constructed by generating a fragment containing FLAG p160 cDNA by PCR from the pALTER-1-p160FLAG plasmid using the SP6 RNA polymerase promoter sequence and 5'-ACGTACTAGTTCAAGGTGTCTGCACTCTC-3' (94°C, 45 seconds; 50°C, 1 minute; 72°C, 4 minutes), digestion with *Nco*I/*Spe*I, and replacement of the c-*myb* gene between the *Nco*I/*Xba*I sites of pact-c-*myb*.

The retroviral construct containing p160FLAG cDNA was generated by digestion of the pALTER-1-p160FLAG plasmid with *Eco*RI/*Sal*I, blunt-ended, and ligated into the *Hpa*I site of pRUF*neo* (60).

In vitro transcription/translation

Proteins were synthesized *in vitro* using the TNT coupled transcription/translation rabbit reticulocyte lysate kit (Promega). Briefly, 1μg of each pGEM3Z-160 and pGEM3Z-67* was incubated at 30°C for 90 minutes in a 50μl reaction volume containing the recommended quantities of kit reagents, including T7 RNA polymerase and 4μl of translation grade $[^{35}S]$-methionine (10.2mCi/ml; NEN Dupont). Labelled proteins were analysed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and detected by autoradiography using Hyperfilm-MP (Amersham).

Proteolysis by cell extracts

Extracts were made by first suspending cells at 10^8/ml in 25mM HEPES pH 7.4, 1mM EDTA, 0.1% CHAPS, 10% sucrose, 5mM DTT. Cells were lysed with three cycles of freezing in liquid nitrogen and thawing at 37°C, then centrifuged at 1000 rpm for 5min at 4°C. The supernatant was clarified by centrifugation at 13,000 rpm for 15 min and stored at -20°C. Proteolysis reactions were carried out with 10μg of extract, 10μl of *in vitro* translated, ^{35}S-methionine-labelled p160 plus buffer in a total of 30μl and incubated at 37°C for varying periods before termination by adding 2X SDS loading buffer, and analysed by SDS-PAGE.

Southern blotting analysis

Genomic DNA was isolated from FDC-P1 cells as described in (2), with the exception that the DNA was spooled onto a Pasteur pipette rather than pelleted. DNA was digested with restriction enzymes, fractionated by electrophoresis in a 1% agarose gel and transferred to Hybond-N nylon membrane (Amersham) according to the manufacturers' instructions. Southern blots were probed with 5' (nucleotides 1-1084) and 3' (nucleotides 3152-4102) p160 cDNA sequences, labelled with $[\alpha^{32}P]$-dATP using a random primer labelling kit (Amersham). Hybridisations were performed as described (63), at the following stringency: 50% Formamide, 5 X SSC, 42°C. Blots were washed with 2 X SSC, 0.1% SDS at 60°C.

Chromosomal localisation

Interspecific mouse backcross mapping was used to determine the chromosomal locations of p160 sequences. Backcross progeny were generated by mating (C57BL/6J x *M.*

Spretus)F_1 females and C57BL/6J males as described (12). A total of 205 N_2 mice were used to map the *p160* loci (see Fig 2B). DNA isolation, restriction enzyme digestion, agarose gel electrophoresis, Southern blot transfer and hybridisation were performed essentially as described (37). All blots were prepared with Zetabind nylon membrane (AMF Cuno). The 3' p160 cDNA probe used is the same as that described above, under 'Southern blotting analysis', and was labelled with [α^{32}P]-dCTP using a nick translation labelling kit (Boehringer Mannheim). Washing was done to a final stringency of 0.8 X SSCP, 0.1% SDS, 65°C. Fragments of 19.0 and 5.0 kb were detected in *Hind*III digested C57BL/6J DNA and fragments of 15.0, 12.0, 7.3, and 5.0 kb were detected in *Hind*III digested *M. Spretus* DNA. The presence or absence of the 15.0, 12.0 and 7.3 kb *Hind*III *M. Spretus*-specific fragments, which segregated independently, was followed in the backcross mice.

A description of the probes and restriction fragment length polymorphisms (RFLPs) for the loci used to position the *p160* loci in the interspecific backcross has been reported. These include: *Ren1* and *Fasl*, chromosome 1 (66); *Myhsf1*, *Glut4*, and *Nf1*, chromosome 11 (9,32); and *Hfhbf1* and *Sos2*, chromosome 12 (3). Recombination distances were calculated as described (29) using the computer program SPRETUS MADNESS. Gene order was determined by minimising the number of recombination events required to explain the allele distribution patterns.

RNA isolation and Northern blotting analysis

Total RNA was isolated from adult mouse tissues or e14 foetuses as described in (10). Poly A^+ RNA was isolated directly from cultured cells as described previously in (26). Total RNA (30μg) or Poly A^+ RNA (2μg) was fractionated by formaldehyde-agarose gel electrophoresis (2). RNA was then transferred to Hybond-N nylon membrane (Amersham) according to the manufacturer's instructions. Northern blots were probed with a 1.1 kb fragment representing the 5' end of the p160 cDNA (nucleotides 1-1083). For probe labelling, hybridisation and washing conditions see 'Southern blotting analysis'. Blots were stripped and reprobed with a murine glyceraldehyde phosphate dehydrogenase (GAPDH) cDNA fragment.

Cell transfection and infection

The 293T cell line, a derivative of the human embryonic kidney 293 cell line (16), was used to express p160FLAG, p67*FLAG and/or Myb. 293T cells were grown in DME medium supplemented with 10% fetal calf serum (FCS). Prior to transfection, cells were seeded at 1.4×10^6 in 60mm culture dishes and left to adhere overnight, after which the medium was changed. Four hours after medium change, calcium phosphate precipitates prepared as described in (48), containing 7μg of pact-160FLAG or pact-67*FLAG DNA were added for 4 hours, after which the cells were trypsinized and replated onto 100mm culture dishes. After incubation for 40 to 48 hours, total cell extracts were prepared. Co-transfections of pact-p160FLAG and pact-WTMyb or L3,4P Myb (38) were performed similarly except that 3 μg of each DNA were used, with the total amount of DNA transfected made up to 10μg with pact vector.

Infection of NIH3T3 cells was performed essentially as described by Pear et al.(57). Briefly, BOSC 23 retroviral packaging cells were transiently transfected with 10μg pRUFneo-160FLAG, using the calcium phosphate procedure. After 24 hours, supernatants were collected and used to infect NIH3T3 cells seeded the previous day at 3×10^5 cells/60mm dish. After an additional 24 hours, cells were selected with G418 at 400μg/ml.

Preparation of cell extracts

Total cell extracts for GST-fusion protein binding assays were prepared by lysing the cells in 450μl NP-40 lysis buffer (10mM TRIS-HCl, pH 8.0, 25mM KCl, 5mM MgCl$_2$, 0.5% NP-40, 2mM PMSF, 40μl/ml Protease Inhibitor Cocktail (Boehringer Mannheim), and 5mM DTT) plus 500mM NaCl. Extracts were incubated on ice for 10 minutes then centrifuged at 45000 rpm (TLA-45 rotor; Beckman TL-100 ultracentrifuge) for 20 minutes, the supernatants removed, aliquotted and stored at -70°C. Nuclear and cytoplasmic extracts for immunoblotting were prepared from FDC-P1 and NIH3T3 cells as described (17). Total cell extracts for immunoprecipitations were prepared essentially as described by Harlow et al. (30), with the addition of 10% glycerol.

Binding reactions and immunoprecipitation

293T total cell extracts containing overexpressed p160FLAG and p67*FLAG were used in binding reactions with GST fusion proteins. These reactions were performed essentially as described (17). Similarly, for co-immunoprecipitation experiments, 293T

cells were transfected with pact vectors encoding p160FLAG, WT Myb or L3,4P Myb as appropriate. Immunoprecipitation from total cell extracts was performed essentially as described by Harlow et al. (30), except that the washing buffer contained 143mM NaCl. Proteins were analysed by immunoblotting with the M2 FLAG monoclonal antibody (IBI) or with anti-Myb monoclonal antibody 1.1 (59). Control immunoprecipitations were carried out with irrelevant monoclonal antibodies.

Antibodies and Immunoblotting

After SDS-PAGE, proteins were transferred to a nitrocellulose membrane (Schleicher and Schuell) and blocked in TBS-T (50mM TRIS-HCl pH 7.4, 135mM NaCl, 0.1% Tween 20) containing 5% skim milk. Epitope tagged proteins were detected using the M2 FLAG monoclonal antibody (1:500; IBI), peroxidase-linked anti-mouse Ig (1:1000; Amersham) and Enhanced Chemiluminescence (ECL) (Pierce).

The GST-p160 5' and 3' fusion constructs (see 'Construction of plasmids') were used to express amino-terminal and carboxyl-terminal p160 fusion proteins, respectively, and purified essentially as described previously (17). The purified GST fusion proteins, containing p160 amino-terminal residues 1-326 or carboxyl-terminal residues 1046-1344, were used to raise polyclonal antisera in rabbits, termed 160N and 160C, respectively. The p160 antisera was used at 1:200 (160N) or 1:500 (160C) for immunoblotting with peroxidase-linked anti-rabbit Ig (1:1000; Amersham) and ECL (Pierce).

Immunofluorescence

NIH3T3 cells stably expressing pRUFneo-160FLAG or uninfected were seeded at 3 x 10^4 cells/well of an 8-well chamber glass slide and left to adhere overnight. After washing with PBS, the cells were fixed with ice-cold 50% (v/v) methanol/acetone for 30 seconds and washed with cold PBS containing 1% BSA. Cells were incubated on ice with primary antibody (anti-FLAG M2 monoclonal (IBI) at 1:300 or carboxyl-terminal p160 polyclonal antiserum; α160C (described under 'Antibodies and immunoblotting') at 1:300) for 45 minutes. FITC-conjugated secondary antibodies were incubated with the cells on ice for 30 minutes at the following dilutions: 1:160 for affinity purified anti-mouse IgG (Silenus) and 1:80 for sheep anti-rabbit Ig F(ab')₂ (Silenus). Cells were mounted with 1% propylgallate/86% glycerol and viewed using a BioRad MRC 600 confocal microscope. Digitized images were manipulated using Adobe Photoshop software.

Transactivation analysis

Myb transactivation experiments were performed essentially as described (54). Briefly, CV-1 cells were cotransfected with 4µg of the CAT reporter plasmid, p*myc*CAT (51), 4µg of each of the effector plasmids pact-c-*myb* or pact-c-*myb*L34P (38), various amounts (0-5µg) of pact-160FLAG or pact-67*FLAG, and 1µg of the internal control plasmid pact-β-gal (44). After 48 hours, cell extracts were normalized for β-galactosidase activity and assayed for CAT activity.

Sequence data

The sequence reported here has been submitted to the Genbank database under the accession number U63648.

RESULTS

Isolation of a p160 cDNA clone

We have described the detection of p67 and p160 as radiolabelled nuclear proteins that bound to a bacterially-expressed glutathione-S-transferase (GST) fusion protein containing the wild-type Myb leucine zipper region, but that failed to bind to similar fusion proteins bearing a mutated leucine zipper (17). By scaling up and slightly modifying this procedure we were able to isolate sufficient p67 from FDC-P1 cells to allow the generation, purification and partial amino acid sequencing of five tryptic peptides (see Materials and Methods and (24)). Comparison of these sequences with protein and DNA sequence databases indicated that they represented a novel protein. Degenerate oligonucleotides corresponding to four of these peptides were used in various combinations as PCR primers for the amplification of potential p67 coding sequences from FDC-P1 cDNA (Fig. 1A; see also Materials and Methods). Nucleotide sequencing of a number of the resultant products identified a 114 bp fragment, generated with primers corresponding to peptides 4 and 5, that contained an open reading frame. This sequence encoded residues from peptide 4 that were not present in the corresponding PCR primer, and also placed a lysine residue immediately preceding the N-terminus of peptide 5, which is consistent with its generation by trypsin cleavage.

An oligonucleotide from within the 114 bp fragment was then used as a PCR primer to obtain a 459 bp sequence from a WEHI-3B cDNA library. This fragment also

contained an open reading frame and was used to screen the same library by hybridisation. One clone containing a 3.1 kbp insert was isolated, and the insert sequenced. This clone contained an open reading frame of 1010 amino acids which started with a good match (nucleotides 1115-1121) to the Kozak consensus (A/GNNATGG (40)). However, the predicted molecular mass of 114 kD and the *in vitro* translation product of ~ 120 kD (data not shown) were both much larger than p67, but smaller than p160. We therefore explored the possibility that this clone represented a partial cDNA clone of the p160 mRNA by performing 5' RACE PCR (19) to extend the cloned sequences. This resulted in an additional ~ 1 kb of sequence being identified which, when joined to the 3.1 kb clone, could encode a predicted protein of 152 kD, again with sequence surrounding the initiation codon (nucleotides 13-19) conforming to the Kozak consensus. The relationship between the various clones described above is illustrated in Fig. 1A, and the nucleotide and predicted amino acid sequences are shown in Fig. 1B. (These sequence data have been deposited in the Genbank database under accession number U63648.) The sequence shown in Fig. 1B may include the 3' terminus of the mRNA since a consensus polyadenlylation signal (AATAAA; beginning at nucleotide 4106) is present 15 nucleotides before the end of the cloned sequence. To confirm that the complete sequence does encode p160, the joined clones were translated *in vitro* and compared to p160 (and p67) from FDC-P1 nuclear extracts. The *in vitro*-translated protein co-migrated with *in vivo*-labelled p160 on an SDS-polyacrylamide gel, and the two proteins showed similar peptide patterns on V8 protease digestion (data not shown).

Features of the predicted p160 sequence

It is apparent from the predicted p160 amino acid sequence (Fig. 1B) that all five p67 tryptic peptides are present in the predicted p160 protein sequence. Inspection of this predicted sequence also indicated the presence of a highly acidic region in the middle of the protein and seven short, highly basic sequences embedded in a proline-serine rich region at the carboxyl terminus (Fig. 1B). These basic repeats conform to a consensus Lys-Lys-Xaa-Xaa-Lys, although some repeats contain arginine instead of lysine at one position, and the "Xaa" residues are also frequently basic. Searches for functional motifs indicated the presence of numerous potential phosphorylation sites (not shown), including those for casein kinase II (S/T X X E/D; single letter code), protein kinase C (S/T X R/K), p34[cdc2] (S/T P X K/R) (reviewed in (39)) and MAP kinase (P X T/S P; (27)). Because p160 and

p67 bind to Myb via the latter's leucine zipper region (17), we particularly looked for potential leucine zippers within the predicted p160 sequence. Two regions - residues 307 to 335 and 900 to 928 respectively - resemble sequences that could conceivably function as leucine zippers (see Discussion). In addition, p160 contains several copies of a recently-described motif termed a "leucine charged domain" (LCD) that is involved in interactions between nuclear hormone receptors and co-activators, possibly by forming a short amphipathic helix (31,68). Five copies of the core motif (LXXLL; single letter code) are present as well as four possible copies in which one of the last two leucines is replaced by another aliphatic amino acid (Fig. 1B).

Searching of nucleic acid and protein sequence databases revealed that several partial p160 cDNAs have been obtained by others using essentially random sequencing approaches (ie as expressed sequence tags). In addition, the deduced sequence of a protein which is almost certainly the rat homologue of p160 has recently appeared in the Genpept database (accession number U83590). This sequence corresponds to a protein termed "PIP" which was apparently identified as interacting with PAR transcription factors (see Discussion). Other than this, homology to sequences of proteins with known functions is very limited. The highest scoring sequences are those of the *Drosophila* lethal 1B(i) protein (Genbank accession number U20542; (69)) and a hypothetical yeast protein of unknown function (Yelo55cp: Genpept accession number U18795). Most of the other (relatively) high-scoring matches, including upstream binding factor (UBF (36)), show homology mainly within the acidic region of p160, suggesting that the database search has preferentially identified a subset of proteins with highly acidic regions (see Discussion).

Chromosomal localisation of P160 and related sequences

As the p160 cDNA sequence appeared to represent a novel gene, it was of interest to determine its chromosomal location and thus to determine whether it maps to a locus corresponding to a previously-identified mouse mutant. This in turn could provide a clue to the function of the gene. Initial Southern analysis of genomic p160 sequences carried out with probes corresponding to both the 5' and 3' ends of the p160 cDNA (Fig. 2A) detected three or more hybridising bands with each probe, suggesting that there may be more than one p160-related gene.

Chromosomal localisation was carried out by Southern hybridisation of the 3' probe to DNA from a panel of interspecific (*Mus musculus* x *M. spretus*) backcrosses (12).

Restriction fragment length polymorphisms (RFLPs) between the two species were used to follow the segregation of p160-related *M. spretus* sequences in the backcross mice. The results of this analysis, illustrated in Fig. 2B, indicate that there are three loci which contain p160-related sequences on chromosomes 1, 11 and 12 respectively. Although three (or more) fragments could be detected with the 5' probe with most restriction enzymes (see also Fig. 2A), only one RFLP could be identified, which allowed the corresponding fragment to be mapped to chromosome 11 (data not shown). As this band did not show recombination with the chromosome 11 fragment detected with the 3' probe (data not shown), both fragments presumably represent the same locus. Moreover, we have found that an intron probe, derived from genomic clones that correspond in sequence to at least 2067 bp of the p160 cDNA, also maps to chromosome 11 and co-segregates with the locus identified with the other probes (data not shown). We have therefore termed the chromosome 11 locus "*P160*", as it appears to represent the entire p160 cDNA sequence, and have termed the loci on chromosomes 1 and 12 "*P160-rs1*" and "*P160-rs2*" ["*rs*" = "related sequence"], respectively. The lack of further RFLPs with the 5' probe did not allow us to determine whether or not the *P160-rs1* and *P160-rs2* loci also contain 5' p160-related sequences.

Expression of p160 mRNA

The cloned p160 cDNA sequence was used to identify and examine the expression of p160 mRNA. Northern blotting of polyadenylated RNA from a number of cell lines with a probe corresponding to the 5'-most 1 kb of p160 (Fig. 3A) or with the entire p160 cDNA (data not shown) revealed a single mRNA species of ~ 4.5 kb in each murine cell line. The size of this transcript is clearly adequate to include all of the cloned sequences (4.1 kb) and a poly(A) tract. Furthermore, the ubiquitous expression of the 4.5 kb mRNA corresponds to the previously determined distribution of p160 (17). Importantly, no additional mRNA species were present in cells (such as FDC-P1 and WEHI-3B) that also expressed p67 (17). Thus, there does not appear to be a separate transcript that could encode p67 (see below). Note also that the p160 cDNA did not detect a transcript in RNA from the human HL-60 cell line; this is consistent with our previous failure to detect p160 or p67 in human cells (including HL-60) by our standard binding assay (17).

We also examined the expression of p160 mRNA in a number of different tissues and organs of the mouse. Fig 3B shows that p160 is expressed in most tissues and organs;

moreover, in other experiments, expression was detected in spleen, liver, intestine and lung (data not shown). Interestingly, the level in fetal liver is particularly high; this probably does not indicate preferential expression in hemopoietic tissue as expression in bone marrow is much lower. Moreover, expression is not substantially greater in hemopoietic cell lines than in fibroblasts (Fig. 3A).

Generation of antisera that recognize p160 and p67

To generate a reagent for the detection of p160 (and p67), antisera were raised against GST fusion proteins containing the N-terminal 326 residues and C-terminal 298 residues of p160, respectively. These were then used to probe nuclear and cytoplasmic extracts of FDC-P1 and NIH3T3 cells. Fig 4 shows that, as expected from our previous work (17), the N-terminal antiserum (160N) recognized proteins of 160kD and 67kD in nuclear, but not cytoplasmic extracts from FDC-P1 cells. Also as expected from our previous report (17), p160 but no p67 was detected in NIH3T3 nuclear extracts. In contrast, the C-terminal antiserum (160C) detected only p160 in nuclear (but not cytoplasmic) extracts from both cell types. Moreover, each antiserum detected these same proteins following binding of nuclear extracts to GST fusion proteins containing the Myb NRD (data not shown). Other bands (of ~ 97 and ~110 kD) detected by the antisera may be due to non-specific reactivities; however it is also possible that they represent other p160-related proteins.

p67 represents an N-terminal proteolytic cleavage product of p160

The clones described here (Fig. 1) were isolated on the basis of peptide sequences derived from p67, but in fact encode p160 (see also below). While this is compatible with the observation that the two proteins are related, as judged by V8 protease peptide mapping (17), it is apparent from the sequence that the N-terminal portion of p160 contains all five of the tryptic peptides that were derived from p67 (see Fig. 1B). Moreover, we show above that the 160N antiserum but not 160C also reacts with p67 (Fig. 4). These observations led us to suspect that p67 is the N-terminal proteolytic cleavage product of p160. The existence of a single mRNA species corresponding to p160 in cells such as FDC-P1 and WEHI-3B that express both p160 and p67 (Fig. 3A) is also consistent with this notion; if p67 were derived from an alternatively-spliced transcript or another gene, we might expect to find an additional mRNA species.

If p67 is generated by proteolysis, cells in which p67 is detected should contain a proteolytic activity capable of cleaving p160, while this activity should be absent from cells that do not contain detectable p67. This prediction was tested by incubating radiolabelled *in vitro* translated p160 with FDC-P1 and NIH3T3 cell extracts. Fig. 5 shows that the FDC-P1 extract generated a number of cleavage products, including one of apparent Mr ~67kD. Moreover, this protein represents the N-terminal portion of p160 since it could be immunoprecipitated with the 160N antiserum described above (data not shown). In contrast, the NIH3T3 extract failed to cleave p160 even after extended incubation (Fig 5).

Binding properties of full length and truncated forms of p160

In order to examine the binding properties of the protein(s) encoded by the p160 cDNA clone, we first constructed a carboxyl-truncated form of p160, which we term p67*, by inserting a termination codon after nucleotide 1755 (see Fig. 1B), resulting in a 580 amino acid protein of predicted Mr ~ 67 kD. (This construct was initially made to approximate the endogenous p67 protein). Epitope-tagged versions of p160 and p67* were then constructed by adding a sequence encoding the FLAG epitope (33) to the 5' ends of each cDNA construct. These constructs were inserted into the pact expression vector (54) and transiently transfected into human embryonic kidney 293T cells. Cell extracts were then analysed directly by immunoblotting with anti-FLAG antibodies (Fig. 6A) or following binding reactions with GST fusion proteins containing Myb or Jun leucine zipper regions (Fig. 6B). In agreement with the properties of endogenous p160 and p67 (17), FLAG-tagged p160 and p67* both bound to the GST-NRD2 protein, which contains a wild-type Myb leucine zipper region, but not to the GST-L3,4P protein (17) which has a mutated leucine zipper. Interestingly, it appears that p67* binds more efficiently than p160 to GST-NRD2 (compare "bound" and "unbound" lanes in Fig. 6B). Also in agreement with studies of the endogenous proteins, p160FLAG bound to the GST-Jun protein which carries the leucine zipper and basic regions of c-Jun, whereas p67*FLAG did not. These observations further validate the authenticity of the p160 cDNA clone, and show that a truncated version of p160 - ie p67* - has similar binding properties to p67.

To determine whether full-length p160 and c-Myb proteins could interact in the intracellular milieu, we utilized expression vectors encoding wild-type c-Myb or the L3,4P leucine zipper mutant and p160FLAG to co-express these proteins in 293T cells. Cell extracts were then subjected to immunoprecipitation with anti-FLAG antibodies, and the

immunoprecipitates analysed by immunoblotting with anti-FLAG and anti-Myb antibodies. Fig. 7 shows that anti-FLAG immunoprecipitates from cells expressing both p160 and c-Myb contained WT c-Myb protein, whilst those from transfections of p160FLAG or Myb alone did not. Neither protein was precipitated when an irrelevant control antibody was used (data not shown). Moreover, Fig. 7 shows that, in contrast to WT c-Myb, the leucine zipper mutant did not interact specifically with p160FLAG. The small amount of Myb detected in the immunoprecipitate from the p160FLAG plus L3,4P Myb co-transfection probably reflects a degree of "stickiness" of the latter protein, since a similar amount was observed in precipitates from cells transfected with L3,4P Myb alone). These data confirm that full-length c-Myb and p160 can indeed associate specifically *in vivo* and that their association requires the c-Myb leucine zipper motif.

P67* but not p160 inhibits transactivation by Myb

As the original rationale for isolating proteins that interact with the Myb leucine zipper was that such proteins may be inhibitors of Myb function, we examined the effects of p160 on one of Myb's key activities - its ability to enhance transcription from promoters containing Myb binding sites. p67* was also tested in this assay since although it can interact with the Myb leucine zipper (Fig. 6), it presumably lacks other functional domains of p160, and thus may have had the potential to act as a "dominant negative" antagonist of p160 function. For this purpose we used the c-*myc* promoter region coupled to the chloramphenicol acetyltransferase (CAT) reporter gene (51,54). This construct was co-transfected into CV-1 cells along with expression vectors for either wild-type or leucine zipper-mutated (L3,4P) Myb and with various amounts of p67* or p160 expression vector. Note that as expected (38) a greater degree of transactivation was observed when cells were transfected with the leucine zipper-mutated Myb (Fig. 8A). Fig. 8B shows that increasing amounts of p67* substantially inhibited transactivation by WT Myb, down to a level of just 12% of that obtained in the absence of p67*. On the other hand, p160 had relatively little effect, giving a slight stimulation at low levels and a slight (25%) inhibition at the highest level. This is unlikely to reflect cleavage of p67 from transfected p160 because p67, and the proteolytic activity that generates it, appears to be confined to myeloid hemopoietic cells ((17) and Fig. 5). Thus, for example, transfection of p160FLAG into 293T or NIH3T3 cells does not give rise to detectable p67 (Fig. 6A and data not shown). Most of the inhibition by p67* was dependent on the Myb leucine zipper, although nearly 50%

inhibition of transactivation by L3,4P Myb was obtained with the highest level of p67*. Again, p160 had little effect on transactivation by the mutant form of Myb (Fig.8). These results are unlikely to reflect a non-specific suppression of transcriptional activity because expression of the control reporter construct (pact-β-galactosidase) showed no inhibition by p67* or p160 (data not shown).

p160 is a predominantly nucleolar protein

The subcellular localisation of p160 was further examined using the p160 antisera described above in immunofluorescence studies. NIH3T3 fibroblasts were used for these studies because our antisera were raised against murine p160 and because the morphology of these cells allowed clear visualisation of the various subcellular compartments. To our surprise, staining with the 160C antiserum revealed that p160 is located predominantly in the nucleolus (Fig. 9B), although significantly (see Discussion), some nucleoplasmic staining was also evident. To independently confirm this observation we examined NIH3T3 fibroblasts that had been infected with a retroviral vector encoding p160FLAG. Again, staining with monoclonal anti-FLAG antibody showed predominantly nucleolar localisation (Fig. 9D). The localization of p67* could not be determined by immunofluorescence in similar studies because we have been unable to stably express p67* in NIH3T3 or other suitable cells.

DISCUSSION

Features of the p160 protein

We report here the molecular cloning of cDNA corresponding to p160, one of two nuclear proteins that were identified by their ability to bind to the c-Myb negative regulatory domain and specifically, to the leucine zipper motif within this domain (17). Although the p160 cDNA was isolated on the basis of peptide sequences derived from p67, evidence based on size, peptide mapping, mRNA expression pattern and the binding specificity of its translation product, demonstrates that the cDNA represents p160.

The predicted p160 amino acid sequence is novel and shows no close homology to other sequences present in databases other than the presumptive rat homologue (Genpept accession no. U83590); this sequence may represent a partial clone since it lacks the first 67 residues of the mouse protein . Many of the other higher-scoring matches show

homology only to the acidic region of the protein (data not shown), which casts some doubt on the biological significance of such similarities. However, the similarities between p160 and the highest-scoring match, the product of the Drosophila lethal (1) 1Bi gene, extend beyond this region and may be of more significance (see below).

Inspection of the predicted p160 sequence reveals some potentially interesting features (see Results and Fig. 1). As mentioned above, there are two leucine zipper-like sequences, starting at amino acids 307 and 900, respectively, with the following residues in a possible heptad repeat: L-L-L-I-M and L-L-Q-L-L (single letter code). However, mutagenesis of this region, which replaced two leucine residues in each of these sequences with prolines, had no effect on the ability of p160 to bind to the Myb or Jun leucine zipper regions (R.A.K and T.J.G., unpublished observations). Thus the structures that mediate the interaction of p160 with Myb and Jun remain to be identified; one intriguing possibility is that these interactions involve some of the LCD motifs. Our results do show, however, that the Myb leucine zipper- and Jun bZip domain-binding regions of p160 are distinct. The Myb binding region must lie within the first 580 amino acids since p67*, like endogenous p67, efficiently and specifically binds to the Myb leucine zipper; similarly, the Jun binding region is carboxy-terminal of residue 580 because p67* fails to bind the Jun leucine zipper region. The functions of the other structurally distinguishable regions of p160 ie the highly acidic region in the centre of the protein and the basic repeats at the C-terminus, are currently uncertain. It is possible that the acidic region indicates a function in transcriptional regulation, since many transcription factors have acidic transactivation domains (65). It is also interesting to note that the basic repeats bear a close similarity to nuclear localization signals (NLS) (reviewed in (20,35)) and indeed, preliminary data suggest that their deletion results in the cytoplasmic localization of at least a proportion of the protein, which as we show here (Fig. 4) is normally exclusively nuclear (J. Gardner, F.J.T., J.L. and T.J.G., unpublished observations).

Relationship between p160 and p67

Earlier data clearly showed that p160 and p67 are functionally and structurally related since they both bind the c-Myb leucine zipper and share many peptides generated by partial digestion with V8 protease (17). The present work extends those results, and provides several lines of evidence that p67 is generated by proteolysis of p160. This raises the possibility that, despite its reproducibility and strict cell type specificity, the detection

of p67 in nuclear extracts is a consequence of the nuclear isolation procedure. In fact, little of this protein is detected when whole cells are lysed directly in the presence of SDS (F.J.T. and T.J.G., unpublished observations). Nevertheless, it is conceivable that there may be physiological conditions that also activate or release the protease which cleaves p160. This in turn could have biological consequences as p67, like the p67* construct, may similarly function as an inhibitor of transactivation by Myb. Examination of the effects of various cell treatments, eg growth factor deprivation or induction of differentiation, on the generation of p67 will be required to resolve this issue.

Similarities of p160 with other nucleolar proteins and possible implications for p160 function

An unexpected finding from this study was the nucleolar localisation of p160. Although this was not anticipated for a potential regulator of c-Myb activity, since Myb is not detected in the nucleolus (4), this observation is consistent with the homology between p160 and the product of the Drosophila lethal (1) 1Bi gene ((69); Genbank accession number U20542). Unfortunately, little information is available about the function of this gene, but interestingly, its product is reportedly also a nucleolar protein (cited in the Genbank entry), as is PIP, the presumptive rat homologue of p160 (again, as cited in the Genpept entry). It may also be of significance that one of the other higher-scoring matches from BLAST searches of protein databases is the upstream binding factor, UBF (36). UBF is a polymerase I transcription factor and thus is also a nucleolar protein. As remarked above, homology with p160 is mostly within the acidic region but it is of interest that the acidic region of UBF is essential for its transactivating/anti-repressing capacity and is involved in displacing histones from DNA (41,42). Another nucleolar protein, Nopp140 (45,46) also shows some other general similarities to p160 although there is no significant sequence homology. Like p160, Nopp140 also has an acidic region containing multiple casein kinase II phosphorylation sites and multiple, basic NLS-like sequences toward its C-terminus (46). Although predominantly nucleolar, it was noted that a small proportion of Nopp140 can be detected in the nucleoplasm (46), as also seen here with p160. The significance of this has been highlighted by a recent report (47) that Nopp140 can function as a transcriptional co-activator of the bZip transcription factor C/EBPβ in stimulating transcription of a polymerase II-transcribed gene. Thus, the similarities between p160 and Nopp140 (and between p160 and UBF) raise the possibility that p160 may play a role - eg

as a co-activator - in the transcription of both polymerase I- and some polymerase II-transcribed genes. In this light, the predominantly nucleolar localization of p160 can be reconciled with its interaction with Myb, Jun and possibly and other transcription factors; ie, the small fraction of p160 in the nucleoplasm may be involved in such interactions. This notion is further supported by the reported (ie, as stated in the Genpept entry) identification of the presumptive rat homologue of p160, PIP, by its interaction with members of the PAR family of transcription factors. Moreover, the presence of several LCD motifs, which have recently been reported to mediate the interaction between certain co-activators and nuclear hormone receptors (31,68), is also consistent with a role for p160 in the regulation of polymerase II-mediated transcription.

Effects of p67 and p160 on transactivation by Myb*

Despite the fact that overexpression of p160 had no significant effect on transactivation by Myb, our findings remain consistent with a role for p160 as a potential co-factor in Myb-regulated transcription. It may simply be that ectopically expressed Myb interacts with endogenous p160 in CV-1 cells (as it does with exogenous p160; Fig. 7) and that the levels of endogenous p160 in these cells are sufficiently high as not to be rate-limiting. In this case excess p160 would have little or no effect; a similar observation, ie a lack of effect of overexpression, was made in the case of the co-activator p/CIP (68). We have been unable look for such an interaction directly because our antisera are directed against murine (not simian) p160. On the other hand, inhibition by a truncated form of p160, ie p67*, supports a role for p160 in transcriptional regulation by Myb. One possible mechanism is, as we envisaged, that p67* is acting as a "dominant negative" antagonist of a co-activator function of (endogenous) p160, since it lacks many of the potential functional domains of the latter. Alternatively, truncation may unmask or enhance an inhibitory function of p160 that is normally subject to regulation (eg by phosphorylation or by other protein-protein interactions). In any case, the inhibition of transactivation by p67* almost certainly involves direct interaction with Myb via the latter's leucine zipper motif, as it is largely dependent on the integrity of this motif. The ability of p67* to bind to the Myb NRD in a leucine zipper-dependent manner is also consistent with a direct interaction *in vivo*. However, the precise mechanism by which p67* inhibits transactivation by Myb remains to be determined in further studies. For example, it may be possible to determine whether p67* inhibits the binding of p160 to Myb. [We noted that a small amount (~50%)

of inhibition of transactivation by mutant (L3,4P) Myb was consistently observed upon co-transfection with p67* (Fig. 8). Two possible explanations for this are that p67* can interact with another region of c-Myb in addition to the leucine zipper, and/or that it has a more general transcriptional repressor activity, due for example to interactions with other transcription factors.]

If p160 does function as a co-regulator of c-Myb-mediated transcription, it would represent a third such protein known to interact with Myb, as CBP (13,56) and p100 (14) have already been shown to do so. These may function together as part of a transcription complex or alternatively, there may be some exclusivity in these interactions. The nett effect of all of the possible interactions may be complex and difficult to predict. For example, p100 appears to repress transactivation by Myb (14), yet it was originally identified as a transcriptional co-activator (67). Moreover, while p160 has at least some characteristics of a transcriptional co-activator (see above), it binds to a region of c-Myb that is involved in negative regulation. It is possible that *in vivo*, these potential regulators differentially affect transcription of various Myb "target genes"; for instance, p160 may be required for Myb to enhance transcription of a subset of target genes involved in the function of mature hemopoietic cells but *not* required for its transforming effects (see discussion of target genes by Ness (52)).

Expression of P160

The expression of p160 mRNA corresponds to that of p160 as detected by binding to GST-Myb fusion protein in a series of cell lines (17). Our present work extends analysis of expression to a range of tissues and organs; together, these data suggest that p160 is present in most or all cell types. Thus we might suspect that p160 serves a function or functions required in all cell types. This is consistent with a role in rRNA synthesis as suggested by its nucleolar localisation, and also with its potential to interact with c-Jun, which is expressed more widely than in the limited range of cells that express c-Myb. The high level of expression in fetal liver, which is a very active site of cell proliferation, is similarly consistent with a role in rRNA synthesis. Neither of these possibilities excludes a role in interacting with c-Myb; in fact the high p160 level in fetal liver, which is enriched for hemopoietic progenitor cells that express c-Myb, is compatible with this role also.

Chromosomal localisation of P160

Finally, we should consider the potential significance of the presence in the mouse genome of three loci that hybridize with the p160 cDNA. We have termed the chromosome 11 locus *P160* because it contains sequences related to both 5' and 3' ends of the p160 cDNA, and because genomic clones that map to this locus correspond in sequence to the cDNA. Furthermore, by exploiting the weak cross-reaction of mouse p160 cDNA with human genomic sequences (which was not sufficient to allow detection of human p160 mRNA; see Fig. 3A), we have identified a human *P160* homologue located in a syntenic region of the human genome (R.K., T.J.G., E Woollat and G Sutherland, unpublished observations). Thus it is highly unlikely that one of the other loci (*P160-rs1 or -rs2*) contains the sequence which we have cloned. It is somewhat surprising that we did not detect other transcripts that could correspond to the other *P160*-related loci; this might indicate that all the transcripts are of the same size, or that expression of the related genes is much more restricted. However, it should also be borne in mind that one (or both) of these loci could represent a pseudogene(s).

The existence of the three *P160* -related genes complicates the identification of potential correlations between these genes and loci corresponding to mutant phenotypes. Without knowing the precise functions of p67 and p160, it is unclear what kinds of phenotypes one might expect. However, the ability of p160 and p67 to interact with the NRD of c-Myb raises the possibility that *P160* could potentially act as a tumor suppressor gene. To the best of our knowledge, there are no obvious murine mutants which exhibit neoplastic or pre-neoplastic phenotypes that map close to any of the three *P160* loci. Current studies on the human homologue of *P160* (R.A.K. and T.J.G., unpublished), together with the studies on the activities of murine p160 using the reagents described in this report, should help to determine whether lesions in P160 play a role in human disease.

REFERENCES

1. **Anfossi, G., A. M. Gewirtz, and B. Calabretta**. 1989. An oligomer complementary to c-myb-encoded mRNA inhibits proliferation of human myeloid leukemia cell lines. Proc Natl Acad Sci U.S.A. **86**:3379-3383.

2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl**. 1994. Current Protocols in Molecular Biology. John Wiley and Sons Inc. New York.

3. **Avraham, K. B., C. Fletcher, D. G. Overdier, D. E. Clevidence, E. Lae, R. H. Costa, N. A. Jenkins, and N. G. Copeland**. 1995. Murine chromosomal location of eight members of the hepatocyte nuclear factor 3/forkhead winged helix family of transcription factors. Genomics **25**:388-393.

4. **Bading, H., E. W. Rauterberg, and K. Moelling**. 1989. Distribution of c-myc, c-myb, and Ki-67 antigens in interphase and mitotic human cells evidenced by immunofluorescence staining technique. Exp. Cell Res. **185**:50-59.

5. **Baeuerle, P. A. and D. Baltimore**. 1988. IkappaB: A specific inhibitor of the NF-kappaB transcription factor. Science **242**:540-546.

6. **Benezra, R., R. L. Davis, D. Lockshon, D. L. Turner, and H. Weintraub**. 1990. The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. Cell **61**:49-59.

7. **Biedenkapp, H., U. Borgmeyer, A. E. Sippel, and K. H. Klempnauer**. 1988. Viral myb oncogene encodes a sequence-specific DNA-binding activity. Nature **335**:835-837.

8. **Bies, J., R. Mukhopadhyaya, J. Pierce, and L. Wolff**. 1995. Only late, nonmitotic stages of granulocyte differentiation in 32Dcl3 cells are blocked by ectopic expression of murine c-myb and its truncated forms. Cell Growth & Differentiation **6**:59-68.

9. **Buchberg, A. M., E. Brownell, S. Nagata, N. A. Jenkins, and N. G. Copeland**. 1989. A comprehensive genetic map of murine chromosome 11 reveals extensive linkage conservation between mouse and human. Genetics **122**:153-161.

10. **Chomczynski, P. and N. Sacchi**. 1987. Single-step method of RNA isolation by acid guanidinium thiocynate-phenol-chloroform extraction. Anal. Biochem. **162**:156-159.

11. **Clarke, M. F., J. F. Kukowska-Latallo, E. Westin, M. Smith, and E. V. Prochownik**. 1988. Constitutive expression of a c-myb cDNA blocks Friend murine erythroleukemia cell differentiation. Mol. Cell Biol. **8**:884-892.

12. **Copeland, N. G. and N. A. Jenkins**. 1991. Development and applications of a molecular genetic linkage map of the genome. Trends. Genet. **7**:113-118.

13. **Dai, P., H. Akimaru, Y. Tanaka, D. X. Hou, T. Yasukawa, C. Kanei-Ishii, T. Takahashi, and S. Ishii**. 1996. CBP as a transcriptional coactivator of c-Myb. Genes Dev. **10**:528-540.

14. **Dash, A. B., F. C. Orrico, and S. A. Ness**. 1996. The EVES motif mediates both intermolecular and intramolecular regulation of c-Myb. Genes Dev. **10**:1858--69.

15. **Dubendorff, J. W., L. J. Whittaker, J. T. Eltman, and J. S. Lipsick**. 1992. Carboxy-terminal elements of c-Myb negatively regulate transcriptional activation in cis and in trans. Genes Dev. **6**:2524-2535.

16. **DuBridge, R. B., P. Tang, H. C. Hsia, P. M. Leong, J. H. Miller, and M. P. Calos**. 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. Mol. Cell Biol. **7**:379-387.

17. **Favier, D. and T. J. Gonda**. 1994. Detection of proteins that bind to the leucine zipper motif of c-Myb. Oncogene **9**:305-311.

18. **Ferrao, P., E. M. Macmillan, L. K. Ashman, and T. J. Gonda**. 1995. Enforced expression of full length c-Myb leads to density-dependent transformation of murine haemopoietic cells. Oncogene **11**:1631-1638.

19. **Frohman, M. A., M. K. Dush, and G. R. Martin**. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific primer. Proc. Natl. Acad. Sci. U. S. A. **85**:8998-9002.

20. **Garcia-Bustos, J., J. Heitman, and M. N. Hall**. 1991. Nuclear protein localisation. Biochim. Biophys. Acta **1071**:83-101.

21. **Gewirtz, A. M. and B. Calabretta**. 1988. A c-myb antisense oligodeoxynucleotide inhibits normal human hematopoiesis in vitro. Science **242**:1303-1306.

22. **Gonda, T. J.** 1991. Molecular and cellular activities of Myb: regulation in normal haemopoiesis and transformation. Seminars in Virology **2**:351-361.

23. **Gonda, T. J., C. Buckmaster, and R. G. Ramsay**. 1989. Activation of c-myb by carboxy-terminal truncation: relationship to transformation of murine haemopoietic cells in vitro. EMBO J **8**:1777-1783.

24. **Gonda, T. J., D. Favier, P. Ferrao, E. M. Macmillan, R. Simpson, and F. Tavner**. 1996. The c-myb negative regulatory domain. Curr. Top. Microbiol. Immunol. **211**:99-109.

25. **Gonda, T. J. and D. Metcalf**. 1984. Expression of myb, myc and fos proto-oncogenes during the differentiation of a murine myeloid leukaemia. Nature **310**:249-251.

26. **Gonda, T. J., D. K. Sheiness, and J. M. Bishop**. 1982. Transcripts from the cellular homologs of retroviral oncogenes: distribution among chicken tissues. Mol. Cell Biol. **2**:617-624.

27. **Gonzalez, F. A., D. L. Raden, and R. J. Davis**. 1991. Identification of substrate recognition determinants for human ERK1 and ERK2 protein kinases. J. Biol. Chem. **266**:22159-22163.

28. **Grasser, F. A., T. Graf, and J. S. Lipsick**. 1991. Protein truncation is required for the activation of the c-myb proto-oncogene. Mol. Cell Biol. **11**:3987-3996.

29. **Green, E. L.** 1981. Linkage, recombination and mapping, p. 77-113. In AnonymousGenetics and probability in animal breeding experiments. Oxford University Press, New York.

30. **Harlow, E., P. Whyte, B. R. Franza, and C. Schely**. 1986. Association of adenovirus early-region 1A proteins with cellular polypeptides. Mol. Cell. Biol. **6**:1579-1589.

31. **Heery, D. M., E. Kalkhoven, S. Hoare, and M. G. Parker**. 1997. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. Nature **387**:733-736.

32. **Hogan, A., S. Heyner, M. J. Charron, N. G. Copeland, D. J. Gilbert, N. A. Jenkins, B. Thorens, and G. A. Schultz**. 1991. Glucose transporter gene expression in early mouse embryos. Development **113**:363-372.

33. **Hopp, T. P., K. S. Pickett, V. Price, R. T. Libby, C. J. March, P. Cerretti, D. L. Urdal, and P. J. Conlon**. 1988. A short polypeptide marker sequence useful for recombinant protein identification and purification. Biotechnology **6**:1205-1210.

34. **Hu, Y. L., R. G. Ramsay, C. Kanei Ishii, S. Ishii, T. J. Gonda, and C. Kanei-Ishii**. 1991. Transformation by carboxyl-deleted Myb reflects increased transactivating capacity and disruption of a negative regulatory domain. Oncogene **6**:1549-1553.

35. **Jans, D. A. and S. Huebner**. 1996. Regulation of protein transport to the nucleus: central role of phosphorylation. Physiol. Rev. **76**:651-685.

36. **Jantzen, H. M., A. Admon, S. P. Bell, and R. Tjian**. 1990. Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. Nature **344**:830-836.

37. **Jenkins, N. A., N. G. Copeland, B. A. Taylor, and B. K. Lee**. 1982. Organization, distribution,and stability of endogenous ecotropic murine leukemia virus DNA sequences in chromosomes of *Mus musculus*. J. Virol. **43**:26-36.

38. **Kanei-Ishii, C., E. M. Macmillan, T. Nomura, A. Sarai, R. G. Ramsay, S. Aimoto, S. Ishii, and T. J. Gonda**. 1992. Transactivation and transformation by Myb are negatively regulated by a leucine-zipper structure. Proc Natl Acad Sci U S A **89**:3088-3092.

39. **Kennelly, P. J. and E. G. Krebs**. 1991. Consensus sequences as substrate specificity determinants for protein kinases and protein phosphatases. J. Biol. Chem. **266**:15555-15558.

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

41. **Kuhn, A. and I. Grummt**. 1992. Dual role of the nucleolar transcription factor UBF: trans-activator and antirepressor. Proc Natl Acad Sci U S A **89**:7340-7344.

42. **Kuhn, A., R. Voit, V. Stefanovsky, R. Evers, M. Bianchi, and I. Grummt**. 1994. Functional differences between the two splice variants of the nucleolar transcription factor UBF: the second HMG box determines specificity of DNA binding and transcriptional activity. EMBO J **13**:416-424.

43. **Lane, T., C. Ibanez, A. Garcia, T. Graf, and J. Lipsick**. 1990. Transformation by v-myb correlates with trans-activation of gene expression. Mol. Cell Biol. **10**:2591-2598.

44. **Maekawa, T., S. Matsuda, J. Fujisawa, M. Yoshida, and S. Ishii**. 1991. CRE-binding protein, CRE-BP1, mediates the E1A-induced but not the Tax-induced transactivation. Oncogene **6**:627-632.

45. **Meier, U. T. and G. Blobel**. 1990. A nuclear localization signal binding protein in the nucleolus. J. Cell Biol. **111**:2235-2245.

46. **Meier, U. T. and G. Blobel**. 1992. Nopp140 shuttles on tracks between nucleolus and cytoplasm. Cell **70**:127-138.

47. **Miau, L.-H., C.-J. Chang, W.-H. Tsai, and S.-C. Lee**. 1997. Identification and characterization of a nucleolar phosphoprotein, Nopp140, as a transcription factor. Mol. Cell Biol. **17**:230-239.

48. **Miller, A. D., D. R. Trauber, and C. Buttimore**. 1986. Factors involved in the production of helper virus-free retrovirus vectors. Somat. Cell Mol. Genet. **12**:175-183.

49. **Moritz, R. L., J. H. Eddes, G. E. Reid, and R. J. Simpson**. 1996. S-Pyridylethylation of intact polyacrylamide gels and *in situ* digestion of electrophoretically separated proteins: a rapid mass spectrometric method for identifying cysteine-containing peptides. Electrophoresis **17**:907-917.

50. **Mucenski, M. L., K. McLain, A. B. Kier, S. H. Swerdlow, C. M. Schreiner, T. A. Miller, D. W. Pietryga, W. J. Scott,Jr., and S. S. Potter**. 1991. A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. Cell **65**:677-689.

51. **Nakagoshi, H., C. Kanei Ishii, T. Sawazaki, G. Mizuguchi, S. Ishii, and C. Kanei-Ishii**. 1992. Transcriptional activation of the c-myc gene by the c-myb and B-myb gene products. Oncogene **7**:1233-1240.

52. **Ness, S. A.** 1996. The Myb oncoprotein: regulating a regulator. Biochim. Biophys. Acta **1288**:F123-F129.

53. **Ness, S. A., A. Marknell, and T. Graf**. 1989. The v-myb oncogene product binds to and activates the promyelocyte-specific mim-1 gene. Cell **59**:1115-1125.

54. **Nishina, Y., H. Nakagoshi, F. Imamoto, T. J. Gonda, and S. Ishii**. 1989. Transactivation by the c-myb proto-oncogene. Nucleic Acids Res **17**:107-117.

55. **Nomura, T., N. Sakai, A. Sarai, T. Sudo, C. Kanei Ishii, R. G. Ramsay, D. Favier, T. J. Gonda, and S. Ishii**. 1993. Negative autoregulation of c-Myb activity by homodimer formation through the leucine zipper. J. Biol. Chem. **268**:21914-21923.

56. **Oelgeschlager, M., R. Janknecht, J. Krieg, S. Schreek, and B. Luscher**. 1996. Interaction of the co-activator CBP with Myb proteins: effects on Myb-specific transactivation and on the cooperativity with NF-M. EMBO J **15**:2771--80.

57. **Pear, W. S., G. P. Nolan, M. L. Scott, and D. Baltimore**. 1993. Production of high-titer helper-free retroviruses by transient transfection. Proc. Natl. Acad. Sci. U. S. A. **90**:8392-8396.

58. **Ramsay, R. G., S. Ishii, and T. J. Gonda**. 1991. Increase in specific DNA binding by carboxyl truncation suggests a mechanism for activation of Myb. Oncogene **6**:1875-1879.

59. **Ramsay, R. G., S. Ishii, Y. Nishina, G. Soe, and T. J. Gonda**. 1989. Characterization of alternate and truncated forms of murine c-myb proteins. Oncogene Research **4**:259-269.

60. **Rayner, J. R. and T. J. Gonda**. 1994. A simple and efficient procedure for generating stable expression libraries by cDNA cloning in a retroviral vector. Mol. Cell Biol. **14**:880-887.

61. **Reid, G. E., H. Ji, J. H. Eddes, R. L. Moritz, and R. J. Simpson**. 1995. Nonreducing two-dimensional polyacrylamide gel electrophoretic analysis of human colonic proteins. Electrophoresis **16**:1120-1130.

62. **Sakura, H., C. Kanei-Ishii, T. Nagase, H. Nakagoshi, T. J. Gonda, and S. Ishii.** 1989. Delineation of three functional domains of the transcriptional activator encoded by the c-myb protooncogene. Proc Natl Acad Sci U S A **86**:5758-5762.

63. **Sambrook, J., E. F. Frtisch, and T. Maniatis.** 1989. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

64. **Shen-Ong, G. L.** 1990. The myb oncogene. Biochimica et Biophysica Acta **1032**:39-52.

65. **Sigler, P. B.** 1988. Acid blobs and negative noodles. Nature **333**:210-212.

66. **Takahashi, T., M. Tanaka, C. I. Brennan, N. A. Jenkins, N. G. Copeland, T. Suda, and S. Nagata.** 1994. Generalized lymphoproliferative disease in mice caused by a point mutation in the Fas ligand. Cell **76**:969-976.

67. **Tong, X., R. Drapkin, R. Yalamanchili, G. Mosialos, and E. Kieff.** 1995. The Epstein-Barr virus nuclear protein 2 acidic domain forms a complex with a novel cellular coactivator that can interact with TFIIE. Mol. Cell Biol. **15**:4735-4744.

68. **Torchia, J., D. W. Rose, J. Inostroza, Y. Kamei, S. Westin, C. K. Glass, and M. G. Rosenfeld.** 1997. The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. Nature **387**:677-684.

69. **Voelker, R. A., S.-M. Huang, G. B. Wisely, J. F. Sterling, S. P. Bainbridge, and K. Hiraizumi.** 1989. Molecular and genetic organization of the *suppressor of sable* and *minute (1) 1B* region in *Drosophila melanogaster*. Genetics **122**:625-642.

70. **Westin, E. H., R. C. Gallo, S. K. Arya, A. Eva, L. M. Souza, M. A. Baluda, S. A. Aaronson, and F. Wong-Staal.** 1982. Differential expression of the *amv* gene in human hematopoietic cells. Proc. Natl. Acad. Sci. U. S. A. **79**:2194-2198.

71. **Weston, K. and J. M. Bishop.** 1989. Transcriptional activation by the v-myb oncogene and its cellular progenitor, c-myb. Cell **58**:85-93.

72. **Yanagisawa, H., T. Nagasawa, S. Kuramochi, T. Abe, Y. Ikawa, and K. Todokoro.** 1991. Constitutive expression of exogenous c-myb gene causes maturation block in monocyte-macrophage differentiation. Biochim. Biophys. Acta **1088**:380-384

FIGURE 1.

Molecular cloning, nucleotide sequence and predicted amino acid sequence of p160 (see text and "Materials and Methods" for details). (A) Overlapping cDNA clones that comprise the p160 cDNA and location of p67-derived peptide sequences. The top four (open) boxes represent cDNA clones derived by PCR, library screening or 5'-RACE PCR as indicated. The bottom box (shaded) represents the entire sequence and shows the location of the sequences corresponding to the five p67-derived tryptic peptides (filled boxes). (B) Nucleotide sequence of p160 cDNA and predicted amino acid sequence of p160. The p67-derived tryptic peptide sequences are **underlined and bold**, the acidic region is boxed, and the basic repeats are *italicized and bold* (Note that two of these sequences contain overlapping repeat motifs while that starting at residue 1332 is incomplete). Potential LCD (leucine charged domain; see text) motifs are indicated by dashed overlining. The symbol ⏌ indicates the carboxyl terminus of p67*.
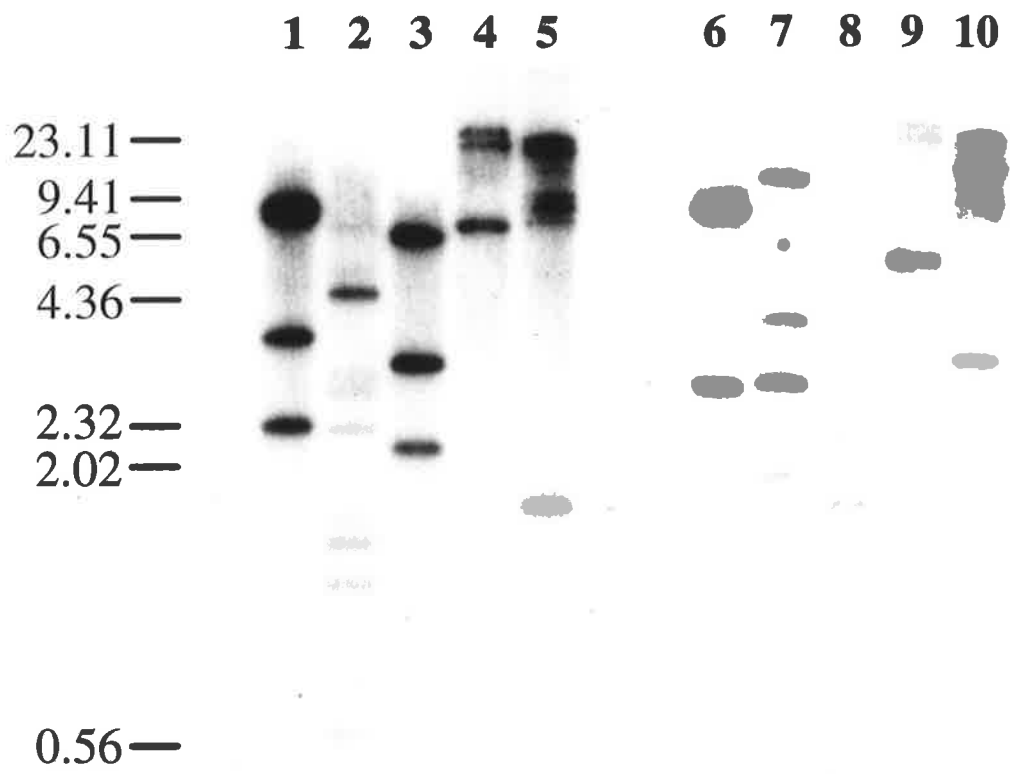
B
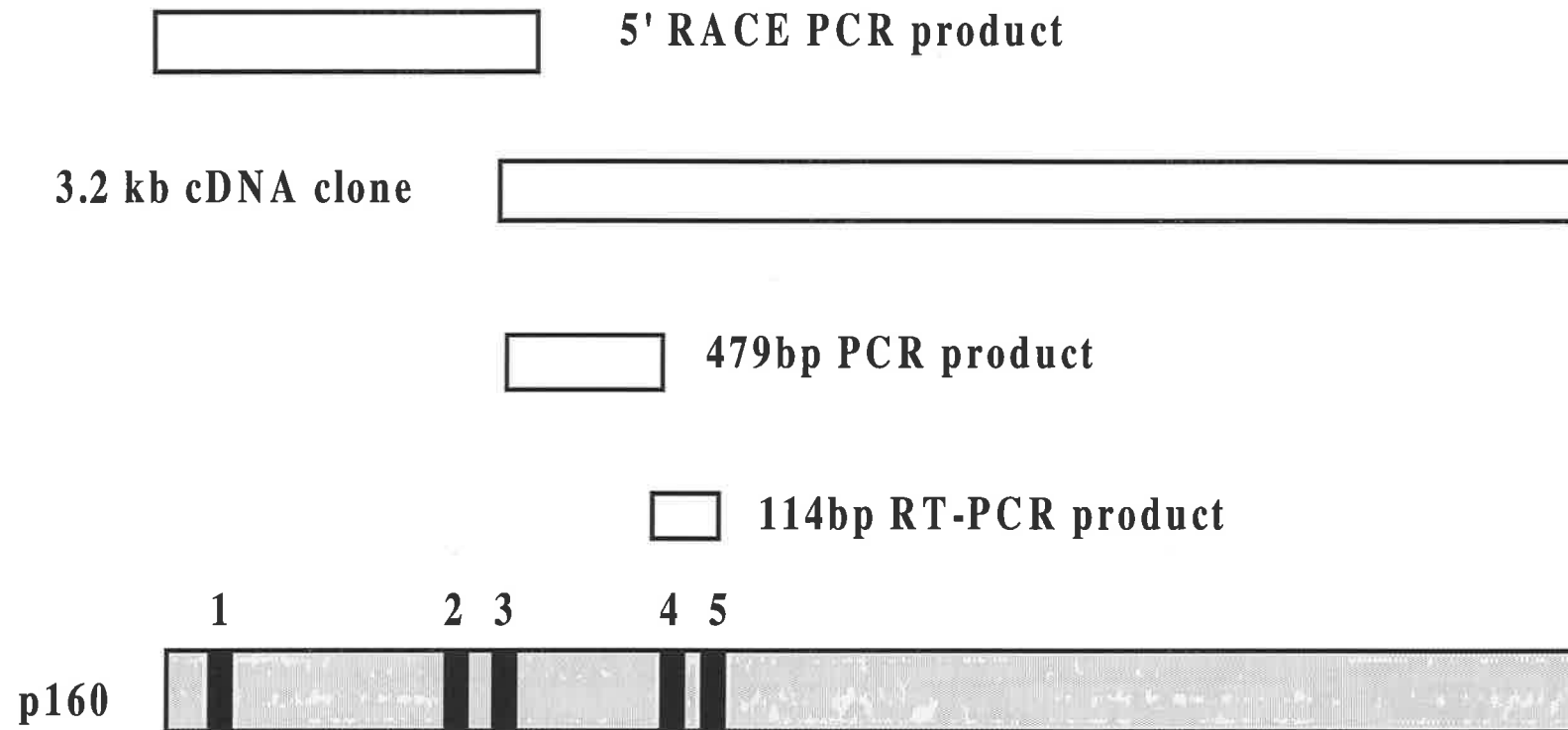
```
1        MetAlaGluMetLysSerProThrLysAlaGluProAlaThrProAlaGluAlaAlaGlnSerAspArgHisSerLeuLeuLeuHisSerArgGluPheLeuAsp    35
1    CGTGTTTGGCTCAGCATGGCGGAGATGAAGAGCCCTACGAAAGCTGAGCCTGCGACTCCCGCAGAAGCGGCGCAAAGCGACCGCCACAGCCTGCTGGAGCACAGCCGCGAGTTCTTGGAC   120

36       PhePheTrpAspIleAlaLysProAspGlnGluThrArgLeuArgAlaThrGluLysLeuLeuGluTyrLeuArgThrArgProAsnAspSerGluMetLysTyrAlaLeuLysArgLeu    75
121  TTCTTCTGGGACATTGCGAAACCGGATCAGGAAACGCGGCTCCGGGCCACGGAGAAGTTGTTGGAGTACTTGCGCACAAGGCCCAATGATTCGGAGATGAAATATGCCCTGAAGCGCCTA    240

76       IleThrGlyLeuGlyValGlyArgGluAlaAlaThrAlaCysTyrSerLeuAlaLeuAlaGlnLeuGlnSerPheGluAspIleProLeuCysAspIleLeuAspGlnIleGlnGlu      115
241  ATCACTGGGCTTGGGGTGGGCCGAGAAGCCGCTACGGCCTGCTACAGCCTGGCGCTGGCACAGCTGTTGCAGTCTTTTGAAGACATCCCATTGTGTGACATCCTGGATCAGATACAAGAA    360

116      LysTyrSerLeuGlnAlaMetAsnLysAlaMetMetArgProSerLeuPheAlaAsnLeuPheGlyValLeuAlaLeuAlaLeuPheGlnSerGlyArgLeuValLysAspLysAlaLeuMet   155
361  AAATACAGCCTACAAGCCATGAACAAGGCAATGATGAGACCTAGTCTCCTTTGCCAAACCTTTTTGGAGTGCTAGCCCTCCTTTCAGTCAGGCCGCCTAGTGAAGGACAAAGAGGCCCTCATG    480

156      LysSerValGlnLeuLeuLysIleLeuSerGlnHisProAsnHisLeuGlnGlyGlnProIleLysAlaLeuValAspIleLeuSerGluValProGluSerMetPheGlnGluIleLeu    195
481  AAGTCCGTGCAATTGCTGAAGATCCTGTCCCAACACCCCAACCACTTACAGGGACAGCCAATAAAGGCTCTGGTGGACATCCTCTCTGAGGTCCCAGAGTCCATGTTCCAGGAGATCCTG    600

196      ProLysValLeuLysGlyAsnMetLysValIleLeuLeuArgSerProLysTyrLeuGluGluProLeuLeuAlaLysGlnArgValProThrGlnSerLeuMetGlySerValAsp      235
601  CCAAAGGTCCTCAAGGGTAACATGAAAGTGATCCTCCGCTCTCCCAAGTACTTGGAGCTCTTCCTCCTGGCTAAGCAGAGGGTGCCGACAAAGCTCGAGTCACTCATGGGCTCGGTTGAC    720

236      LeuPheSerGluAspAsnIleProSerLeuValLeuAsnIleLeuLysLeuValAlaAlaAlaAsnSerValLysLysGluHisLysLeuProAsnValAlaLeuAspLeuLeuArgLeuAlaLeuLys   275
721  CTATTCTCAGAAGACAATATTCCCAGTCTGGTGAACATCCTGAAGGTGGCCGCCAACTCTGTCAAGAAGGAGCACAAGCTGCCTAATGTGGCTCTGGACCTGCTCCGCCTGGCTCTCAAG    840

276      GluSerArgPheGluLeuPheTrpLysLysValLeuGluGluGlyLeuLeuLysAsnProSerTrpThrSerSerTyrMetCysPheArgLeuLeuLeuGlyAlaSerLeuProLeuLeuSer   315
841  GAGAGCAGATTCGAACTGTTCTGGAAGAAGGTTTTGGAGGAGGGGCTGCTGAAGAATCCGTCCTGGACATCCAGCTACATGTGCTTCCGCCTACTGGGTGCGTCTCTGCCGCCTTCTGTCA    960

316      GluGluGlnLeuGlnLeuValMetArgGlyAspLeuIleArgHisPheGlyGluAsnMetValIleSerLysProGlnAsnLeuPheLysIleIleProGluIleSerThrTyrValGly    355
961  GAGGAGCAGTTGCAGTTGGTGATGCGAGGAGACTTGATCCGCCATTTTGGGGAGAACATGGTTATTTCTAAGCCCCAAAACCTATTTAAGATCATCCCAGAGATAAGTACATACGTGGGT   1080

356      ThrPheLeuGlnGlyCysGlnAspAspProLysArgGlnLeuThrMetMetValAlaPheThrThrIleThrAsnGlnGlyLeuProValMetProThrPheTrpArgValThrArgPhe    395
1081 ACCTTCCTAGAGGGGTGCCAGGATGACCCTAAGCGGCAGTTGACTATGATGGTGGCCTTTACAACCATCACCAATCAAGGTCTCCCTGTCATGCCTACCTTCTGGCGTGTCACGCGGTTT   1200

396      LeuAsnAlaGluAlaAlaLeuGlnSerTyrValAlaTrpLeuArgAspMetPheLeuGlnProAspLeuAsnSerLeuValAspPheSerThrAlaAsnGlnLysArgAlaGlnAspAlaSer   435
1201 TTGAATGCTGAAGCCCTGCAGAGCTATGTGGCCTGGTTGCGGGACATGTTCCTGCAGCCTGACCTGAACTCCTTGGTTGACTTCAGCACTGCCAACCAGAAGAGAGCTCAGGACGCCTCG   1320

436      LeuAsnValProGluArgAlaValPheArgLeuArgLysTrpIleIleHisArgLeuValSerLeuValAspHisLeuGluLysAspGluAlaValValGluGlnIleAlaArg      475
1321 TTGAATGTGCCTGAGCGAGCTGTATTCCGGCTCCGGAAGTGGATCATCCACCGCCTGGTCAGCCTTGTGGATCATTTGCATCTGGAGAAGGATGAAGCTGTGGTTGAGCAAATAGCCAGG   1440

476      PheCysLeuPheHisAlaPhePheLysThrLysLysAlaThrProGlnIleProGluProThrLysGlnHisPheSerPheProLeuAspAspArgAsnArgGlyValPheValSerAlaPhe   515
1441 TTTTGCTTGTTCCATGCCTTCTTTAAGACGAAGAAGGCTACGCCCCAGATCCCAGAGCCGACGAAGCAGCACTTCTCCTTCCCTTTGGACGACGCCAACCGTGGGGTCTTTGTCAGTGCCTTC   1560

516      PheSerLeuLeuGlnThrLeuSerValLysPheArgGlnThrProAspLeuAlaGluAsnGlyLysProTrpThrTyrArgLeuValGlnLeuAlaAspMetLeuLeuAsnHisAsnArg    555
1561 TTCAGCCTACTGCAGACGCTCAGTGTGAAGTTCAGGCAGACACCAGACCTGGCTGAAAATGGGAAGCCTTGGACTTACCGCCTGGTTCAGTTGGCAGACATGCTGTTGAACCATAACCGC   1680

556      AsnValThrSerValThrSerLeuThrThrGlnIleAspHisGlnSerLeuThrLeuAspGlnAlaTrpAspGlnMetMetSerThrLeuLysGluLeuAlaArgSerGluThrArgAlaIleAlaPheGlnHis   595
1681 AATGTAACCAGTGTGACATCCTTGACAACACACGCAGCGTCAGGCCTGGGACCAGATGATGAGTACTCTGAAGGAATTAGAGGCCCGCTCCTCTGAGACGCAGGGCCATTGCCTTCCAGCAC   1800

596      LeuLeuLeuGluValGlyLeuHisIlePheLysSerProAlaGluSerCysSerAspValLeuGlyAspIleGlnThrCysIleLysLysSerMetGluGlnAsnProArgArgSerArgSer   635
1801 CTGCTGCTTCTGGTGGGCCTCCACATCTTCAAGTCCCCTGCAGGAAAGCTGTGATGTCCTAGGAGACATTCAGACTTGCATCAAGAAAAGCATGGAGCAGAATCCCCGCCGATCACGCTCT   1920

636      ArgAlaLysAlaSerGlnGluProValTrpValGluValMetValGluIleLeuSerLeuLeuLeuAlaGlnProSerAsnLeuMetArgGlnValValArgSerValPheGlyHisIle    675
1921 AGAGCCAAAGCCTCCCAGGAGCCAGTATGGGTGGAGGTGATGGTGGAGATCTTGCTGTCCTTGCTGGCTCAACCCAGCAACTTGATGCGCCAGGTGGTCCGGAGTGTATTTGGTCATATC   2040

676      CysProHisAsnLeuThrProArgCysLeuGlnLeuIleLeuAlaValAlaLeuAlaLeuSerProValAsnGluAspGluAspAspAsnValValValValThrAspAspAlaAspGluLysGlnLeuGln   715
2041 TGTCCCCACCTTACTCCACGTTGTCTGCAGCTAATCCTGGCTGTGCTCAGCCCTGTGACAAAGAGGATGAGGATGACAACGTGGTGGTCACTGATGACGCTGATGAGAAGCAGCTGCAG    2160

716      HisGlyGluAspGluSerAspAsnGluAspAsnLysAsnSerGluSerAspMetAspSerGluAspGlyGluGluSerGluGluGluAspArgAspLysAspValAspProGlyPhe     755
2161 CATGGAGAGGACGAAGACTCAGATAATGAGGACAATAAGAACTCAGAGAGTGACATGGACAGTGAGGATGGGGAAGAAAGTGAAGAGGAGGACCGTGACAAAGATGTGGACCCAGGCTTC   2280

756      ArgGlnGlnLeuMetGluValLeuLysAlaGlyAsnAlaLeuGlyGlyValAlaAsnGluGluGluGluLeuGlyAspGluAlaMetMetAlaLeuAspGlnAsnLeuAlaSerLeu     795
2281 CGTCAACAGTTGATGGAAGTGTTAAAAGCTGGGAATGCATTGGGTGGAGTGGACAACGAGGAGGAGGAGGAGCTTGGGGATGAGGCCATGATGGCCCTGGACCAGAACCTGGCCAGCCTA   2400

796      PheLysLeuGlnLysMetArgIleGlnAlaArgAsnGlyGluGluLysLysAsnLysSerAspLysLysAlaThrArgArgAspPheGlnAlaArgAlaLeuAspLeuGluGluValLeuVal   835
2401 TTTAAAGAGCAGAAGATGCGCATCCAGGCCCGGAATGAGGAGAAAAACAAGCTACAGAAGGAGAAGAAGTCCGACGGGACTTCCAAATCAGGGCACTAGACCTGATCGAGGTGCTGGTG   2520

836      ThrLysGlnProGluHisProLeuIleLeuGluLeuLeuGluProLeuLeuAsnValIleGlnHisSerMetArgSerLysGlySerThrLysGlnGluGlnAspLeuLeuHisLysThr    875
2521 ACCAAGCAGCCTGAGCACCCCCTGATCCTGGAACTACTTGAGCCACTGCTGAACGTGATCCAGCACAGTATGCGCAGCAAAGGCTCCACCAAGCAGGAGCAGGACCTCCTGCACAAGACC   2640

876      AlaArgIlePheMetHisHisLeuCysArgAlaArgArgTyrCysHisGluValGlyProCysAlaGluAlaLeuHisAlaGlnValGluArgLeuValGlnGlnLeuAlaGlySerGlnAla   915
2641 GCCCGCATCTTCATGCACCACCTGTGTCGTGCCCGCCGCTACTGCCACGAGGTGGGGCCGTGTGCAGAGGCTCTGCATGCCCAGGTGGAGAGGCTTGTGCAGCAGGCTGGCAGCCAGGCT   2760

916      AspAlaSerValAlaLeuTyrTyrPheAsnAlaSerLeuTyrLeuLeuArgValLeuAsnThrAsnLysArgHisGlnAspGlyHisLysLeuHisGlyValAlaAspThrGluAsp     955
2761 GATGCCTCTGTCGCCCTATACTATTTCAATGCCTCTCTGTACCTGCTGCGAGTCCTCAAGGGCAACACCAATAAGAGGCACCAAGATGGTCATAAGTTACACGGAGCTGACACAGAGGAC   2880

956      SerGluAspGlnAlaAlaAsnCysLeuAspLeuAspPheValThrArgValTyrSerAlaSerLeuGluSerLeuLeuThrLysArgAsnSerSerLeuThrValProMetPheLeuSer    995
2881 TCAGAGGACCAGGCTGCTAACTGCTTAGACTTGGACTTTGTGACCCGGGTGTATTCAGCAAGCCTCGAACTCCTTCTGACCAAGCGTAACAGCTCACTTACGGTCCCCATGTTCCTCAGC   3000

996      LeuPheSerArgTyrProValIleCysLysLysLeuProValLeuAlaGlnHisValAlaGlyProSerArgProArgHisGlnHisLeuAlaGlnHisAlaCysLeuMetLeuGlnLysThrLeu   1035
3001 CTCTTCTCCAGATACCCAGTGATCTGTAAGAACCTGCTTCCCGTCCTGGCTCAGCATGTGGCTGGCCCATCTCGGCCCCGCCATCAGGCCCCAGGCATGCCTGATGCTTCAGAAGACTCTG   3120

1036     SerAlaArgGluLeuArgValPheGluAspProGluTrpGluGlnLeuIleLeuAsnLeuLysArgThrGlnIleLeuGlnThrLeuGlnHisThrGlnGluAlaGlnSerLysGlyGlu   1075
3121 TCTGCACGAGAGCTGAGAGTCTTGTTTGAGGATCCTGAGTGGGAGCAGCTGATTACCCAACTCTTGGGAAAGGCCACCCAGACCCTGCAGACTCTTCAGGGAGGCACAGAGCAAGGGGGAG   3240

1076     HisGlnLysGluLeuSerIleLeuGluLeuLeuAsnThrLeuLeuThrValAsnHisGluLysLeuValSerAlaProLeuGlyValLeuGlnSerLysSerGlnLysLys   1115
3241 CACCAGAAGGAGCTATCCATCTTGGAGCTGCTGAACACTCTGTTGAGGACAGTAAGTCATCAAGTGGTCAGCGCTCCGCTGGGCGTGCTTCAGAGCAAGCAACAGAAG   3360

1116     LeuGlnGlnSerLeuGlnIleGlyAsnHisSerSerGlySerAsnArgLeuTyrAspLeuTyrTrpGlnAlaMetArgMetLeuGlyValGlnArgProLysSerGluGlyLysLysAsnAla   1155
3361 CTGCAGCAAAGCCTGCAGCAGGGGAATCACTCATCTGGCTCCAATCGCCTCTATGATCTCTACTGGCAGGCCATGAGGATGCTAGGAGTGCAACGTCCAAAGTCAGAAAAGGAAGAATGCC   3480

1156     LysAspIleProSerAspThrGlnSerProValSerThrLysArgLysLysLysGlyPheLeuProGluThrLysLysArgLysLysLeuLysSerGluGlyThrThrProGluLysAsn   1195
3481 AAGGATATTCCTAGTGACACCCAGAGCCCCGTCAGCACAAAGCGGAAGAAAAAGGGATTCTTGCCAGAGACCAAGAAGCGAAAGAAACTTAAATCTGAGGGCACCACACCAGAAAAGAAT   3600

1196     AlaAlaSerGlnGlnAsnAspAlaValThrGluGlyAlaMetProAlaAlaThrGlyLysAspGlnProProSerThrGlyLysLysLysArgArgValLysAlaSerThrProSerGln   1235
3601 GCTGCGTCACAGCAGGATGCAGTGACAGAGGGTGCCATGCCTGCTGCCACTGGTAAAGACCAGCCCCCCAGCACAGGCAAGAAGAAAAGGAGAGGGTAAAGGCCAGCACCCCATCCCAG   3720

1236     ValAsnGlyIleThrGlyAlaLysSerProAlaProSerAsnProThrLeuSerProSerSerThrProAlaLysValThrProLysLeuGlnLysLysLysArgLeuLeuSerGlnValAsnGly   1275
3721 GTGAATGGGATAACTGGGGCCAAGAGTCCAGCTCCCAGTAACCCCACCCTAAGCCCCAGCAGCACCCCTGCCAAGCACCCCAAAACTGCAGAAGAAAAAGAGAAGCTGTCACAGGTGAATGGA   3840

1276     AlaThrProValSerProIleLeuProGluSerLysLysHisHisGlnLeuAlaLeuSerThrLysGluValIleArgLysSerProHisProGlnSerAlaLeuProLysLysArgAla   1315
3841 GCCACTCCTGTGTCCCCCATAGAGCCTGAAAGCAAAAAGCATCATCAGGAGGCACTAAGCACCAAGGAGGTCATAAGAAAGTCCCCCCACCCCCAGTCTGCCCCTGCCAAGAAAAGAGCA   3960

1316     ArgLeuSerProLeuValSerArgSerProSerProLeuLeuGlnSerGlyValLysLysArgArgValAlaAlaSerArgArgValGlnThrPro***   1344
3961 AGGCTGTCTCCTGGTGAGCAGGAGCCCCAGCCTGTTACAGAGTGGGGTCAAGAAAAGGAGAGTGGCCAGCAGGAGAGTGCAGACACCTTGAGTGTGTATAGTCTGTTTCCCTGCCCTAGAG   4080

4081 ACTCCTATTTTTTCACCAATATTTTAATAAACAATCCATGATGCTA   4126
```
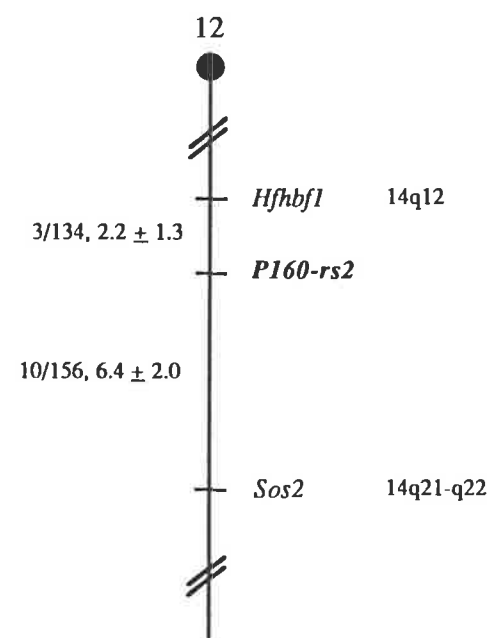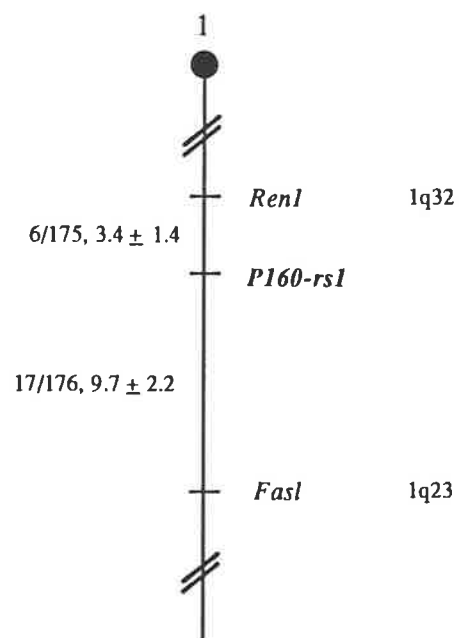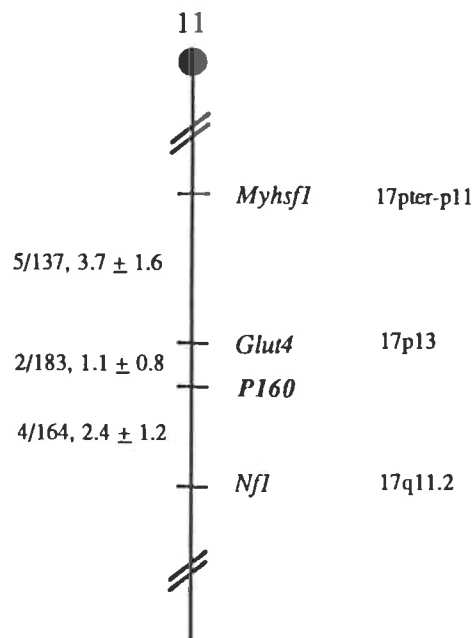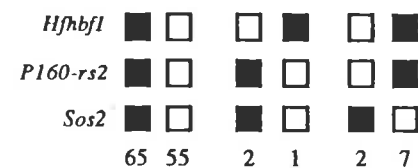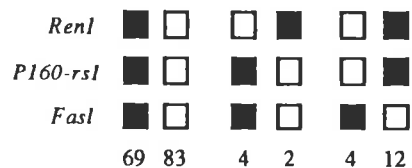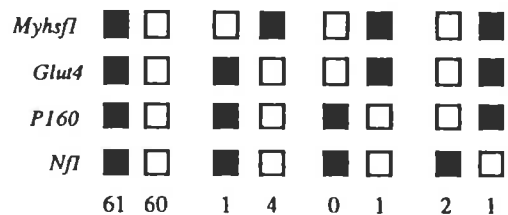
FIGURE 2.

Genomic organisation and chromosomal localisation of *P160* and related genes, *P160-rs1* and *P160-rs2*. (A) Southern blot analysis of murine DNA with probes corresponding to (left) the 5'-most 1.0 kb and (right) the 3'-most 1.0 kb of the p160 cDNA sequence. DNA was digested with the following restriction endonucleases: lanes 1 and 6, *Xba*I; lanes 2 and 7, *Sac*I; lanes 3 and 8, *Pst*I; lanes 4 and 9, *Hind*III and lanes 5 and 10, *Eco*RI. The positions and sizes (in kb) of DNA markers are shown at the left. (B) Summary of chromosomal localisation of p160-related loci in the mouse genome, as determined by interspecific backcross analysis. The grids above each chromosome map represent the possible genotypes of the backcross mice, with closed squares representing the *M. spretus* alleles and open squares representing the *M. musculus* alleles; the numbers of recombinants of the indicated genotypes are shown below the grid. The number of recombinant N2 animals, over the total number of N2 animals is shown at the left of each chromosome map. The recombination frequencies between each *P160*-related locus and the adjacent markers, expressed as genetic distance in centimorgans (± one standard error), are also shown. The positions of corresponding loci on human chromosomes, where known, are shown at the right of each chromosome map.

A

| | 1 | 2 | 3 | 4 | 5 | | 6 | 7 | 8 | 9 | 10 |

23.11 —

9.41 —
6.55 —

4.36 —

2.32 —
2.02 —

0.56 —

B

Myhsf1	■	□		□	■		□	■		□	■
Glut4	■	□		■	□		□	■		□	■
P160	■	□		■	□		■	□		□	■
Nf1	■	□		■	□		■	□		■	□
	61	60		1	4		0	1		2	1

11

5/137, 3.7 ± 1.6		Myhsf1	17pter-p11
2/183, 1.1 ± 0.8		Glut4	17p13
		P160	
4/164, 2.4 ± 1.2			
		Nf1	17q11.2

Ren1	■	□		□	■		□	■
P160-rs1	■	□		■	□		□	■
Fasl	■	□		■	□		■	□
	69	83		4	2		4	12

1

6/175, 3.4 ± 1.4		Ren1	1q32
		P160-rs1	
17/176, 9.7 ± 2.2			
		Fasl	1q23

Hfhbf1	■	□		□	■		□	■
P160-rs2	■	□		■	□		□	■
Sos2	■	□		■	□		■	□
	65	55		2	1		2	7

12

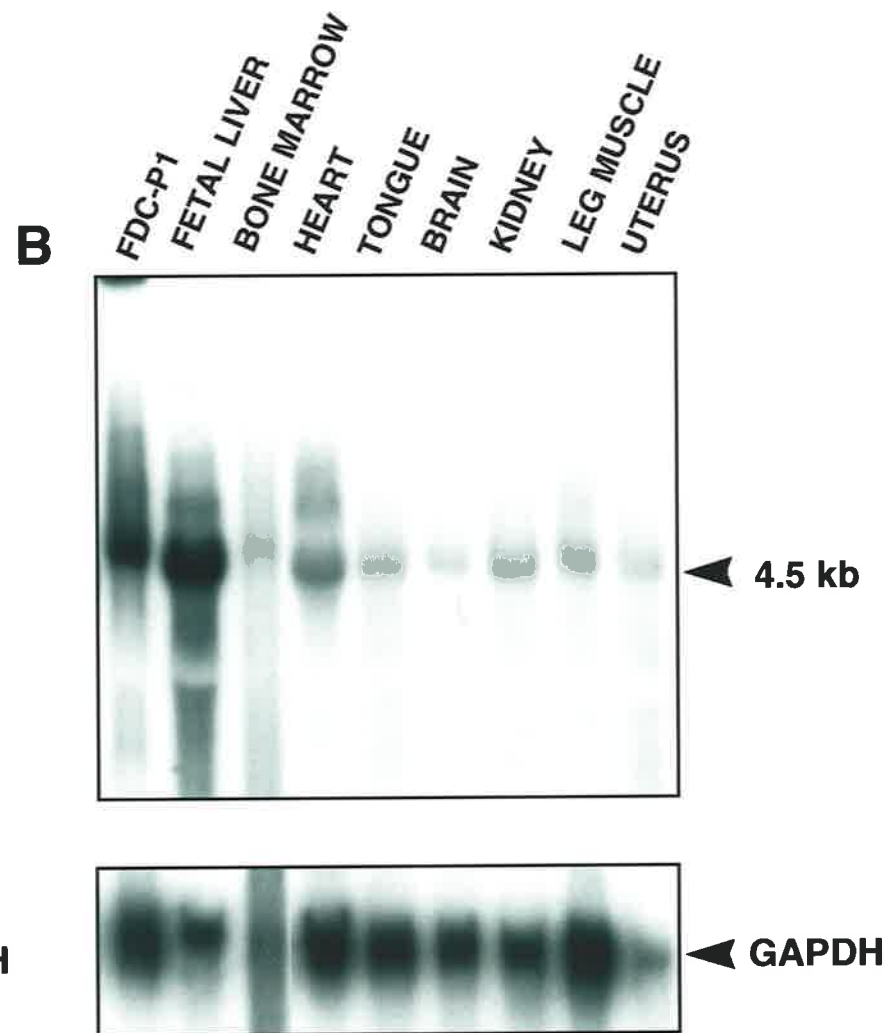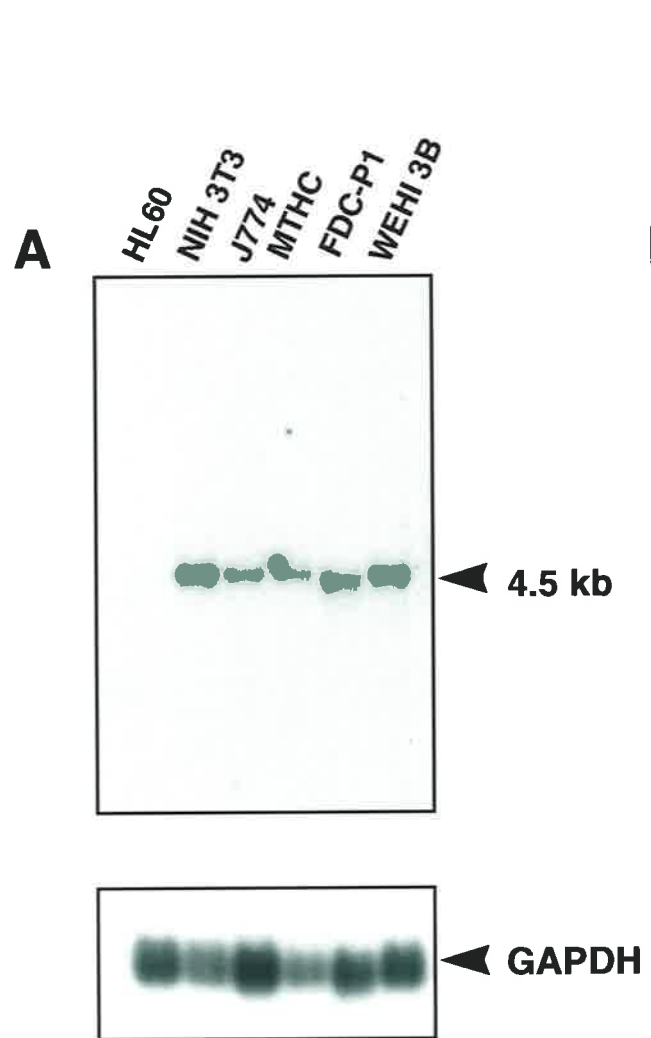| 3/134, 2.2 ± 1.3 | | Hfhbf1 | 14q12 |
		P160-rs2	
10/156, 6.4 ± 2.0			
		Sos2	14q21-q22

FIGURE 3.

Northern blot analysis of p160 mRNA expression with the p160 cDNA probe. Upper panels: (A), 2µg of polyA$^+$ RNA from the indicated cell lines were loaded in each lane (Abbreviation: MTHC - *myb* transformed hemopoietic cells); (B), 30µg of total RNA from the indicated mouse tissues (or FDC-P1 cells) were loaded in each lane. The size of the p160 mRNA was determined by comparison with RNA markers (not shown). Lower panels: the Northern blots shown in the upper panels were stripped and re-probed with a GAPDH cDNA as a control for mRNA loading; only the relevant area of each blot is shown.

A

HL60　NIH 3T3　J774　MTHC　FDC-P1　WEHI 3B

◄ 4.5 kb

◄ GAPDH

B

FDC-P1　FETAL LIVER　BONE MARROW　HEART　TONGUE　BRAIN　KIDNEY　LEG MUSCLE　UTERUS

◄ 4.5 kb

◄ GAPDH

FIGURE 4.

Proteins detected by antisera raised to the N- and C-termini of p160. Cytoplasmic and nuclear extracts from FDC-P1 and NIH3T3 cells, as indicated, were fractionated by SDS-PAGE and immunoblotted with antisera against (A), the N-terminal (160N) or (B), the C-terminal (160C) regions of p160. The positions of p160 and p67 are indicated by arrows and those of size standards (in kD) are indicated at the left.
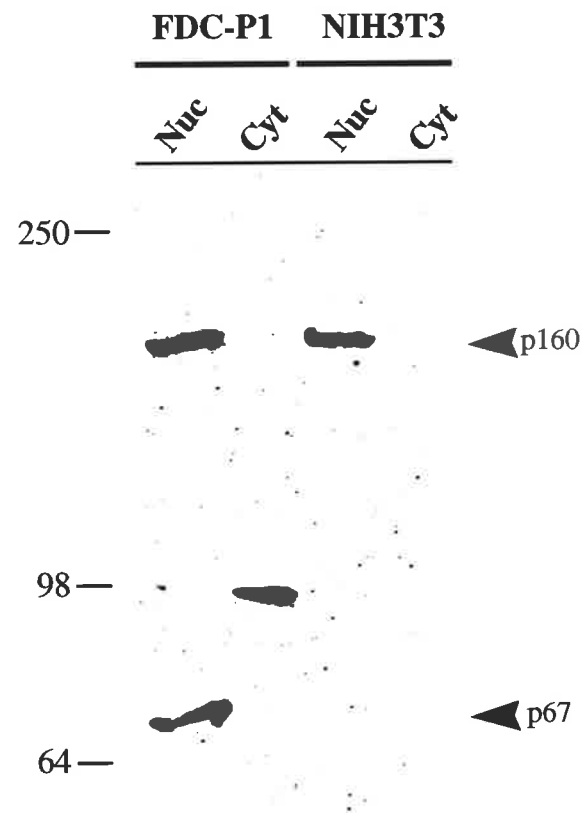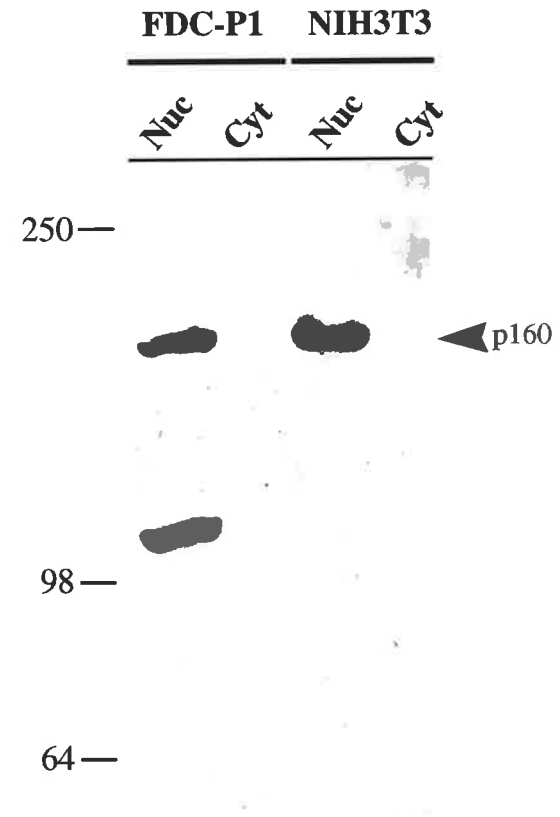
FIGURE 5.

Proteolysis of p160 by cell extracts. *In vitro*-translated p160 labelled with ^{35}S-Methionine was incubated with extracts from FDC-P1 or NIH3T3 cells for the indicated times (in minutes), following which the products were analysed by SDS-PAGE. The positions and Mr of size standards are indicated at the right, and the bands corresponding to p160 and the 67kD product (p67) are indicated at the left.
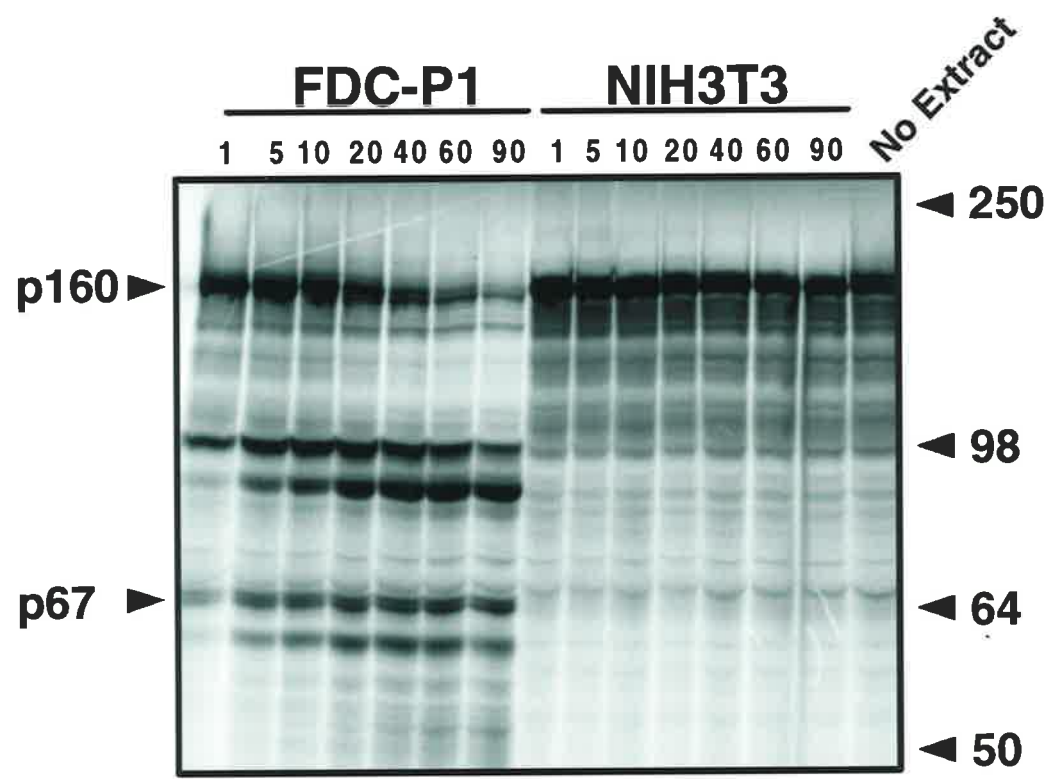
FIGURE 6.

Expression and binding properties of p160FLAG and p67*FLAG following transfection into 293T cells. In each panel, proteins were detected by immunoblotting with anti-FLAG monoclonal antibody M2. (A) SDS-PAGE of extracts from cells transfected with pact expression vectors containing p160FLAG, p67*FLAG constructs, or mock-transfected cells, as indicated. (B) Analysis of bound and unbound p160FLAG and p67*FLAG following incubation of the cell extracts shown in panel (A) with sepharose beads carrying the GST-NRD2 (lanes M), GST-L3,4P (lanes M') or GST-Jun (lanes J) fusion proteins. Note that only 50% of each unbound fraction was loaded.
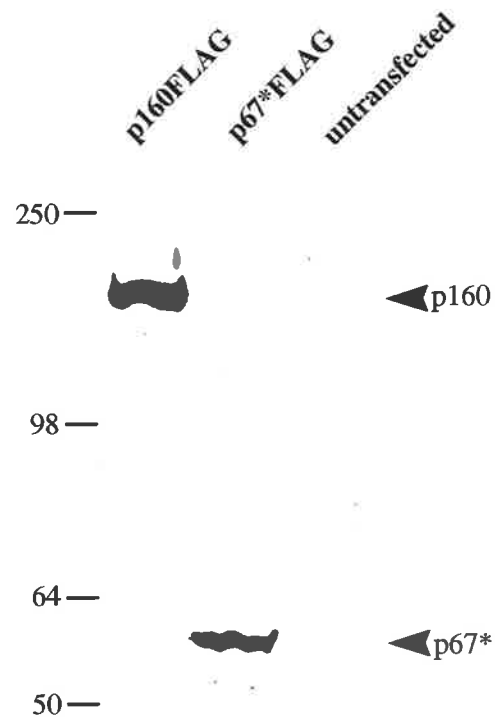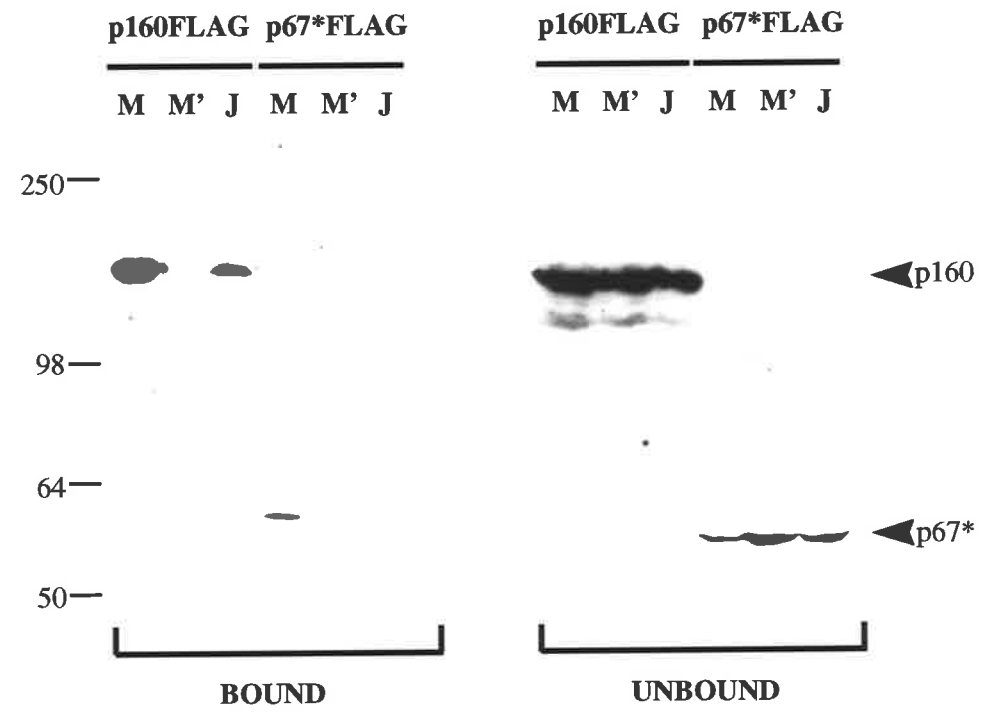
FIGURE 7.

Co-immunoprecipitation of Myb and p160. Extracts from 293T cells co-transfected with expression vectors encoding p160FLAG, wild-type Myb (Myb), L3,4P mutant Myb (L3,4PMyb) or combinations of these as indicated, were subject to immunoprecipitation with anti-FLAG antibodies. The resultant immunoprecipitates were then analysed by SDS-PAGE and immunoblotting with anti-Myb monoclonal antibody 1.1. The immunoblots were then stripped and reprobed with M2 anti-FLAG antibodies. Also shown is a direct immunoblot of a portion of each cell lysate used for immunoprecipitation ("cell lysate").

Immunoblot: M2

L3,4P Myb   p160+ L3,4P Myb   L3,4P Myb   p160+ L3,4P Myb

Myb   p160   p160+Myb   Myb   p160   p160+Myb

◄ p160FLAG

1.1

◄ Myb

Cell Lysate   **FLAG IP**
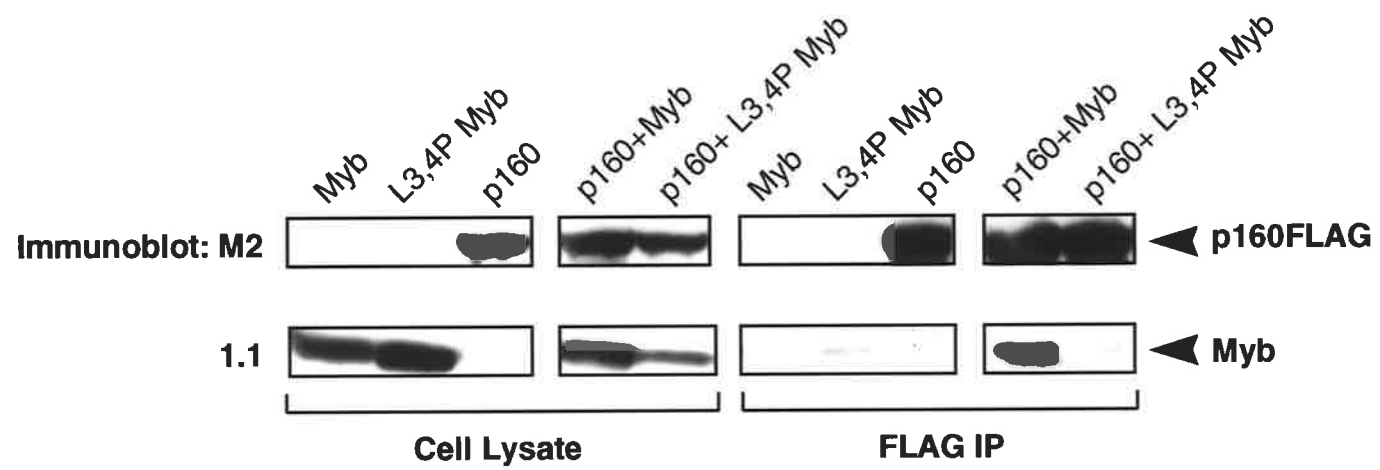
FIGURE 8.

Effect of p160 and p67* on transactivation by wild-type (WT) and L3,4P mutant (Mut) Myb. (A) CAT activity (measured by thin layer chromatography) in CV-1 cells transfected with the pc-myc-CAT reporter plasmid, pact-βgal, and with pact expression vectors encoding Myb, p160FLAG or p67*FLAG, as indicated at the bottom of the chromatograms. Where the p160FLAG or p67*FLAG vectors were included, the amount of plasmid used (in μg) is shown. The degree of activation, expressed as the ratio of CAT activity obtained in each transfection compared to that in the control transfection with no Myb vector (left-most lane), is shown above each lane; these values are normalized with respect to β-galactosidase activity to correct for differences in transfection efficiency. See Materials and Methods for details. (B) Representation of the data from (A) to indicate the degree of inhibition of Myb-stimulated transactivation by p160 and p67*; which is taken as 100% for each form of Myb (WT or Mut, as indicated). Error bars represent the standard error as determined from triplicate transfections.

A

Fold Activation

	1.0	7.2	11.5	8.1	6.7	8.0	5.7	1.4	14.1	19.1	12.9	12.5	16.7	14.3	7.3
WT Myb	-	+	+	+	+	+	+	+	-	-	-	-	-	-	-
Mut Myb	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+
μg p67*	-	-	-	-	-	1	3	5	-	-	-	-	1	3	5
μg p160	-	-	1	3	5	-	-	-	-	1	3	5	-	-	-

B

μg p67* expression plasmid

μg p160 expression plasmid
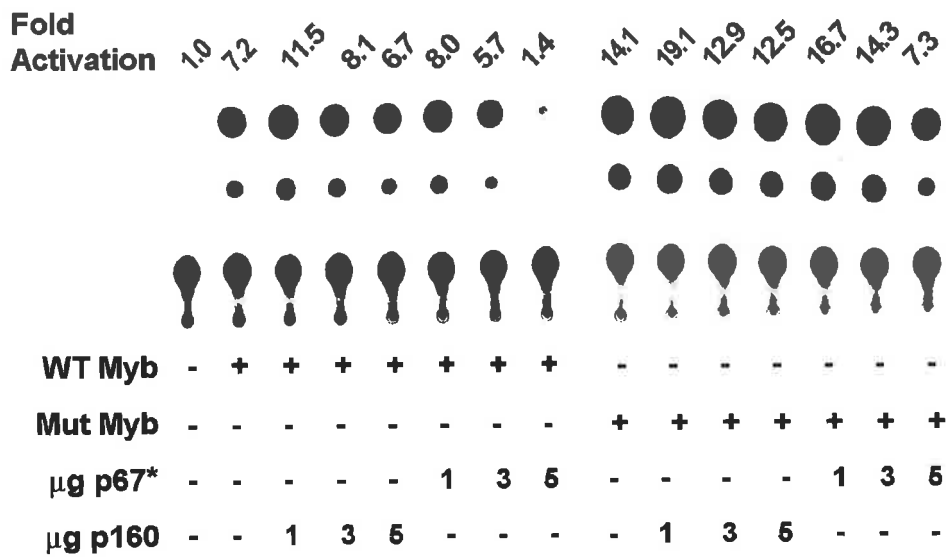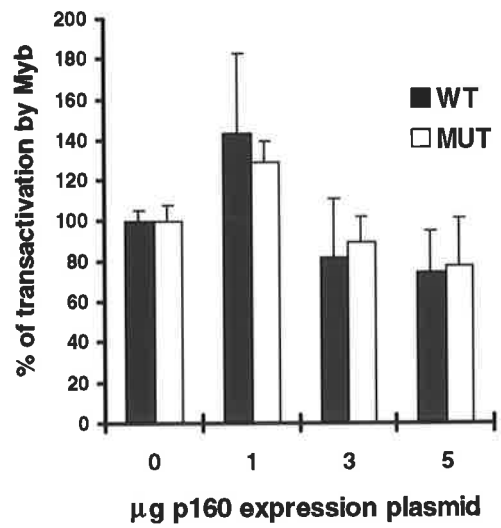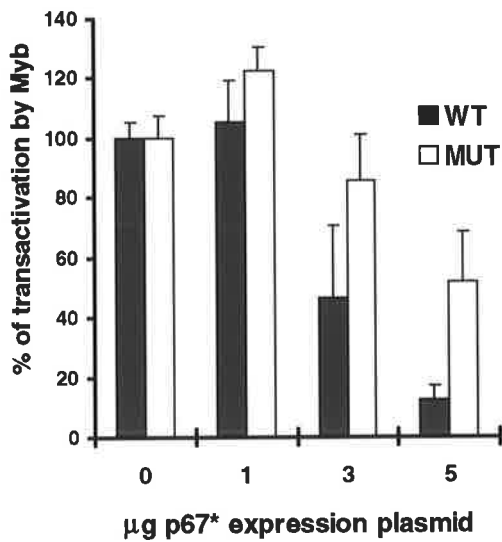
FIGURE 9.

Immunofluorescence analysis of the subcellular localization of p160. Each panel (A-D) shows phase contrast (left) and fluorescence (right) images of the same field of cells. The cells shown in (A), (B) and (C) are NIH3T3 fibroblasts and in (D), NIH3T3 fibroblasts expressing p160FLAG (see text for details). These were stained with (A), pre-immune rabbit serum; (B) 160C rabbit antiserum; (C) and (D) anti-FLAG monoclonal antibody M2.

BIBLIOGRAPHY

Anfossi, G., Gerwitz, A.M., and Calabretta, B. (1989). An oligomer complementary to c-*myb*-encoded mRNA inhibits proliferation of human myeloid leukemia cell lines. *Proc. Natl. Acad. Sci. USA.* **86**, 3379-3383.

Anton, I.A. and Frampton, J. (1988). Tryptophans in *myb* proteins. *Nature.* **336**, 719

Anzick, S.L., Kononen, J., Walker, R.L., Azorsa, D.O., Tanner, M.M., Guan, X.-Y., Sauter, G., Kallioniemi, O.-P., Trent, J.M., and Meltzer, P.S. (1997). AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer. *Science.* **277**, 965-968.

Aota, S.-I., Gojobori, T., Ishibashi, F., Maruyama, T., and Ikemura, T. (1988). Codon usage tabulated from the GenBank genetic sequence data. *Nucleic Acids Res.* **16**, r315-r402.

Arsura, M., Luchetti, M.M., Erba, E., Golay, J., Rambaldi, A., and Introna, M. (1994). Dissociation between $p^{93}B$-*myb* and $p^{75}c$-*myb* expression during the proliferation and differentiation of human myeloid cell lines. *Blood.* **83**, 1778-1790.

Auble, D.T. and Hahn, S. (1993). An ATP-dependent inhibitor of TBP binding to DNA. *Genes Dev.* **7**, 844-856.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Smith, J.A., Seidman, J.G., and Struhl, K. (1994). *Current Protocols in Molecular Biology.* (John Wiley and Sons, Inc.: New York).

Aziz, N., Wu, J., Dubendorff, J.W., Lipsick, J.S., Sturgill, T.W., and Bender, T.P. (1993). c-Myb and v-Myb are differentially phosphorylated by $p42^{mapk}$ *in vitro. Oncogene.* **8**, 2259-2265.

Aziz, N., Miglarese, M.R., Hendrickson, R.C., Shabanowitz, J., Sturgill, T.W., Hunt, D.F., and Bender, T.P. (1995). Modulation of c-Myb-induced transcription activation by a phosphorylation site near the negative regulatory domain. *Proc. Natl. Acad. Sci. USA.* **92**, 6429-6433.

Badiani, P., Corbella, P., Kioussis, D., Marvel, J., and Weston, K. (1994). Dominant interferring alleles define a role for c-Myb in T-cell development. *Genes Dev.* **8**, 770-782.

Bading, H., Beutler, C., and Moelling, K. (1989a). Mapping of a small phosphopeptide at the carboxyterminus of the viral *myb* protein by monoclonal antibodies. *Oncogene.* **4**, 33-38.

Bading, H., Rauterberg, E.W., and Moelling, K. (1989b). Distribution of c-*myc*, c-*myb*, and Ki-67 antigens in interphase and mitotic human cells evidenced by immunofluorescence staining technique. *Exp. Cell Res.* **185**, 50-59.

Bannister, A.J. and Kouzarides, T. (1996). The CBP co-activator is a histone acetyltransferase. *Nature.* **384**, 641-643.

Bender, T.P., Thompson, C.B., and Kuehl, W.M. (1987). Differential expression of c-*myb* mRNA in murine B lymphomas by a block to transcription elongation. *Science.* **237**, 1473-1476.

Bender, T.P. and Kuehl, W.M. (1986). Structure and expression of c-*myb* protooncogene mRNA in murine B-cells. *Curr. Topics Microbiol. Immunol.* **132**, 153-158.

Beug, H., von Kirchbach, A., Doderlein, G., Conscience, J., and Graf, T. (1979). Chicken hemopoietic cells transformed by seven strains of defective avian leukemia viruses display three distinct phenotypes of differentiation. *Cell.* **18**, 375-390.

Beug, H., Leutz, A., Kahn, P., and Graf, T. (1984). *Ts* mutants of E26 leukemia virus allow transformed myeloblasts, but not erythroblasts or fibroblasts, to differentiate at the nonpermissive temperature. *Cell.* **39**, 579-588.

Beug, H., Blundell, P.A., and Graf, T. (1987). Reversibility of differentiation and proliferative capacity in avian myelomonocytic cells transformed by *ts*E26 leukemia virus. *Genes Dev.* **1**, 277-286.

Biedenkapp, H., Borgmeyer, U., Sippel, A.E., and Klempnauer, K.-H. (1988). Viral *myb* oncogene encodes a sequence-specific DNA-binding activity. *Nature.* **335**, 835-837.

Bies, J., Mukhopadhyaya, R., Pierce, J., and Wolff, L. (1995). Only late, nonmitotic stages of granulocyte differentiation in 32Dcl3 cells are blocked by ectopic expression of murine c-*myb* and its truncated forms. *Cell Growth Diff.* **6**, 59-68.

Bies, J. and Wolff, L. (1997). Oncogenic activation of c-Myb by carboxyl-terminal truncation leads to decreased proteolysis by the ubiquitin-26S proteasome pathway. *Oncogene.* **14**, 203-212.

Bishop, J.M. (1985). Trends in oncogenes. *Trends in Genetics.* **1**, 245-249.

Bister, K., Nunn, M., Moscovici, C., Perbal, B., Baluda, M.A., and Duesberg, P.H. (1982). Acute leukemia viruses E26 and avian myeloblastosis virus have related transformation-specific RNA sequences but different genetic structures, gene products, and oncogenic properties. *Proc. Natl. Acad. Sci. USA.* **79**, 3677-3681.

Boise, L.H., Gorse, K.M., and Westin, E.H. (1992). Multiple mechanisms of regulation of human c-*myb* gene during myelomonocytic differentiation. *Oncogene.* **7**, 1817-1825.

Borer, R.A., Lehner, C.F., Eppenberger, H.M., and Nigg, E.A. (1989). Major nucleolar proteins shuttle between nucleus and cytoplasm. *Cell.* **56**, 379-390.

Boyle, W.J., Lipsick, J.S., and Baluda, M.A. (1984). Avian myeloblastosis virus and E26 virus oncogene products are nuclear proteins. *Proc. Natl. Acad. Sci. USA.* **81**, 4265-4269.

Boyle, W.J., Lampert, M.A., Li, A.C., and Baluda, M.A. (1985). Nuclear compartmentilization of the v-*myb* oncogene product. *Mol. Cell. Biol.* **5**, 3017-3023.

Brendeford, E.M., Myrset, A.H., Hegvold, A.B., Lundin, M., and Gabrielsen, O.S. (1997). Oncogenic point mutattions induce altered conformation, redox sensitivity, and DNA binding in the minimal DNA binding domain of avian myeloblastosis virus v-Myb. *J. Biol. Chem.* **272**, 4436-4443.

Britos-Bray, M. and Friedman, A.D. (1997). Core binding factor cannot synergistically activate the myeloperoxidase proximal enhancer in immature myeloid cells without c-Myb. *Mol. Cell. Biol.* **17**, 5127-5135.

Brown, M.S. and Goldstein, J.L. (1997). The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell.* **89**, 331-340.

Buratowski, S. (1994). The basics of basal transcription by RNA polymerase II. *Cell.* **77**, 1-3.

Burgess, T.L., Fisher, E.F., Ross, S.L., Bready, J.V., Qian, Y.-X., Bayewitch, L.A., Cohen, A.M., Herrera, C.J., Hu, S.S.-F., Kramer, T.B., Lott, F.D., Martin, F.H., Pierce, G.F., Simonet, L., and Farrell, C.L. (1995). The antiproliferative activity of c-*myb* and c-*myc* antisense oligonucleotides in smooth muscle cells is caused by a nonantisense mechanism. *Proc. Natl. Acad. Sci. USA.* **92**, 4051-4055.

Burk, O., Mink, S., Ringwald, M., and Klempnauer, K.-H. (1993). Synergistic activation of the chicken *mim-1* gene by v-*myb* and C/EBP transcription factors. *EMBO J.* **12**, 2027-2038.

Burk, O., Worpenberg, S., Haenig, B., and Klempnauer, K.-H. (1997). *tom-1*, a novel v-Myb target gene expressed in AMV- and E26-transformed myelomonocytic cells. *EMBO J.* **16**, 1371-1380.

Calabretta, B., Sims, R.B., Valtieri, M., Caracciolo, D., Szczylik, C., Venturelli, D., Ratajczak, M., Beran, M., and Gerwitz, A.M. (1991). Normal and leukemic hematopoietic cells manifest differential sensitivity to inhibitory effects of c-*myb* antisense oligodeoxynucleotides: an *in vitro* study relevant to bone marrow purging. *Proc. Natl. Acad. Sci. USA.* **88**, 2351-2355.

Caracciolo, D., Venturelli, D., Valtieri, M., Peschle, C., Gerwitz, A.M., and Calabretta, B. (1990). Stage-related proliferative activity determines c-*myb* functional requirements during normal human hematopoiesis. *J. Clin. Invest.* **85**, 55-61.

Catron, K.M., Purkerson, J.M., Isakson, P.C., and Bender, T.P. (1992). Constitutive versus cell cycle regulation of c-*myb* mRNA expression correlates with developmental stages in murine B lymphoid tumours. *J. Immunology.* **148**, 934-942.

Clarke, M.F., Kukowska-Latallo, J.F., Westin, E., Smith, M., and Prochownik, E.V. (1988). Constitutive expression of a c-*myb* cDNA blocks Friend murine erythroleukemia cell differentiation. *Mol. Cell. Biol.* **8**, 884-892.

Cleveland, D.W. (1983). Peptide mapping in one dimension by limited proteolysis of sodium dodecyl sulfate-solubilized proteins. *Methods in Enzymology.* **96**, 222-229.

Cogswell, J.P., Cogswell, P.C., Kuehl, W.M., Cuddihy, A.M., Bender, T.M., Engelke, U., Marcu, K.B., and Ting, J.P.-Y. (1993). Mechanism of c-*myc* regulation by c-Myb in different cell lineages. *Mol. Cell. Biol.* **13**, 2858-2869.

Dai, P., Akimaru, H., Tanaka, Y., Hou, D.-X., Yasukawa, T., Kanei-Ishii, C., Takahashi, T., and Ishii, S. (1996). CBP as a transcriptional coactivator of c-Myb. *Genes Dev.* **10**, 528-540.

Danish, R., El-Awar, O., Weber, B.L., Langmore, J., Turka, L.A., Ryan, J.J., and Clarke, M.F. (1992). c-*myb* effects on kinetic events during MEL cell differentiation. *Oncogene.* **7**, 901-907.

Dash, A.B., Orrico, F.C., and Ness, S.A. (1996). The EVES motif mediates both intermolecular and intramolecular regulation of c-Myb. *Genes Dev.* **10**, 1858-1869.

Deng, Q.-L., Ishii, S., and Sarai, A. (1996). Binding site analysis of c-Myb: screening of potential binding sites by using the mutation matrix derived from systematic binding affinity measurements. *Nucleic Acids Res.* **24**, 766-774.

Dingwall, C. and Laskey, R.A. (1991). Nuclear targeting sequences-a consensus? *Trends Biochem. Sci.* **16**, 478-481.

Dini, P.W., Eltman, J.T., and Lipsick, J.S. (1995). Mutations in the DNA-binding and transcriptional activation domains of v-Myb cooperate in transformation. *J. Virology.* **69**, 2515-2524.

Dini, P.W. and Lipsick, J.S. (1993). Oncogenic truncation of the first repeat of c-Myb decreases DNA binding in vitro and in vivo. *Mol. Cell. Biol.* **13**, 7334-7348.

Dubendorff, J.W., Whittaker, L.J., Eltman, J.T., and Lipsick, J.S. (1992). Carboxyl-terminal elements of c-Myb negatively regulate transcriptional activation in *cis* and in *trans*. *Genes Dev.* **6**, 2524-2535.

DuBridge, R.B., Tang, P., Hsia, H.C., Leong, P.-M., Miller, J.H., and Calos, M.P. (1987). Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Mol. Cell. Biol.* **7**, 379-387.

Dudek, H., Tantravahi, R.V., Rao, V.N., Reddy, S.P., and Reddy, E.P. (1992). Myb and Ets proteins cooperate in transcriptional activation of the *mim-1* promoter. *Proc. Natl. Acad. Sci. USA.* **89**, 1291-1295.

Ebneth, A., Adermann, K., and Wolfes, H. (1994). Does a synthetic peptide containing the leucine-zipper domain of c-myb form an a-helical structure in solution? *FEBS Letters.* **337**, 265-268.

Engelke, U., Whittaker, L., and Lipsick, J.S. (1995). Weak transcriptional activation is sufficient for transformation by v-Myb. *Virology.* **208**, 467-477.

Engelke, U. and Lipsick, J.S. (1994). Transformation of myelomonocytic cells by the avian myeloblastosis virus is determined by the v-*myb* oncogene, not by the long terminal repeats of the virus. *J. Virology.* **68**, 2752-2755.

Evans, J.L., Moore, T.L., Kuehl, W.M., Bender, T., and Ting, J.P.-Y. (1990). Functional analysis of c-Myb protein in T-lymphocytic cell lines shows that it *trans*-activates the c-*myc* promoter. *Mol. Cell. Biol.* **10**, 5747-5752.

Favier, D. and Gonda, T.J. (1994). Detection of proteins that bind to the leucine zipper motif of c-Myb. *Oncogene.* **9**, 305-311.

Feinberg, A.P. and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry.* **132**, 6-13.

Ferrao, P., Macmillan, E.M., Ashman, L.K., and Gonda, T.J. (1995). Enforced expression of full length c-Myb leads to density-dependent transformation of murine haemopoietic cells. *Oncogene.* **11**, 1631-1638.

Finco, T.S. and Baldwin, A.S. (1995). Mechanistic aspects of NF-κB regulation: the emerging role of phosphorylation and proteolysis. *Immunity.* **3**, 263-272.

Fischer, U., Huber, J., Boelens, W.C., Mattaj, I.W., and Luhrmann, R. (1995). The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell.* **82**, 475-483.

Foos, G., Grimm, S., and Klempnauer, K.-H. (1992). Functional antagonism between members of the *myb* family: B-*myb* inhibits v-*myb*-induced gene activation. *EMBO J.* **11**, 4619-4629.

Foos, G., Grimm, S., and Klempnauer, K.-H. (1994). The chicken A-*myb* protein is a transcriptional activator. *Oncogene.* **9**, 2481-2488.

Frampton, J., Leutz, A., Gibson, T.J., and Graf, T. (1989). DNA-binding domain ancestry. *Nature.* **342**, 134

Frampton, J., Kouzarides, T., Doderlain, G., Graf, T., and Weston, K. (1993). Influence of the v-Myb transactivation domain on the oncoprotein's transformation specificity. *EMBO J.* **12**, 1333-1341.

Frampton, J., Ramqvist, T., and Graf, T. (1996). v-Myb of E26 leukemia virus up-regulates *bcl-2* and suppresses apoptosis in myeloid cells. *Genes Dev.* **10**, 2720-2731.

Fridell, R.A., Fischer, U., Luhrmann, R., Meyer, B.E., Meinkoth, J.L., Malim, M.H., and Cullen, B.R. (1996). Amphibian transcription factor IIIA proteins contain a sequence element functionally equivalent to the nuclear export signal of human immunodeficiency virus type 1 Rev. *Proc. Natl. Acad. Sci. USA.* **93**, 2936-2940.

Frohman, M.A., Dush, M.K., and Martin, G.R. (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific primer. *Proc. Natl. Acad. Sci. USA.* **85**, 8998-9002.

Fu, S.-L. and Lipsick, J.S. (1996). FAETL motif required for leukemic transformation by v-Myb. *J. Virology.* **70**, 5600-5610.

Fu, S.-L. and Lipsick, J.S. (1997). Constitutive expression of full-length c-Myb transforms avian cells characteristic of both monocytic and granulocytic lineages. *Cell Growth Diff.* **8**, 35-45.

Garcia, A., LaMontagne, K., Reavis, D., Stober-Gasser, U., and Lipsick, J.S. (1991). Determinants of sequence-specific DNA-binding by p48^{v-myb}. *Oncogene.* **6**, 265-273.

Garcia-Bustos, J., Heitman, J., and Hall, M.N. (1991). Nuclear protein localization. *Biochim. Biophys. Acta.* **1071**, 83-101.

Gardner, J.M. (1996). *An Investigation of the Properties of p160: Nuclear Localisation and Binding to c-Myb and c-Jun.* Honours Thesis. University of Adelaide, South Australia.

Gerace, L. (1995). Nuclear export signals and the fast track to the cytoplasm. *Cell.* **82**, 341-344.

Gerondakis, S. and Bishop, J.M. (1986). Structure of the protein encoded by the chicken proto-oncogene c-*myb*. *Mol. Cell. Biol.* **6**, 3677-3684.

Gerwitz, A.M., Anfossi, G., Venturelli, D., Valpreda, S., Sims, R., and Calabretta, B. (1989). G$_1$/S transition in normal human T-lymphocytes requires the nuclear protein encoded by c-*myb*. *Science.* **245**, 180-183.

Gerwitz, A.M. and Calabretta, B. (1988). A c-*myb* antisense oligodeoxynucleotide inhibits normal human hematopoiesis *in vitro*. *Science.* **242**, 1303-1306.

Golay, J., Capucci, A., Arsura, M., Castellano, M., Rizzo, V., and Introna, M. (1991). Expression of c-*myb* and B-*myb*, but not A-*myb*, correlates with proliferation in human hemopoietic cells. *Blood.* **77**, 149-158.

Gonda, T.J., Sheiness, D.K., and Bishop, J.M. (1982). Transcripts from the cellular homologs of retroviral oncogenes: distribution among chicken tissues. *Mol. Cell. Biol.* **6**, 617-624.

Gonda, T.J., Gough, N.M., Dunn, A.R., and de Blaquiere, J. (1985). Nucleotide sequence of cDNA clones of the murine *myb* proto-oncogene. *EMBO J.* **4**, 2003-2008.

Gonda, T.J., Buckmaster, C., and Ramsay, R.G. (1989a). Activation of c-*myb* by carboxy-terminal truncation: relationship to transformation of murine haemopoietic cells *in vitro*. *EMBO J.* **8**, 1777-1783.

Gonda, T.J., Ramsay, R.G., and Johnson, G.R. (1989b). Murine myeloid cell lines derived by *in vitro* infection with recombinant c-*myb* retroviruses express *myb* from rearranged vector proviruses. *EMBO J.* **8**, 1767-1775.

Gonda, T.J. (1991). Molecular and cellular activities of Myb: regulation in normal hemopoiesis and transformation. *Seminars in Virology.* **2**, 351-361.

Gonda, T.J., Macmillan, E.M., Townsend, P.V., and Hapel, A.J. (1993). Differentiation state and responses to hematopoietic growth factors of murine myeloid cells transformed by *myb*. *Blood*. **82**, 2813-2822.

Gonda, T.J., Favier, D., Ferrao, P., Macmillan, E.M., Simpson, R., and Tavner, F. (1996). The c-Myb negative regulatory domain. *Curr. Topics Microbiol. Immunol*. **211**, 99-109.

Gonda, T.J. and Metcalf, D. (1984). Expression of *myb*, *myc* and *fos* proto-oncogenes during the differentiation of a murine myeloid leukaemia. *Nature*. **310**, 249-251.

Gonzalez, F.A., Raden, D.L., and Davis, R.J. (1991). Identification of substrate recognition determinants for human ERK1 and ERK2 protein kinases. *J. Biol. Chem*. **266**, 22159-22163.

Grasser, F.A., Graf, T., and Lipsick, J.S. (1991). Protein truncation is required for the activation of the c-*myb* proto-oncogene. *Mol. Cell. Biol*. **11**, 3987-3996.

Griffin, C.A. and Baylin, S.B. (1985). Expression of the c-*myb* oncogene in human small cell lung carcinoma. *Cancer Res*. **45**, 272-275.

Guehmann, S., Vorbrueggen, G., Kalkbrenner, F., and Moelling, K. (1992). Reduction of a conserved Cys is essential for Myb DNA-binding. *Nucleic Acids Res*. **20**, 2279-2286.

Guerra, J., Withers, D.A., and Boxer, L.M. (1995). MYB binding sites mediate negative regulation of c-*MYB* expression in T-cell lines. *Blood*. **86**, 1873-1880.

Harlow, E., Whyte, P., Franza, B.R., and Schely, C. (1986). Association of adenovirus early-region 1A proteins with cellular polypeptides. *Mol. Cell. Biol*. **6**, 1579-1589.

Harlow, E. and Lane, D.P. (1988). *Antibodies: A Laboratory Manual*. (Cold Spring Harbor Laboratory Press: Cold Spring Harbor Laboratory, New York).

Heery, D.M., Kalkhoven, E., Hoare, S., and Parker, M.G. (1997). A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature*. **387**, 733-736.

Hengen, P.N. (1996). Purification of GST fusion proteins. *Trends Biochem. Sci*. **21**, 400-401.

Hernandez-Munain, C. and Krangel, M.S. (1994). Regulation of the T-cell receptor δ enhancer by functional cooperation between c-Myb and core-binding factors. *Mol. Cell. Biol*. **14**, 473-483.

Hernandez-Munain, C. and Krangel, M.S. (1995). c-Myb and core-binding factor/PEBP2 display functional synery but bind independently to adjacent sites in the T-cell receptor δ enhancer. *Mol. Cell. Biol*. **15**, 3090-3099.

Hogg, A., Schirm, S., Nakagoshi, H., Bartley, P., Ishii, S., Bishop, J.M., and Gonda, T.J. (1997). Inactivation of a c-Myb/estrogen receptor fusion protein in transformed primary cells leads to granulocyte/macrophage differentiation and down regulation of c-*kit* but not c-*myc* or *cdc2*. *Oncogene*. **15**, 2885-2898.

Howe, K.M., Reakes, C.F.L., and Watson, R.J. (1990). Characterization of the sequence-specific interaction of mouse c-*myb* protein with DNA. *EMBO J.* **9**, 161-169.

Howe, K.M. and Watson, R.J. (1991). Nucleotide preferences in sequence-specific recognition of DNA by c-*myb* protein. *Nucleic Acids Res.* **19**, 3913-3919.

Hu, Y., Ramsay, R.G., Kanei-Ishii, C., Ishii, S., and Gonda, T.J. (1991). Transformation by carboxyl-deleted Myb reflects increased transactivating capacity and disruption of a negative regulatory domain. *Oncogene.* **6**, 1549-1553.

Hunter, T. and Karin, M. (1992). The regulation of transcription by phosphorylation. *Cell.* **70**, 375-387.

Ibanez, C.E. and Lipsick, J.S. (1990). *Trans* activation of gene expression by v-*myb*. *Mol. Cell. Biol.* **10**, 2285-2293.

Introna, M., Golay, J., Frampton, J., Nakano, T., Ness, S.A., and Graf, T. (1990). Mutations in v-*myb* alter the differentiation of myelomonocytic cells transformed by the oncogene. *Cell.* **63**, 1287-1297.

Jacobs, S.M., Gorse, K.M., and Westin, E.H. (1994). Identification of a second promoter in the human c-*myb* proto-oncogene. *Oncogene.* **9**, 227-235.

Jans, D.A., Ackermann, M., Bischoff, J.R., Beach, D.H., and Peters, R. (1991). p34^{cdc2}-mediated phosphorylation at T^{124} inhibits nuclear import of SV40 T-antigen proteins. *J. Cell Biol.* **115**, 1203-1212.

Jans, D.A. and Hubner, S. (1996). Regulation of protein transport to the nucleus: central role of phosphorylation. *Physiological Reviews.* **76**, 651-685.

Jantzen, H.-M., Admon, A., Bell, S.P., and Tjian, R. (1990). Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. *Nature.* **344**, 830-836.

Kalkbrenner, F., Guehmann, S., and Moelling, K. (1990). Transcriptional activation by human c-*myb* and v-*myb* genes. *Oncogene.* **5**, 657-661.

Kamei, Y., Xu, L., Heinzel, T., Torchia, J., Kuokawa, R., Gloss, B., Lin, S.-C., Heyman, R.A., Rose, D.W., Glass, C.K., and Rosenfeld, M.G. (1996). A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. *Cell.* **85**, 403-414.

Kanei-Ishii, C., Sarai, A., Sawazaki, T., Nakagoshi, H., He, D.-N., Ogata, K., Nishimura, Y., and Ishii, S. (1990). The tryptophan cluster: a hypothetical structure of the DNA-binding domain of the *myb* protooncogene product. *J. Biol. Chem.* **265**, 19990-19995.

Kanei-Ishii, C., Macmillan, E.M., Nomura, T., Sarai, A., Ramsay, R.G., Aimoto, S., Ishii, S., and Gonda, T.J. (1992). Transactivation and transformation by Myb are negatively regulated by a leucine-zipper structure. *Proc. Natl. Acad. Sci. USA.* **89**, 3088-3092.

Kanei-Ishii, C., Yasukawa, T., Morimoto, R.I., and Ishii, S. (1994). c-Myb-induced *trans*-activation mediated by heat shock elements without sequence-specific DNA binding of c-Myb. *J. Biol. Chem.* **269**, 15768-15775.

Kanei-Ishii, C., Tanikawa, J., Nakai, A., Morimoto, R.I., and Ishii, S. (1997). Activation of heat shock transcription factor 3 by c-Myb in the absence of cellular stress. *Science.* **277**, 246-248.

Karin, M. (1994). Signal transduction from the cell surface to the nucleus through the phosphorylation of transcription factors. *Curr. Opin. Cell Biol.* **6**, 415-424.

Kennelly, P.J. and Krebs, E.G. (1991). Consensus sequences as substrate specificity determinants for protein kinases and protein phosphatases. *J. Biol. Chem.* **266**, 15555-15558.

Kenney, S.C., Holley-Guthrie, E., Quinlivan, E.B., Gutsch, D., Zhang, Q., Bender, T., Giot, J.-F., and Sergeant, A. (1992). The cellular oncogene c-*myb* can interact synergistically with the Epstein-Barr virus BZLF1 transactivator in lymphoid cells. *Mol. Cell. Biol.* **12**, 136-146.

Kermekchiev, M., Workman, J.L., and Pikaard, C.S. (1997). Nucleosome binding by the polymerase I transactivator upstream binding factor displaces linker histone H1. *Mol. Cell. Biol.* **17**, 5833-5842.

Klempnauer, K.-H., Gonda, T.J., and Bishop, J.M. (1982). Nucleotide sequence of the retroviral leukemia gene v-*myb* and its cellular progenitor c-*myb*: the architecture of a transduced oncogene. *Cell.* **31**, 453-463.

Klempnauer, K.-H., Ramsay, G., Bishop, J.M., Moscovici, M.G., Moscovici, C., McGrath, J.P., and Levinson, A.D. (1983). The product of the retroviral transforming gene v-myb is a truncated version of the protein encoded by the cellular oncogene c-*myb*. *Cell.* **33**, 345-355.

Klempnauer, K.-H., Symonds, G., Evan, G.I., and Bishop, J.M. (1984). Subcellular loclalization of proteins encoded by oncogenes of avian myeloblastosis virus and avian leukemia virus E26 and by the chicken c-*myb* gene. *Cell.* **37**, 537-547.

Klempnauer, K.-H., Bonifer, C., and Sippel, E. (1986). Identification and characterization of the protein encoded by the human c-*myb* proto-oncogene. *EMBO J.* **5**, 1903-1911.

Koleske, A.J. and Young, R.A. (1995). The RNA polymerase II holoenzyme and its implications for gene regulation. *Trends Biochem. Sci.* **20**, 113-116.

Kowenz-Leutz, E., Herr, P., Niss, K., and Leutz, A. (1997). The homeobox gene *GBX2*, a target of the *myb* oncogene, mediates autocrine growth and monocyte differentiation. *Cell.* **91**, 185-195.

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

Krieg, J., Oelgeschlager, M., Janknecht, R., and Luscher, B. (1995). High affinity DNA binding of native full length c-Myb and differential proteolytic sensitivity of its N- and C-terminal domains. *Oncogene.* **10**, 2221-2228.

Krylov, D., Mikhailenko, I., and Vinson, C. (1994). A thermodynamic scale for leucine zipper stability and dimerization specificity: e and g interhelical interactions. *EMBO J.* **13**, 2849-2861.

Ku, D.-H., Wen, S.-C., Engelhard, A., Nicolaides, N.C., Lipson, K.E., Marino, T.A., and Calabretta, B. (1993). c-*myb* transactivates cdc2 expression via Myb binding sites in the 5'-flanking region of the human *cdc2* gene. *J. Biol. Chem.* **268**, 2255-2259.

Kuhn, A., Voit, R., Stefanovsky, V., Evers, R., Bianchi, M., and Grummt, I. (1994). Functional differences between the two splice variants of the nucleolar transcription factor UBF: the second HMG box determines specificity of DNA binding and transcriptional activity. *EMBO J.* **13**, 416-424.

Kuhn, A. and Grummt, I. (1992). Dual role of the nucleolar transcription factor UBF: trans-activator and antirepressor. *Proc. Natl. Acad. Sci. USA.* **89**, 7340-7344.

Kumar, S. and Harvey, N.L. (1995). Role of multiple cellular proteases in the execution of programmed cell death. *FEBS Letters.* **375**, 169-173.

Laget, M.-P., Callebaut, I., de Launoit, Y., Stehelin, D., and Mornon, J.-P. (1993). Predicted common structural features of DNA-binding domains from Ets, Myb and HMG transcription factors. *Nucleic Acids Res.* **21**, 5987-5996.

Landschulz, W.H., Johnson, P.F., and McKnight, S.L. (1988). The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science.* **240**, 1759-1764.

Lane, S., Farlie, P., and Watson, R. (1997). B-Myb function can be markedly enhanced by cyclin A-dependent kinase and protein truncation. *Oncogene.* **14**, 2445-2453.

Lane, T., Ibanez, C., Garcia, A., Graf, T., and Lipsick, J. (1990). Transformation by v-*myb* correlates with *trans*-activation of gene expression. *Mol. Cell. Biol.* **10**, 2591-2598.

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

Lin, H.H., Sternfeld, D.C., Shinpock, S.G., Popp, R.A., and Mucenski, M.L. (1996). Functional analysis of the c-*myb* proto-oncogene. *Curr. Topics Microbiol. Immunol.* **211**, 79-87.

Lin, H.H., Stubbs, L., J., and Mucenski, M.L. (1997). Identification and characterization of a seven transmembrane hormone receptor using differential display. *Genomics.* **41**, 301-308.

Lipsick, J.S. (1996). One billion years of Myb. *Oncogene.* **13**, 223-235.

Lobb, R.R. and Hemler, M.E. (1994). The pathophysiologic role of a4 integrins *in vivo*. *J. Clin. Invest.* **94**, 1722-1728.

Lupas, A. (1996). Coiled coils: new structures and new functions. *Trends Biochem. Sci.* **21**, 375-382.

Luscher, B., Christenson, E., Litchfield, D.W., Krebs, E.G., and Eisenman, R.N. (1990). Myb DNA binding inhibited by phosphorylation at a site deleted during oncogenic activation. *Nature.* **344**, 517-522.

Luscher, B. and Eisenman, R.N. (1992). Mitosis-specific phosphorylation of the nuclear oncoproteins Myc and Myb. *J. Cell Biol.* **118**, 775-784.

Macmillan, E.M. and Gonda, T.J. (1994). Murine myeloid cells transformed by *myb* require fibroblast-derived or autocrine growth factors in addition to granulocyte-macrophage colony-stimulating factor for proliferation. *Blood.* **83**, 209-216.

Majello, B., Kenyon, L.C., and Dalla-Favera, R. (1986). Human c-*myb* protooncogene: nucleotide sequence of cDNA and organization of the genomic locus. *Proc. Natl. Acad. Sci. USA.* **83**, 9636-9640.

Marcand, S., Gasser, S.M., and Gilson, E. (1996). Chromatin: a sticky silence. *Current Biol.* **6**, 1222-1225.

McCann, S., Sullivan, J., Guerra, J., Arcinas, M., and Boxer, L.M. (1995). Repression of the c-*myb* gene by WT1 protein in T and B cell lines. J. *Biol. Chem.* **270**, 23785-23789.

McClinton, D., Stafford, J., Brents, L., Bender, T.P., and Kuehl, W.M. (1990). Differentiation of mouse erythroleukemia cells is blocked by late up-regulation of a c-*myb* transgene. *Mol. Cell. Biol.* **10**, 705-710.

McCracken, S., Leung, S., Bosselut, R., Ghysdael, J., and Miyamoto, N.G. (1994). Myb and Ets related transcription factors are required for activity of the human *lck* type I promoter. *Oncogene.* **9**, 3609-3615.

Meier, U.T. and Blobel, G. (1992). Nopp140 shuttles on tracks between nucleolus and cytoplasm. *Cell.* **70**, 127-138.

Melotti, P., Ku, D.-H., and Calabretta, B. (1994). Regulation of the expression of the hematopoietic stem cell antigen CD34: role of c-*myb*. *J. Exp. Med.* **179**, 1023-1028.

Melotti, P. and Calabretta, B. (1994). Ets-2 and c-Myb act independently in regulating expression of the hematopoietic stem cell antigen CD34. *J. Biol. Chem.* **269**, 25303-25309.

Melotti, P. and Calabretta, B. (1996). Induction of hematopoietic commitment and erythromyeloid differentiation in embryonal stem cells constitutively expressing c-*myb*. *Blood.* **87**, 2221-2234.

Miau, L.-H., Chang, C.-J., Tsai, W.-H., and Lee, S.-C. (1997). Identification and characterization of a nucleolar phosphoprotein, Nopp140, as a transcription factor. *Mol. Cell. Biol.* **17**, 230-239.

Miglarese, M.R., Richardson, A.F., Aziz, N., and Bender, T.P. (1996). Differential regulation of c-Myb-induced transcription activation by a phosphorylation site in the negative regulatory domain. *J. Biol. Chem.* **271**, 22697-22705.

Mink, S., Kerber, U., and Klempnauer, K.-H. (1996). Interaction of C/EBPb and v-Myb is required for synergistic activation of the *mim-1* gene. *Mol. Cell. Biol.* **16**, 1316-1325.

Mitchell, P.J. and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science.* **243**, 371-378.

Mizuguchi, G., Kanei-Ishii, C., Takahashi, T., Yasukawa, T., Nagase, T., Horikoshi, M., Yamamoto, T., and Ishii, S. (1995). c-Myb repression of *c-erbB-2* transcription by direct binding to the *c-erbB-2* promoter. *J. Biol. Chem.* **270**, 9384-9389.

Moscovici, C., Samarut, J., Gazzolo, L., and Moscovici, G. (1981). Myeloid and erythroid neoplastic responses to avian defective leukemia viruses in chickens and in quail. *Virology.* **113**, 765-768.

Moscovici, M.G. and Moscovici, C. (1983). Isolation and characterization of a temperature-sensitive mutant of avian myeloblastosis virus. *Proc. Natl. Acad. Sci. USA.* **80**, 1421-1425.

Mucenski, M.L., McLain, K., Kier, A.B., Swerdlow, S.H., Schreiner, C.M., Miller, T.A., Pietryga, D.W., Scott, W.J., and Potter, S.S. (1991). A functional c-*myb* gene is required for normal murine fetal hepatic hematopoiesis. *Cell.* **65**, 677-689.

Myrset, A.H., Bostad, A., Jamin, N., Lirsac, P.-N., Toma, F., and Gabrielsen, O.S. (1993). DNA and redox state induced conformational changes in the DNA-binding domain of the Myb oncoprotein. *EMBO J.* **12**, 4625-4633.

Nakagoshi, H., Nagase, T., Ueno, Y., and Ishii, S. (1989). Transcriptional *trans*-repression by the c-*myb* proto-oncogene product. *Nucleic Acids Res.* **17**, 7315-7324.

Nakagoshi, H., Nagase, T., Kanei-Ishii, C., Ueno, Y., and Ishii, S. (1990). Binding of the c-*myb* proto-oncogene product to the simian virus 40 enhancer stimulates transcription. *J. Biol. Chem.* **265**, 3479-3483.

Nakagoshi, H., Kanei-Ishii, C., Sawazaki, T., Mizuguchi, G., and Ishii, S. (1992). Transcriptional activation of the c-*myc* gene by the c-*myb* and B-*myb* gene products. *Oncogene.* **7**, 1233-1240.

Nakagoshi, H., Takemoto, Y., and Ishii, S. (1993). Functional domains of the human B-*myb* gene product. *J. Biol. Chem.* **268**, 14161-14167.

Ness, S.A., Marknell, A., and Graf, T. (1989). The v-*myb* oncogene product binds to and activates the promyelocyte-specific *mim-1* gene. *Cell.* **59**, 1115-1125.

Ness, S.A., Kowenz-Leutz, E., Casini, T., Graf, T., and Leutz, A. (1993). Myb and NF-M: combinatorial activators of myeloid genes in heterologous cell types. *Genes Dev.* **7**, 749-759.

Ness, S.A. (1996). The Myb oncoprotein: regulating a regulator. Biochim. Biophys. Acta. **1288**, F123-F139.

Nicolaides, N.C., Gualdi, R., Casadevall, C., Manzella, L., and Calabretta, B. (1991). Positive autoregulation of c-*myb* expression via Myb binding sites in the 5' flanking region of the human c-*myb* gene. *Mol. Cell. Biol.* **11**, 6166-6176.

Nigg, E.A. (1997). Nucleocytoplasmic transport: signals, mechanisms and regulation. *Nature.* **386**, 779-787.

Nishina, Y., Nakagoshi, H., Imamoto, F., Gonda, T.J., and Ishii, S. (1989). *Trans*-activation by the c-*myb* proto-oncogene. *Nucleic Acids Res.* **17**, 107-117.

Nomura, N., Takahashi, M., Matsui, M., Ishii, S., Date, T., Sasamoto, S., and Ishizaki, R. (1988). Isolation of human cDNA clones of *myb*-related genes, A-*myb* and B-*myb*. *Nucleic Acids Res.* **16**, 11075-11089.

Nomura, T., Sakai, N., Sarai, A., Sudo, T., Kanei-Ishii, C., Ramsay, R.G., Favier, D., Gonda, T.J., and Ishii, S. (1993). Negative autoregulation of c-Myb activity by homodimer formation through the leucine zipper. *J. Biol. Chem.* **268**, 21914-21923.

Nuchprayoon, I., Simkevich, C.P., Luo, M., Friedman, A.D., and Rosmarin, A.G. (1997). GABP cooperates with c-Myb and C/EBP to activate the neutrophil elastase promoter. *Blood.* **89**, 4546-4554.

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

Nunn, M.F. and Hunter, T. (1989). The *ets* sequence is required for induction of erythroblastosis in chickens by avian retrovirus E26. *J. Virology.* **63**, 398-402.

O'Shea, E.K., Rutkowski, R., and Kim, P.S. (1989). Evidence that the leucine zipper is a coiled coil. *Science.* **243**, 538-542.

Oelgeschlager, M., Krieg, J., Luscher-Firzlaff, J.M., and Luscher, B. (1995). Casein kinase II phosphorylation site mutations in c-Myb affect DNA binding and transcriptional cooperativity with NF-M. *Mol. Cell. Biol.* **15**, 5966-5974.

Oelgeschlager, M., Janknecht, R., Krieg, J., Schreek, S., and Luscher, B. (1996). Interaction of the co-activator CBP with Myb proteins: effects on Myb-specific transactivation and on the cooperativity with NF-M. *EMBO J.* **15**, 2771-2780.

Ogata, K., Hojo, H., Aimoto, S., Nakai, T., Nakamura, H., Sarai, A., Ishii, S., and Nishimura, Y. (1992). Solution structure of a DNA-binding unit of Myb: a helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core. *Proc. Natl. Acad. Sci. USA.* **89**, 6428-6432.

Ogata, K., Morikawa, S., Nakamura, H., Sekikawa, A., Inoue, T., Kanai, H., Sarai, A., Ishii, S., and Nishimura, Y. (1994). Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. *Cell.* **79**, 639-648.

Ogata, K., Morikawa, S., Nakamura, H., Hojo, H., Yoshimura, S., Zhang, R., Aimoto, S., Ametani, Y., Hirata, Z., Sarai, A., Ishii, S., and Nishimura, Y. (1995). Comparison of the free and DNA-complexed forms of the DNA-binding domain from c-Myb. *Nature Struct. Biol.* **2**, 309-319.

Ogata, K., Kanei-Ishii, C., Sasaki, M., Hatanaka, H., Nagadoi, A., Enari, M., Nakamura, H., Nishimura, Y., Ishii, S., and Sarai, A. (1996). The cavity in the hydrophobic core of Myb DNA-binding domain is reserved for DNA recognition and *trans*-activation. *Nature Struct. Biol.* **3**, 178-187.

Ogryzko, V.V., Schiltz, R.L., Russanova, V., Howard, B.H., and Nakatani, Y. (1996). The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell.* **87**, 953-959.

Onate, S.A., Tsai, S.Y., Tsai, M.-J., and O'Malley, B.W. (1995). Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. *Science.* **270**, 1354-1357.

Ookata, K., Hisanaga, S., Okano, T., Tchibana, K., and Kishimoto, T. (1992). Relocation and distinct subcellular localization of p34^{cdc2}-cyclin B complex at meiosis reinitiation in starfish oocytes. *EMBO J.* **11**, 1763-1772.

Pabo, C.O. and Sauer, R.T. (1992). Transcription factors: structural families and principles of DNA recognition. *Annu. Rev. Biochem.* **61**, 1053-1095.

Patel, G., Kreider, B., Rovera, G., and Reddy, E.P. (1993). v-*myb* blocks granulocyte colony-stimulating factor-induced myeloid cell differentiation but not proliferation. *Mol. Cell. Biol.* **13**, 2269-2276.

Pazin, M.J. and Kadonaga, J.T. (1997). What's up and down with histone deacetylation and transcription? *Cell.* **89**, 325-328.

Pelicci, P.-G., Lanfrancone, L., Brathwaite, M.D., Wolman, S.R., and Dalla-Favera, R. (1984). Amplification of the c-*myb* oncogene in a case of human acute myelogenous leukemia. *Science.* **224**, 1117-1121.

Peters, C.W.B., Sippel, A.E., Vingron, M., and Klempnauer, K.-H. (1987). *Drosophila* and vertebrate *myb* proteins share two conserved regions, one of which functions as a DNA-binding domain. *EMBO J.* **6**, 3085-3090.

Petrij, F., Giles, R.H., Dauwerse, H.G., Saris, J.J., Hennekam, R.C.M., Masuno, M., Tommerup, N., van Ommen, G.-J.B., Goodman, R.H., Peters, D., J.M., and Breuning, M.H. (1995). Rubenstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. *Nature.* **376**, 348-351.

Phan, S.-C., Feeley, B., Withers, D., and Boxer, L.M. (1996). Identification of an inducible regulator of c-*myb* expression during T-cell activation. *Mol. Cell. Biol.* **16**, 2387-2393.

Plaza, S., Turque, N., Dozier, C., Bailly, M., and Saule, S. (1995). c-Myb acts as a transcriptional activator of the quail PAX6 (PAX-QNR) promoter through two different mechanisms. *Oncogene.* **10**, 329-340.

Postigo, A.A., Sheppard, A.M., Mucenski, M.L., and Dean, D.C. (1997). c-Myb and Ets proteins synergize to overcome transcriptional repression by ZEB. *EMBO J.* **16**, 3924-3934.

Press, R.D., Reddy, E.P., and Ewert, D.L. (1994). Overexpression of C-terminally but not N-terminally truncated Myb induces fibrosarcomas: a novel nonhematopoietic target cell for the *myb* oncogene. *Mol. Cell. Biol.* **14**, 2278-2290.

Press, R.D., Wisner, T.W., and Ewert, D.L. (1995). Induction of B cell lymphomas by overexpression of a Myb oncogene truncated at either terminus. *Oncogene.* **11**, 525-535.

Ptashne, M. and Gann, A. (1997). Transcriptional activation by recruitment. *Nature.* **386**, 569-577.

Radke, K., Beug, H., Kornfeld, S., and Graf, T. (1982). Transformation of both erythroid and myeloid cells by E26, an avian leukemia virus that contains the *myb* gene. *Cell.* **31**, 643-653.

Ramsay, R.G., Ishii, S., Nishina, Y., Soe, G., and Gonda, T.J. (1989). Characterization of alternate and truncated forms of murine c-*myb* proteins. *Oncogene Research.* **4**, 259-269.

Ramsay, R.G., Ishii, S., and Gonda, T.J. (1991). Increase in specific DNA binding by carboxyl truncation suggests a mechanism for activation of Myb. *Oncogene.* **6**, 1875-1879.

Ramsay, R.G., Ishii, S., and Gonda, T.J. (1992a). Interaction of the Myb protein with specific DNA binding sites. *J. Biol. Chem.* **267**, 5656-5662.

Ramsay, R.G., Thompson, M.A., Hayman, J.A., Reid, G., Gonda, T.J., and Whitehead, R.H. (1992b). Myb expression is higher in malignant human colonic carcinoma and premalignant adenomatous polyps than in normal mucosa. *Cell Growth Differ.* **3**, 723-730.

Ramsay, R.G., Morice, N., Van Eeden, P., Kanagasundaram, V., Nomura, T., de Blaquiere, J., Ishii, S., and Wettenhall, R. (1995). Regulation of c-*Myb* through protein phosphorylation and leucine zipper interactions. *Oncogene.* **11**, 2113-2120.

Raschella, G., Negroni, A., Skorski, T., Pucci, S., Nieborowska-Skorska, M., Romeo, A., and Calabretta, B. (1992). Inhibition of proliferation by c-*myb* antisense RNA and oligodeoxynucleotides in transformed neuroectodermal cell lines. *Cancer Res.* **52**, 4221-4226.

Ratajczak, M.Z., Luger, S.M., DeRiel, K., Abrahm, J., Calabretta, B., and Gerwitz, A.M. (1992). Role of the *KIT* protooncogene in normal and malignant human hematopoiesis. *Proc. Natl. Acad. Sci. USA.* **89**, 1710-1714.

Rayner, J.R. and Gonda, T.J. (1994). A simple and efficient procedure for generating stable expression libraries by cDNA cloning in a retroviral vector. *Mol. Cell. Biol.* **14**, 880-887.

Reddy, C.D. and Reddy, E.P. (1989). Differential binding of nuclear factors to the intron 1 sequences containing the transcriptional pause site correlates with c-*myb* expression. *Proc. Natl. Acad. Sci. USA.* **86**, 7326-7330.

Reddy, M.A., Yang, B.-S., Yue, X., Barnett, C.J.K., Ross, I.L., Sweet, M.J., Hume, D.A., and Ostrowski, M.C. (1994). Opposing actions of c-*ets*/PU.1 and c-*myb* protooncogene products in regulating the macrophage-specific promoters of the human and mouse colony-stimulating factor-1 receptor (c-*fms*) genes. *J. Exp. Med.* **180**, 2309-2319.

Richards, S.A., Lounsbury, K.M., Carey, K.L., and Macara, I.G. (1996). A nuclear export signal is essential for the cytosolic localization of the Ran binding protein, RanBP1. *J. Cell Biol.* **134**, 1157-1168.

Roberts, B.L., Richardson, W.D., and Smith, A.E. (1987). The effect of protein context on nuclear location signal function. *Cell.* **50**, 465-475.

Rosenthal, M.A., Thompson, M.A., Ellis, S., Whitehead, R.H., and Ramsay, R.G. (1996). Colonic expression of c-*myb* is initiated in *utero* and continues throughout adult life. *Cell Growth Differ.* **7**, 961-967.

Roussel, M., Saule, S., Lagrou, C., Rommens, H., Beug, H., Graf, T., and Stehelin, D. (1979). Three new types of viral oncogene of cellular origin specific for hematopoietic cell transformation. *Nature.* **281**, 452-455.

Sachs, A.B. and Buratowski, S. (1997). Common themes in translational and transcriptional regulation. *Trends Biochem. Sci.* **22**, 189-192.

Sakura, H., Kanei-Ishii, C., Nagase, T., Nakagoshi, H., Gonda, T.J., and Ishii, S. (1989). Delineation of three functional domains of the transcriptional activator encoded by the c-*myb* protooncogene. *Proc. Natl. Acad. Sci. USA.* **86**, 5758-5762.

Sala, A., Bellon, T., Melotti, P., Peschle, C., and Calabretta, B. (1995). Inhibition of erythro-myeloid differentiation by constitutive expression of a DNA binding-deficient c-*myb* mutant: implication for c-*myb* function. *Blood.* **86**, 3404-3412.

Sala, A., Kundu, M., Casella, I., Engelhard, A., Calabretta, B., Grasso, L., Paggi, M.G., Giordano, A., Watson, R.J., Khalili, K., and Peschle, C. (1997). Activation of human B-MYB by cyclins. *Proc. Natl. Acad. Sci. USA.* **94**, 532-536.

Salomoni, P., Perrotti, D., Martinez, R., Franceschi, C., and Calabretta, B. (1997). Resistance to apoptosis in CTLL-2 cells constitutively expressing c-Myb is asscociated with induction of BCL-2 expression and Myb-dependent regulation of *bcl-2* promoter activity. *Proc. Natl. Acad. Sci. USA.* **94**, 3296-3301.

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

Shapiro, L.H. (1995). Myb and Ets proteins cooperate to transactivate an early myeloid gene. *J. Biol. Chem.* **270**, 8763-8771.

Shaw, P.J. and Jordan, E.G. (1995). The nucleolus. *Annu. Rev. Cell Dev. Biol.* **11**, 93-121.

Sheiness, D. and Gardinier, M. (1984). Expression of a proto-oncogene (proto-*myb*) in hemopoietic tissues of mice. *Mol. Cell. Biol.* **4**, 1206-1212.

Shen-Ong, G.L.C. (1990). The *myb* oncogene. *Biochim. Biophys. Acta.* **1032**, 39-52.

Shikama, N., Lyon, J., and La Thangue, N.B. (1997). The p300/CBP family: integrating signals with transcription factors and chromatin. *Trends Cell Biol.* **7**, 230-236.

Sigler, P.B. (1988). Acid blobs and negative noodles. *Nature.* **333**, 210-212.

Sitzmann, J., Noben-Trauth, K., and Klempnauer, K.-H. (1995). Expression of mouse c-*myb* during embryonic development. *Oncogene.* **11**, 2273-2279.

Siu, G., Wurster, A.L., Lipsick, J.S., and Hedrick, S.M. (1992). Expression of the CD4 gene requires a Myb transcription factor. *Mol. Cell. Biol.* **12**, 1592-1604.

Slamon, D.J., deKernion, J.B., Verma, I.M., and Cline, M.J. (1984). Expression of cellular oncogenes in human malignancies. *Science.* **224**, 256-262.

Slamon, D.J., Boone, T.C., Murdock, D.C., Keith, D.E., Press, M.F., Larson, R.A., and Souza, L.M. (1986). Studies of the human c-*myb* gene and its product in human acute leukemias. *Science.* **233**, 347-351.

Smith, A., Metcalf, D., and Nicola, N.A. (1997). Cytoplasmic domains of the common b-chain of the GM-CSF/IL-3/IL-5 receptors that are required for inducing differentiation or clonal suppression in myeloid leukaemic cell lines. *EMBO J.* **16**, 451-464.

Smith, C.L., Onate, S.A., Tsai, M.-J., and O'Malley, B.W. (1996). CREB binding protein acts synergistically with steroid receptor coactivator-1 to enhance steroid receptor-dependent transcription. *Proc. Natl. Acad. Sci. USA.* **93**, 8884-8888.

Sollner-Webb, B. and Mougey, E.B. (1991). News from the nucleolus: rRNA gene expression. *Trends Biochem. Sci.* **16**, 58-62.

Sternglanz, R. (1996). Histone acetylation: a gateway to transcriptional activation. *Trends Biochem. Sci.* **21**, 357-358.

Sudo, T., Miyazawa, H., Hanaoka, F., and Ishii, S. (1992). The c-*myb* proto-oncogene product binds to but does not activate the promoter of the DNA polymerase α gene. *Oncogene.* **7**, 1999-2006.

Sullivan, J., Feeley, B., Guerra, J., and Boxer, L.M. (1997). Identification of the major positive regulators of c-*myb* expression in haemopoietic cells of different lineages. *J. Biol. Chem.* **272**, 1943-1949.

Tanaka, Y., Nomura, T., and Ishii, S. (1997). Two regions in c-*myb* proto-oncogene product negatively regulating its DNA binding activity. *FEBS Letters.* **413**, 162-168.

Tanikawa, J., Yasukawa, T., Enari, M., Ogata, K., Nishimura, Y., Ishii, S., and Sarai, A. (1993). Recognition of specific DNA sequences by the c-*myb* protooncogene product: role of the three repeat units in the DNA-binding domain. *Proc. Natl. Acad. Sci. USA.* **90**, 9320-9324.

Tashiro, S., Takemoto, Y., Handa, H., and Ishii, S. (1995). Cell type-specific trans-activation by the B-*myb* gene product: requirement of the putative cofactor binding to the C-terminal conserved domain. *Oncogene.* **10**, 1699-1707.

Tavner, F.J., Simpson, R., Tashiro, S., Favier, D., Jenkins, N.A., Gilbert, D.J., Copeland, N.G., Macmillan, E.M., Lutwyche, J., Keough, R.A., Ishii, S., and Gonda, T.J. (1998). Molecular cloning reveals that the p160 Myb-binding protein is a novel, predominantly nucleolar protein which may play a role in transactivation by Myb. *Mol. Cell. Biol.* **18**, In Press.

Taylor, D., Badiani, P., and Weston, K. (1996). A dominant interferring Myb mutant causes apoptosis in T cells. *Genes Dev.* **10**, 2732-2744.

Thiele, C.J., Cohen, P.S., and Israel, M.A. (1988). Regulation of c-*myb* expression in human neuroblastoma cells during retinoic acid-induced differentiation. *Mol. Cell. Biol.* **8**, 1677-1683.

Thompson, C.B., Challoner, P.B., Neiman, P.E., and Groudine, M. (1986). Expression of the c-*myb* proto-oncogene during cellular proliferation. *Nature.* **319**, 374-380.

Thompson, M.A., Flegg, R., Westin, E.H., and Ramsay, R.G. (1997). Microsatellite deletions in the c-*myb* transcriptional attenuator region associated with over-expression in colon tumour cell lines. *Oncogene.* **14**, 1715-1723.

Thompson, M.A. and Ramsay, R.G. (1995). Myb: an old oncoprotein with new roles. *BioEssays.* **17**, 341-350.

Tjian, R. and Maniatis, T. (1994). Transcriptional activation: a complex puzzle with few easy pieces. *Cell.* **77**, 5-8.

Todokoro, K., Watson, R.J., Higo, H., Amanuma, H., Kuramochi, S., Yanagisawa, H., and Ikawa, Y. (1988). Down-regulation of c-*myb* gene expression is a prerequisite for erythropoietin-induced erythroid differentiation. *Proc. Natl. Acad. Sci. USA.* **85**, 8900-8904.

Torchia, J., Rose, D.W., Inostroza, J., Kamei, Y., Westin, S., Glass, C.K., and Rosenfeld, M.G. (1997). The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. *Nature.* **387**, 677-684.

Torelli, G., Venturelli, D., Colo, A., Zanni, C., Selleri, L., Moretti, L., Calabretta, B., and Torelli, U. (1987). Expression of c-*myb* protooncogene and other cell cycle-related genes in normal and neoplastic human colonic mucosa. *Cancer Res.* **47**, 5266-5269.

Toscani, A., Mettus, R.V., Coupland, R., Simpkins, H., Litvin, J., Orth, J., Hatton, K.S., and Reddy, E.P. (1997). Arrest of spermatogenesis and defective breast development in mice lacking A-*myb*. *Nature.* **386**, 713-717.

Toth, C.R., Hostutler, R.F., Baldwin, A.S., and Bender, T.P. (1995). Members of the nuclear factor kB family transactivate the murine c-*myb* gene. *J. Biol. Chem.* **270**, 7661-7671.

Trauth, K., Mutschler, B., Jenkins, N.A., Gilbert, D.J., Copeland, N.G., and Klempnauer, K.-H. (1994). Mouse A-*myb* encodes a trans-activator and is expressed in mitotically active cells of the developing central nervous system, adult testis and B lymphocytes. *EMBO J.* **13**, 5994-6005.

Ullman, K.S., Powers, M.A., and Forbes, D.J. (1997). Nuclear export receptors: from importin to exportin. *Cell.* **90**, 967-970.

Valtieri, M., Venturelli, D., Care, A., Fossati, C., Pelosi, E., Labbaye, C., Mattia, G., Gerwitz, A.M., Calabretta, B., and Peschle, C. (1991). Antisense *myb* inhibition of purified erythroid progenitors in development and differentiation is linked to cycling activity and expression of DNA polymerase α. *Blood.* **77**, 1181-1190.

Venturelli, D., Travali, S., and Calabretta, B. (1990). Inhibition of T-cell proliferation by a Myb antisense oligomer is accompanied by selective down-regulation of DNA polymerase α expression. *Proc. Natl. Acad. Sci. USA.* **87**, 5963-5967.

Vinson, C.R., Hai, T., and Boyd, S.M. (1993). Dimerization specificity of the leucine zipper-containing bZIP motif on DNA binding: prediction and rational design. *Genes Dev.* **7**, 1047-1058.

Voelker, R.A., Huang, S.-M., Wisely, G.B., Sterling, J.F., Bainbridge, S.P., and Hiraizumi, K. (1989). Molecular and genetic organization of the *suppressor of sable* and *Minute (1) 1B* region in *Drosophila melanogaster. Genetics.* **122**, 625-642.

Voit, R., Schnapp, A., Kuhn, A., Rosenbauer, H., Hirschmann, P., Stunnenberg, H.G., and Grummt, I. (1992). The nucleolar transcription factor mUBF is phosphorylated by casein kinase II in the C-terminal hyperacidic tail which is essential for transactivation. *EMBO J.* **11**, 2211-2218.

Voit, R., Kuhn, A., Sander, E.E., and Grummt, I. (1995). Activation of mammalian ribosomal gene transcription requires phosphorylation of the nucleolar transcription factor UBF. *Nucleic Acids Res.* **23**, 2593-2599.

Vorbrueggen, G., Kalkbrenner, F., Guehmann, S., and Moelling, K. (1994). The carboxyterminus of human c-myb protein stimulates activated transcription in *trans. Nucleic Acids Res.* **22**, 2466-2475.

Wallrapp, C., Muller-Pillasch, F., Solinas-Toldo, S., Lichter, P., Friess, H., Buchler, M., Fink, T., Adler, G., and Gress, T.M. (1997). Characterization of a high copy number amplification at 6q24 in pancreatic cancer identifies c-*myb* as a candidate oncogene. *Cancer Res.* **57**, 3135-3139.

Watson, R.J. (1988). Expression of the c-*myb* and c-*myc* genes is regulated independently in differentiating mouse erythroleukemia cells by common processes of premature transcription arrest and increased mRNA turnover. *Mol. Cell. Biol.* **8**, 3938-3942.

Wen, W., Meinkoth, J.L., Tsien, R.Y., and Taylor, S.S. (1995). Identification of a signal for rapid export of proteins from the nucleus. *Cell.* **82**, 463-473.

Westin, E.H., Gallo, R.C., Arya, S.K., Eva, A., Souza, L.M., Baluda, M.A., Aaronson, A., and Wong-Staal, F. (1982). Differential expression of the *amv* gene in human hematopoietic cells. *Proc. Natl. Acad. Sci. USA.* **79**, 2194-2198.

Weston, K. (1992). Extension of the DNA binding consensus of the chicken c-Myb and v-Myb proteins. *Nucleic Acids Res.* **20**, 3043-3049.

Weston, K. and Bishop, J.M. (1989). Transcriptional activation by the v-*myb* oncogene and its cellular progenitor, c-*myb*. *Cell.* **58**, 85-93.

Wolffe, A.P. (1994a). Transcription: in tune with the histones. *Cell.* **77**, 13-16.

Wolffe, A.P. (1994b). Nucleosome positioning and modification: chromatin structures that potentiate transcription. *Trends Biochem. Sci.* **19**, 240-244.

Woodgett, J.R. (1991). A common denominator linking glycogen metabolism, nuclear oncogenes and development. *Trends Biochem. Sci.* **16**, 177-181.

Xue, Z. and Melese, T. (1994). Nucleolar proteins that bind NLSs: a role in nuclear import or ribosome biogenesis. *Trends Cell Biol.* **4**, 414-417.

Yanagisawa, H., Nagasawa, T., Kuramochi, S., Abe, T., Ikawa, Y., and Todokoro, K. (1991). Constitutive expression of exogenous c-*myb* gene causes maturation block in monocyte-macrophage differentiation. *Biochim. Biophys. Acta* **1088**, 380-384.

Yang, T.-H., Tsai, W.-H., Lee, Y.-M., Lei, H.-Y., Lai, M.-Y., Chen, D.-S., Yeh, N.-H., and Lee, S.-C. (1994). Purification and characterization of nucleolin and its identification as a transcription repressor. *Mol. Cell. Biol.* **14**, 6068-6074.

Zawel, L. and Reinberg, D. (1995). Common themes in assembly and function of eukaryotic transcription factors. *Annu. Rev. Biochem.* **64**, 533-561.

Zhou, D.-J., Ahuja, H., and Cline, M.J. (1989). Proto-oncogene abnormalities in human breast cancer: c-ERBB-2 amplification does not correlate with recurrence of disease. *Oncogene.* **4**, 105-108.

Ziebold, U., Bartsch, O., Marais, R., Ferrari, S., and Klempnauer, K.-H. (1997). Phosphorylation and activation of B-Myb by cyclin A-Cdk2. *Current Biology.* **7**, 253-260.