TRANSCRIPTIONAL REGULATION OF THE HUMAN ALPHA 2(I) PROCOLLAGEN GENE

Thesis presented by

VIRNA DRUCILLE LEANER

in fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MEDICAL BIOCHEMISTRY

in the

FACULTY OF MEDICINE
UNIVERSITY OF CAPE TOWN

September 1997

CERTIFICATE OF SUPERVISOR

In terms of paragraph eight of the "General regulations for the degree of Ph.D.", I, as supervisor of the candidate V.D. LEANER, certify that I approve of the incorporation into this thesis of material that has already been published or submitted for publication.

Signed by candidate

Prof. M.I. Parker

Department of Medical Biochemistry

D.G. Murray Trust Chair of Cancer Research

ACKNOWLEDGEMENTS

Thanks to the Almighty

I wish to express my sincere thanks to the following individuals:

- Prof. M. Iqbal Parker, for excellent supervision, guidance, support and fostering a sense of scientific questioning in me.
- Prof. Wieland Gevers, Prof. Mario Ehlers, Prof. Lutz Thilo, Dr Cynthia Sikakana, Dept Med. Biochem, UCT, as role models for young scientists in all disciplines.
- Malcolm Collins, Howard Donninger, Gael Fenhalls and Agatha Masemola for willingness to be of assistance at all times and useful and often entertaining discussion.
- My other colleagues, past and present, E. Dietzsch, J. Davies, E. Joseph, S. Herman, 'M. Ndlovu, Y. Giyose, H. Karjiker, Y.P. Goldberg, A. Smith, D. Bhoolia, M. Geyp and others in the Department of Medical Biochemistry for helpful discussion and for creating a stimulating work environment.
- My family for their support and encouragement throughout the duration of my studies.
- Zulaiga Rossouw, for friendship, motivation and providing me with opportunities to enjoy many an African sunset in the bush.
- Chrismaur Oppelt, Portia and John Rodgers, for friendship, encouragement and sharing in both the good and bad times.
- Graham Julies for his helpfulness, especially in supplying computer facilities.
- My friends in Natsoc and Protec, for proving that together we can make a difference.
- The Medical Research Council, Foundation for Research and Development (FRD), Duitsche Akademiese Ausl Dienste (DAAD) and the University of Cape Town for scholarships and bursaries in support of this work.

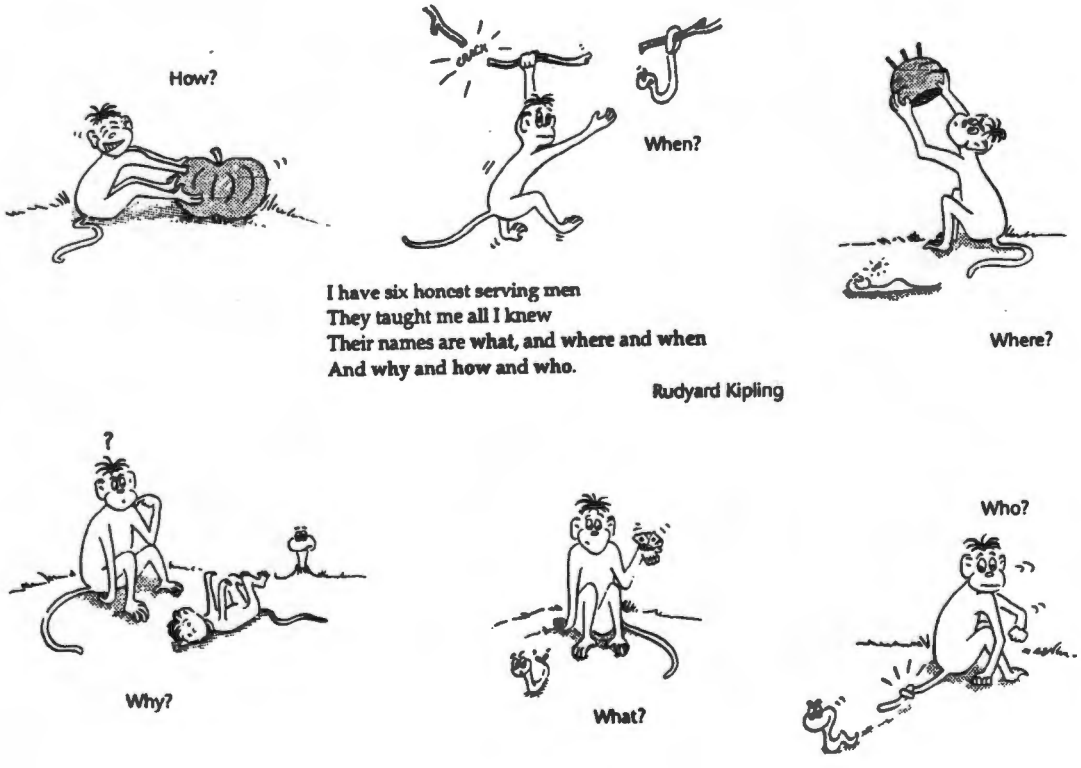
CONTENTS

ABBREVIATIONS

ATCC	American Type Culture Collection
ATP	adenosine triphosphate
BAPN	β-aminopropionitrile
bp	base pair
CAT	chloramphenicol acetyltransferase
cDNA	complementary deoxyribonucleic acid
CIP	calf intestinal phosphatase
CME	collagen modulating element
cpm	counts per minute
CTP	cytidine triphosphate
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
DMEM	Dubelco's minimum essential medium
DNAse1	deoxyribonuclease 1
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
EDTA	ethylediaminetetra-acetic acid
EMSA	electrophoretic mobility shift assay
G/CBE	GGAGG/CCAAT binding element
G/CBF	GGAGG/CCAAT binding factor
GITC	guanidine (aminomethanamide) thiocyanate salt
GTP	guanosine triphosphate
HEPES	N-2-hydroxyethylpiperazine-N'-ehanesulphoic acid
IPTG	isopropyl-1-thio-β-D-galactosidase
Kb	kilobases
kD	kilodalton
ONPG	O-nitrophenyl-β-D-galactopyranosidase
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PIPES	piperazine-N, N'-bis-(2-ethanesulphonic acid)
PMSF	phenylmethylsulphonylfluoride
RNAse	ribonuclease
SV40	simian virus-40
TCA	tricarboxylic acid
Tris	tris(hydroxymethyl) aminomethane
UTP	uridine triphosphate
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactosidase

TO MY PARENTS AND GRANDPARENTS

the honest men and women who supported me in seizing opportunities not available to them in the past



From: The Book of Projects, Derek John Gray, 1989.

ABSTRACT

The objective of this study was to investigate the cell- and species-specific regulation of the α2(1) procollagen gene by analysing trans-acting factor interactions within the proximal promoter of the gene and to identify the genes coding for these trans-acting factors.

α2(1) procollagen gene expression was examined in a number of differentiated cell lines and shown to differ significantly between normal fibroblasts (WI-38, FG_o), transformed fibroblasts (CT-1, SVWI-38), HT1080 fibrosarcoma, HepG2 hepatocellular carcinoma, L77 lymphoblasts and breast cancer epithelial cells (MDA-MB-231, ZR-75-2). These differences were due to changes in transcription of the α2(1) procollagen gene as shown by Northern blot analysis and nuclear run-on transcription experiments. Analysis of DNA-protein interactions with the proximal α2(1) procollagen promoter showed the presence of at least two DNA-protein complexes (complexes I and III) in collagen producing cell lines, while cells where collagen synthesis did not occur contained a third DNA-protein complex (complex II). α2(1) procollagen gene expression was therefore shown to be associated with the presence of complexes I and III while repression of the gene was associated with the presence of complexes I and II and the partial or complete absence of complex III. Complex I is a ubiquitous factor which binds the inverted CCAAT box located between -92 and -80 (G/CBE) with an apparent K_d of 2.9nM. Complexes II and III both bind an adjacent DNA sequence between -78 and -67 (the CME) with K_d values of 4.2 and 3.5nM respectively.

While the CCAAT boxes in the human and mouse promoters are identical, a 3 bp mismatch was detected in the CME. This mismatch abolished the formation of complex II and III on the mouse promoter, even though mouse cells contained

complex II proteins. The difference in the CME binding site between rodent and human promoters implied species-specific regulation of the α2(1) procollagen gene. Transfection of human and mouse proximal α2(1) procollagen promoter/CAT constructs into human cells (CT-1) indicated that the human promoter had higher activity than the mouse promoter, whilst the two promoters had equivalent activities in rodent cells. These promoter activities may be accounted for by the differences in trans-acting factor binding to the two promoters.

Complex I formation was competed out by the mouse CBF and NF-Y consensus oligonucleotides, while the mouse anti-CBF-B antibody resulted in a supershifted complex I. These results indicate that complex I is a member of the heterologous CCAAT-binding proteins and possibly related to or similar to the mouse CBF.

The treatment of nuclear extracts with calf intestinal phosphatase resulted in a loss of complex I formation on the human and CBF binding to the mouse promoters. The Ser/Thr phosphatase, PP2A, specifically inhibited complexes II and III formation. Nuclear extracts from CT-1 and U937 cell lines treated with the kinase inhibitor, staurosporin, was accompanied by a loss in DNA-protein interaction. This inhibition of DNA-binding activity was not observed using the tyrosine kinase inhibitor, genistein, and the PP2A phosphatase inhibitor, okadaic acid. Staurosporin also had a significant inhibitory effect on α2(1) procollagen promoter activity in CT-1 cells transfected with the human proximal α2(1) procollagen promoter and on steady state collagen mRNA levels. These results indicate that phosphorylation is required for the binding of trans-acting factors to the proximal α2(1) procollagen promoter and in transcriptional regulation of this gene. In support of the suggestion that phosphorylation events play a role in transcriptional regulation of the α2(1) procollagen gene, CT-1 cells treated with the protein

kinase C activator, PMA, showed a significant reduction in α2(1) procollagen mRNA levels.

A lambda gt11 expression library was screened to obtain cDNA's encoding proteins that bind the CME in the human α2(1) proximal promoter. A cDNA clone of 958 bp with a predicted open reading frame of 116 amino acids (12.5kD) was obtained. No significant DNA or polypeptide sequence homologies existed in the databank, indicating the possibility of a novel trans-acting factor. Binding of this fusion protein was specific for the CME as observed in South Western blotting and gel shift assays using competitor DNA sequences. Northern blot analysis detected a mRNA transcript of approximately 4kb predominantly in cells where α2(1) procollagen expression is repressed.

1. INTRODUCTION

CHAPTER 1:

INTRODUCTION

1.1 THE COLLAGENS

The collagens are a superfamily of extracellular structural proteins with diverse biological functions, ranging from promotion of cell growth and differentiation, cell migration and tissue development. There are many excellent reviews on the structure /function and pathobiology of the different collagens (Bornstein and Sage, 1980; Burgeson, 1988; Ramirez et al, 1989; Ramirez et al, 1990; Vuorio and de Crombrugghe, 1990; van der Rest and Garrone, 1991; Prockop et al, 1993); hence this chapter will present a brief overview on the fibrillar collagens, with specific reference to transcriptional regulatory mechanisms in type I collagen gene expression.

1.1.1 Introduction

At least 19 different collagen types have been characterised to date and their chromosomal location, tissue specificity and biological function vary considerably (Prockop and Kivirikko, 1995). The collagens are triple helical proteins consisting of three polypeptides forming molecular rod-like structures. The primary structure, is for the most part, a repeating Gly-X-Y amino acid triplet with positions X and Y frequently being occupied by prolyl and hydroxyprolyl residues respectively. This unique molecular structure and arrangement confers on the collagens the ability to polymerise and form highly ordered structures. Many of the collagens contain globular, non-collagenous domains which vary in size from 10 to 150 kD per polypeptide chain. These globular sequences are speculated to make the molecule more flexible although their exact function remains to be identified. Based on the

polymeric structures that the collagens form, they can be subdivided into different classes; (1) The fibrillar collagens, (2) collagens that form network-like structures, (3) the fibril-associated collagens (FACITS), (4) collagens forming beaded filaments, (5) anchoring collagen fibrils, (6) collagens containing transmembrane domains, (7) newly identified collagens and (8) non-collagen collagens (Prockop and Kivirikko, 1995).

1.1.2 The fibrillar collagens - a brief overview

This group of collagens are identified by their formation of highly organised fibers and fibrils which provide structural support in most connective tissues, blood vessels, bones and nerves. The structure of the genes encoding these collagens have common features in that the triple-helical domain is encoded by 42 exons varying in size usually in multiples of 54 bp (1 X 54 bp, 2 X 54 bp, 3 X 54 bp). Exons of 45 bp and 99 bp are also found. The introns seperating the exons vary in size from 100 to 3000 bp (reviewed by Ramirez et al, 1990; Vuorio and de Crombrugghe, 1990 and Prockop and Kivirikko, 1995). Types I, II and III (major fibrillar collagens) are the most abundant collagens in this group, with type V and XI constituting the minor fibrillar collagens. These collagens all consist of long triple-helical domains of approximately 1000 amino acids, with short globular domains flanking the triple-helical portion.

After post-translational modifications and assembly of the procollagen helix, specific N- and C-proteases remove the globular propeptides. Interestingly, such proteases have been demonstrated to have a potential role in the regulation of collagen biosynthesis. The cleavage of the C-propeptide by mast cell chymase is reported to initiate fibril formation (Kofford et al, 1997). A recent study also shows that type III collagen, which is usually associated with type I collagen; is essential for type I collagen fibrillogenesis in the cardiovascular system and other organs (Liu et al, 1997). Similiarly, it has been shown that the interaction between types I and V collagen regulates fibril formation (Birk et al, 1990; Marchant et al, 1996).

Type I collagen synthesis varies significantly under different physiologic conditions, in different cellular tissues and in certain pathologic conditions. An understanding of the mechanisms underlying the expression of the type I collagen genes will contribute to our understanding of these conditions.

1.2 REGULATION OF TYPE I COLLAGEN GENE EXPRESSION

1.2.1 Preamble

One of the earliest studies regarding the regulation of collagen biosynthesis in differentiated cell types was provided by Green et al, (1966). These authors classified differentiated cell types in terms of collagen expression in three levels at which synthesis may be maintained. These include cells of fibroblastic origin where the ability to synthesise collagen is highly developed, non-fibroblastic cell types which synthesize collagen at a rate of 2-3% compared to fibroblasts and cell types where the collagen genes are completely silenced. Interestingly, this paper concluded by stating that it was not possible to say whether any of these collagen states are produced by soluble regulatory factors. After more than thirty years of intensive investigation, a number of protein factors have been identified that have regulatory function in the expression of the collagens identified to date, and more are being discovered. The cell- and tissue-specific transcription of the type I collagen genes is mainly attributed to complex interactions between *cis*-regulatory DNA elements and sequence specific trans-acting factors. These regulatory elements are located predominantly in the promoter region, 5' flanking region, the first intron and the 3' untranslated region of the genes.

The following section will provide an overview of the factors influencing cell- and tissue-specific type I collagen synthesis as well as highlight some of the currently debated issues in the field.

1.2.2. Co-oordinate expression of the type I collagen genes

It has generally been accepted that the expression of the α1(1) and α2(1) procollagen genes are co-ordinately regulated (de Wet et al, 1983). The steady state levels of α1(1) and α2(1) collagen mRNAs are in the ratio of 2:1(Vuust et al, 1985), and are reflected in the synthesis of pro α1(1) and pro α2(1) chains (Bornstein and Sage, 1989). This ratio is, however, not maintained under varying culture conditions. While nuclei from sub-confluent cultures generally transcribe the two genes in a 2:1 ratio, nuclei from post-confluent cultures transcribe the genes at a ratio of approximately 4:1 (Olsen and Prockop, 1989). Steady state mRNA levels remain at a 2:1 ratio and it is likely that in addition to transcriptional regulation, post-transcriptional mechanisms also have a role in maintaining these levels. Co-ordinate expression of type I collagen genes has also been shown to occur during mouse development. Concommitant with this expression, co-ordinate α1(111) collagen gene expression is also detected in fibroblast containing tissues (Niederreither et al, 1995). Transfections into human skin fibroblasts using 500 bp of the proximal α1(1) and α2(1) procollagen promoters fused to the luciferase gene provided additional evidence for the co-ordinate expression of these genes (Hata, 1995).

The mechanism by which the co-ordinate regulation of the α1(1) and α2(1) collagen genes can be achieved is via common transcriptional factors that control both genes. A number of *cis*-elements and the transcription factors that interact with these regions have been identified in the promoters of the α1(1) and α2(1) procollagen genes and are discussed below.

1.2.3 Type I collagen promoter elements and trans-acting factors

1.2.3.1 Transcription factors in co-ordinate regulation

Both the α1(1) and α 2(1) procollagen promoters contain the characteristic TATA box which is the binding site for the general transcription factors, a common CCAAT motif

and binding sites for other transcription factors. These factors have a key role in the co-ordinate regulation of the type I collagen genes.

1.2.3.1.1 The CCAAT binding factor (CBF)

The CCAAT consensus sequence is found in most eukaryotic promoters and is involved in the tissue and cell-specific expression of a number of genes including the α-globin (Cohen et al, 1986; Barnhait et al, 1988), the liver specific aldolase B (Gregori et al, 1993), the rat albumin (Tronche et al, 1991; Wuarin et al, 1990), cell cycle-related genes eg thymidine kinase (Flemington et al, 1987; Pang and Chen, 1993), the major histocompatability complex class II genes (Dorn et al, 1987) and others (Gallinari et al, 1989). Recently the CCAAT binding factor, NF-Y has been reported to be modulated during monocyte-macrophage differentiation and it is required for the transcriptional regulation of the ferritin heavy-chain gene (Marzicialli et al, 1997). CCAAT binding proteins have also been implicated in the regulation of genes involved in stress responses, hsp70 (Wu et al, 1981; Lum et al, 1990) and in regulation of the grp78/BiP promoter under conditions of calcium stress (Roy and Lee, 1996).

Both the α2(1) and α1(1) collagen promoters contain a common inverted CCAAT motif at positions -75 to -98 and -90 to -115 respectively. These inverted CCAAT motifs bind the transcriptional activator, the heterotrimeric CCAAT binding factor identified in NIH-3T3 cells (Hatamochi et al, 1986; Oikarinen et al, 1987; Maity et al, 1988). Other members in the family of heterotrimeric CCAAT-binding factors include NF-Y (Dorn et al, 1987), CP1 (Chodosh et al, 1988a), yeast Hap-2,3,4 (Forsbery and Guarente, 1989) and α CP1 (Kim & Sheffery, 1990). It has generally become accepted that CBF, NF-Y and CP1 are the same or closely related factors. CBF was initially identified as a heterodimeric factor (Hatamochi et al, 1988) but subsequent investigations proved it to be a heterotrimeric protein (Maity et al, 1992; Vuorio et al, 1990).

The three subunits comprising CBF are CBF A, CBF B and CBF C, all of which are required for DNA binding activity. (Maity et al, 1992) and occur in a stoichiometric ratio of 1:1:1 (Sinha et al, 1996). CBF A and CBF B share a striking amino acid homology (99%) in humans, rats and mice. These identities were determined on the basis of their derived cDNA sequences (Li et al, 1992; Maity et al, 1990; Hooft van Huisjduijnen et al, 1990). While CBF A is a heat sensitive protein with a putative polypeptide size of 25 kD, CBF B is heat resistant and has a calculated molecular mass of 41 kD (Hatamochi et al, 1988, Vuorio et al, 1990). CBF A and CBF B have a high sequence identity with the yeast Hap-3 and Hap-2 proteins respectively. The C-terminal domain of CBF B, which shares similarity with the yeast Hap-2 protein, consists of two separate segments, one domain which is required for DNA binding activity while the other for interaction with the other components of CBF (Maity and de Crombrugghe B, 1992). A similar interaction has also been observed for the yeast Hap2 protein with the other components in the yeast CCAAT binding factor, Hap2-3-4 (Oleson et al, 1990; Xing et al, 1994).

CBF C, the third component of the heterotrimeric CBF is required together with CBF A and CBF B for formation of the CBF-DNA complex (Sinha et al, 1995). These authors also showed that CBF A and CBF C initially interact with each other to form a CBF A-CBF C complex. This is followed by the interaction with CBF B and DNA binding of the complete complex is achieved. CBF C shares significant homology with the yeast Hap5 protein (McNabb et al, 1995) and its gene is located on chromosome 4 in mouse with linkage data suggesting that the human NF-Y C (CBF C) will map to 1p32 (Sinha et al, 1996). Using northern blot analysis, CBF C was shown to be ubiquitously expressed in several rat tissues and has a major mRNA species of 2kb (Sinha et al, 1996).

CBF A contains three functional domains which are essential for, (i) interaction with CBF C to form the heterodimeric CBF A/CBF C complex. (ii) interaction of CBF A/CBF C complex with CBF B to form the CBF heterotrimer and (iii) binding of CBF

to DNA (Sihna et al, 1996). Interestingly, the segment of CBF A required for DNA binding of the CBF complex shows sequence similarity with a region of CBF C and these sequences in turn have similarity to the histone fold motifs of histones H2B and H2A (Baxevanis et al, 1995). These motifs are essential for the interaction between CBF A and CBF C and the subsequent binding of the CBF complex to DNA (Kim et al, 1996).

In vivo transcriptional activation by CBF has been reported to require the glutamine rich domain of CBF B, but in an *in vitro* transcription reconstitution system, however, a truncated form of CBF B (lacking the glutamine and ser/thr rich domains) with purified CBF A and CBF C could activate transcription. This suggested that transcriptional activation of CBF may also be achieved via either or both CBF A and CBF C or via other transcription factors interacting with CBF (Coustry et al, 1995). Indeed, a subsequent report showed that CBF transcriptional activity can be mediated by both CBF B and CBF C. Using a CBF deletion protein consisting of the full length CBF A and deleted forms of either CBF B or CBF C, it was shown that these mutant CBF proteins could stimulate transcription but at reduced levels compared to that of the wild type protein (Coustry et al, 1996).

For CBF to participate in cell-specific gene regulation it is thought that CBF functions in conjunction with other transcription factors. Evidence for this has been provided by studies on the CCAAT-binding NF-Y factor which is implicated in the *in vivo* recruitment of upstream DNA binding transcription factors in class II MHC transcription. Constitutive expression of γ-IFN-induced MHC class II DRA expression requires a functional NF-Y CCAAT element (the Y box) and adjacent elements, the S and X boxes (Vilen et al, 1990, 1992). Mutations in the Y box abolished binding of trans-acting factors at the X1 and X2 sites in the HLA-DRA promoters (Wright et al, 1994). These findings suggest that NF-Y stabilises and interacts with X box factors and these interactions are required for gene activation. Sinha et al (1996) speculate that the CBF A/CBF C heterodimer could interact with the TATA binding protein-

associated factors (TAFs) or other TBP-associated proteins to either activate or derepress transcription. Although the above has not been shown as yet, CCAAT binding factor interactions with a number of other proteins have been reported. The B subunit of NF-Y can physically interact with the human T-cell lymphotropic virus type I Tax protein to activate the NF-Y responsive DQB promoter. Tax, an activator of the human T-cell lymphotropic virus type I (HTLV-1) transcription and of many cellular genes, does not bind DNA directly, but regulates transcription via protein-protein interactions (Pise-Masison, 1997). The SV40 large T Antigen (SV40-LT) also transactivates the human cdc 2 promoter via induction of a CCAAT Box binding factor. Although physical interaction between the SV40-LT and CBF has not been shown, it is suggested that CCAAT box binding motifs in the human cdc 2 promoter are the major target for transaction by SV40-LT (Chen et al, 1996). A recent study showed that basal expression of the human α2(1) procollagen promoter may require the binding of a trans-acting factor to sequences adjacent to the CCAAT box which may, in combination with CBF, result in effective activation of the promoter. Mutations in the binding sequence of this factor drastically reduced promoter activity as observed in transient transfection assays (Collins et al, 1997).

The CCAAT binding protein interactions and their role in transcriptional activation are therefore highly complex and involve the interaction of the different components of CBF with each other and the interaction of CBF with other trans-acting factors which may either activate or repress transcription. In terms of the type I collagen genes, sequences flanking the CCAAT motifs have been identified to interact with transcription factors possibly required in the co-ordinate regulation of the genes.

1.2.3.1.2 Transcription factors IF-1 and IF-2

Two DNA binding proteins, Inhibitory factor 1 (IF-1) and Inhibitory factor 2 (IF-2) interact with regions in the proximal α1(1) collagen promoter. Both are proposed transcriptional repressors (Karsenty and de Crombrugghe, 1990). IF-1 binds two adjacent sites (-190 to -170 and -160 to -130) while IF-2 binds a 12bp repeat (5'-

TGGGGGCCGGGC-3') that brackets the proximal CCAAT motif. This 12bp sequence is a GC rich repeat and shows homology to Sp1 binding sites. Sp1, however, appears to bind this region poorly (Brenner et al, 1989). Contrary to these observations, Nehls et al, (1991, 1992) demonstrated that Sp1 binds the 12bp repeat and that overexpression of Sp1 in NIH3T3 cells interferes with basal α1(1) collagen promoter activity. IF-2 is a metallo-protein and mutually exclusive binding between IF-2 and CBF could have a significant role in the regulation of the α1(1) collagen promoter (Karsenty and de Crombrugghe, 1990). The purification and functional characterisation of IF-2 indicated that CBF effectively inhibits IF-2 binding while IF-2 weakly inhibits CBF binding (Karsenty et al, 1991).

The putative transcriptional repressor, IF-1, binds at least two upstream adjacent sites in the mouse α1(1) collagen promoter, located at positions -190 to -170 (element A) and -160 to -130 (element B). EMSAs using a double-stranded oligonucleotide containing the B element indicated that less pronounced binding is detected to the B element than to the A element (Karsenty and de Crombrugghe, 1990). A similar B element is also present in the mouse α2(1) collagen promoter between -173 and -143 and shows a high degree of similarity (10 out of 11 bp) with the B element of the α1(1) promoter. Mutations in this sequence abolish IF-1 binding and result in a 4 fold increase in α2(1) collagen promoter activity as compared to that of the wild type promoter construct (Karsenty and de Crombrugghe, 1991). These studies therefore propose that the IF-1 binding element may in part be involved in the co-ordinate regulation of the α1(1) and α2(1) collagen genes. A homologous IF-1 binding sequence is also present in the human α2(1) collagen promoter at -164 to -159. Mutations in this region result in a 5-6 fold increase in promoter activity, suggesting a possible repressor binding site (Ihn et al, 1996).

In contrast to studies showing that mutations in the IF-1 binding site in the mouse α2(1) collagen promoter (-173 to -143) results in a 4 fold increase in promoter activity

(Karsenty and de Crombrugghe, 1991), a recent study by Hasegawa et al, (1996) showed that deletion of the -170 to -130 region of the mouse α2(1) promoter resulted in a significant reduction in promoter activity. Although this contradiction remains to be explained, the authors provide two possible explainations for how the -173 to -143 region can function as a negative element; (i) that the protein binding this sequence is a true transcriptional repressor or (ii) that the protein binding this sequence is a weak activator and mutations in its binding site may allow the binding of a strong activator, hence the increase in promoter activity observed when the sequence is mutated. These explanations are certainly plausible and it is possible that the relative abundance of transcriptional activators and repressors may be important determinants in the regulation of type I collagen genes.

1.2.3.1.3 Transcription factor Sp1

In addition to the CCAAT elements in the α1(1) and α2(1) collagen promoters, a number of potential Sp1 binding sites have been identified (Rossouw et al, 1987). The same 12bp sequence reported to bind IF-2 (section 1.2.3.1.2) was reported by Nehls et al, (1991, 1992) to bind Sp1, which is thought to bind the α1(1) collagen promoter in a mutually exclusive manner with NF-1. This finding is similar to that reported for IF-2 and CBF (Karsenty et al, 1991), suggesting that IF-2 and Sp1 may be similar factors. Karsenty and de Combrugghe (1991), however, reported that IF-2 and Sp1 are distinct DNA binding factors and it may therefore be possible that sequences flanking the CCAAT sequence interact with Sp1 and other GC-box binding proteins. The isolation of cDNA clones will allow for similarities between IF-2 and Sp1 to be evaluated. The IF-2 binding sequence has also been identified in the α2(1) promoter, suggesting that Sp1 and IF-2 and other GC-rich binding proteins have the ability to bind both the mouse proximal α1(1) and α2(1) collagen promoters. A review by Karsenty and Park, (1995) speculate that Sp1 and other factors binding the sequences flanking the CCAAT motif in the mouse α1(1) procollagen promoter act as activators of

transcription and somehow prevent the binding of a stronger transcriptional activator, the CCAAT binding factor.

Hasegawa et al, (1996) reported that the -170 to -40 region of the mouse α2(1) procollagen promoter contains binding sites (-176 to -152; -136 to -114 and -105 to -65) for a number of nuclear proteins which include Sp1, other proteins that bind to Sp1 consensus sites and proteins that bind to a Krox consensus site. Sp1, however, appears to have a modest role in the activity of the mouse α2(1) collagen promoter, since anti-Sp1 antibodies had little effect on *in vitro* transcription of the WT -350 bp mouse promoter. *In vitro* transcription assays with deletion mutants encomposing the -170 to -40 bp region upstream of the start of transcription, indicated that this region is however essential for transcriptional activity (Hasegawa et al, 1996).

While Sp1 appears to have a modest role in activity of the mouse promoter, it is essential for binding and basal activity of the human α2(1) procollagen promoter. Three short cis-acting GC rich elements between -303 and -274 in the human α2(1) collagen promoter are capable of binding Sp1 and mutations in these boxes abolished trans-acting factor interactions accompanied by a 90% loss in promoter activity in transient transfection assays (Tamaki et al, 1995). Sp1 has also been implicated in the TGF-β response of the human α2(1) collagen promoter (Inagaki et al, 1994; Greenwel et al, 1997) and the debate around this observation will be discussed in section 1.2.3.3.

In a recent report it was shown that purified Sp1 can specifically interact with the human α1(1) procollagen promoter (Poppleton and Raghow, 1997). Sp1 binding sites are also located in the more distal region of the α1(1) collagen promoter (-466 to -459) and in the first intron of the gene (Bornstein et al, 1987; Liska et al, 1992). It is possible that these Sp1 motifs can interact with each other, as it has been shown that Sp1 binding to a promoter can synergistically stimulate transcription via Sp1/Sp1 protein interactions (Courey et al, 1989).

1.2.3.1.4 Recently identified trans-acting factors

Previous reports showed that the sequences between -180 and -136 of the mouse α2(1) collagen promoter are essential for promoter activity and that this region binds a complex set of trans-acting factors. This region of the promoter was used by Hasegawa et al, (1997) to clone transcription factors that bind to specific sequences in the proximal promoter using the yeast one hybrid system. One such cDNA clone, designated BFCOL1, encodes a polypeptide of 775 amino acids of which the N - terminal 400 amino acids share 95% similarity with the human htβ protein. BFCOL1 has four potential zinc fingers and binds the -180 to -152 region of the mouse α 2(1) collagen promoter. This binding site corresponds to that of the IF-1 binding site in the mouse α1(1) and α2(1) collagen promoter (section 1.2.3.1.2). The two IF-1 binding sites in the mouse α1(1) collagen promoter also interact with BFCOL1. However, BFCOL1 had a greater affinity for the -168 to -129 α1(1) oligonucleotide (element B) than to the -194 to -168 α1(1) oligonucleotide (element A). This is in contrast to IF-1 in which the reverse is found. The C- terminal domain of BFCOL1 functions as a potential transcription activator when tested as a GAL4 fusion polypeptide in transfection experiments (Hasegawa et al, 1997).

In addition to BFCOL1, c-Krox also binds the same sequences in the α1(1) and α2(1) collagen promoters. c-Krox appears to bind the -194 to -168 region more effectively than the -168 to -129 region of the α1(1) promoter. This binding correlates with transcriptional activity of a short α1(1) collagen promoter in cotransfection experiments and expression of c-Krox mRNA is tissue-specific with the most abundant expression being in the skin (Galera et al, 1994; reviewed by Karsenty and Park, 1995). Three c-Krox binding sites are identified in the mouse α2(1) collagen promoter, one of which is adjacent to the CCAAT motif in both promoters (Galera et al, 1996). c-Krox binding is the same as that reported for IF-1 (Karsenty and de Crombrugghe, 1990). It is likely that IF-1 and c-Krox are similar, if not the same factor.

The α1(1) and α2(1) procollagen proximal promoter regions therefore interact with a variety of common transcription factors, suggesting a role for these factors in the co-ordinate regulation of the genes. The relationship between IF-1 and c-Krox, IF-2 and Sp1 is not clear as yet, although the preference in recent literature is towards c-Krox and Sp1 as major transcription factors in regulation of the type I collagen genes. It has become increasingly apparent that a number of different transcription factors can interact with the same DNA sequence and the cell type and relative abundance of these proteins may determine which will interact.

1.2.3.2 Additional DNA elements and trans-acting factors involved in constitutive activity of the type I collagen genes

A number of other transcription factors and *cis*-elements in the α1(1) and α2(1) collagen promoters have been identified as being important in the regulation of the genes. An NF-1 binding site (-315 to -295) is located in the mouse α2(1) collagen promoter (Oikarinen et al, 1987). Point mutations which inhibit NF-1 binding exhibited a strong inhibitory effect on α2(1) promoter activity, suggesting that NF-1 (or a protein with similar DNA binding properties) is an activator of α2(1) collagen gene expression. Support for this suggestion was provided in studies where the inclusion of this sequence in the SV40 early promoter resulted in a marked (10 fold) increase in promoter activity (Karsenty et al, 1988). Subsequent studies showed that this region of the promoter is required in TGF-β stimulation of the mouse α2(1) procollagen gene (see section 1.2.3.3). A factor, possibly histone H1, has been purified and was also shown to bind the NF-1-like sequence in the mouse α2(1) procollagen promoter (Ristiniemi and Oikarinen, 1989). The sequences further upstream in the mouse α2(1) collagen promoter -(419 to -399) bind a factor designated as ColF1. Purification of this factor revealed two molecular weight species of 40 and 42 kD. ColF1 binding appears to be distinct from that of CBF and it appears to be more abundant in NIH3T3 fibroblasts than in lymphoid and epithelial cells (Hatamochi et al, 1993). The exact function of this trans-acting factor is unknown as yet.

The sequences -804 t0 -675 in the mouse α1(1) procollagen promoter contain an AP1 binding site and the reduced interaction of AP1 to this sequence results in the increased expression of the α1(1) procollagen gene in tight skin (TSK) mice myocardial fibroblast (Phillips et al, 1995). It would appear that AP1 has a negative regulatory effect within this region of the gene. An AP1 binding site has also been located within the -265 to -241 region of the human α2(1) procollagen gene and there is currently debate as to its involvement in conveying the response to TGF-β (see section 1.2.3.3).

Although cis-elements and trans-acting factors which regulate the rodent type I collagen genes are reasonably well characterised, it is only of late that these elements in the human promoters have been identified. The region up to -376 of the human α2(1) procollagen promoter is sufficient for high level expression of the gene (Boast et al, 1990). Closer analysis of this promoter segment revealed a TGF-β response element (-323 to -186) (Inagaki et al, 1994) and three additional regions involved in nuclear factor binding and transcriptional activity located at -173 to -155; -133 to -119 and -101 to -72 (Ihn et al, 1996). The -101 to -72 region is the previously identified CCAAT motif, -133 to -119 contains a TCCTCC motif which is required for activation of the promoter and -173 to -155 appears to be a repressor binding site. These *cis*-elements and the transcription factors that interact at these regions are fundamental in the regulation of the type I collagen genes.

1.2.3.3 Response of the type I collagen genes to TGF-β

Type I collagen synthesis can be regulated by a number of extracellular signalling molecules. One of these is TGF-β, a well documented regulator of cell cycle progression, cell differentiation and cell adhesion affecting the formation of extracellular matrices (Roberts et al, 1990; Massague,1990). Three ubiquitous cell membrane receptors mediate the effect of TGF-β. Two of these (types I and II) contain novel serine/threonine kinase activities and are responsible for the cellular signalling of

TGF-β (reviewed by Brenner et al, 1994 and Massague and Weis-Garcia, 1996). TGF-β has been shown to have a stimulatory effect on collagen production (Fine and Goldstein, 1989; Ignotz and Massague, 1986). Subcutaneous injections of TGF-β into transgenic mice containing 3.6 kb of the rat α1(1) collagen promoter fused to a reporter CAT gene also resulted in increased CAT expression in the dermis of TGF-β treated animals (Agarwal et al, 1996).

The mechanism underlying stimulation by TGF-β has been extensively investigated. The steady state levels of α1(1) and α2(1) collagen mRNA in 3T3 and Ito cells treated with TGF-β were increased by 15-20 fold (Pentinnen et al, 1988; Weiner et al, 1990). Inhibition of RNA synthesis with actinomycin D showed that the upregulation of collagen gene expression in TGF-β treated cells is due to increased transcriptional activity (Ignotz et al, 1986).

A variety of transcription factors are thought to mediate the TGF-β effect on both the α1(1) and α2(1) collagen genes in rodents and humans. A TGF-β activating element which contains an NF-1-like sequence is located 1.6 kb upstream of the transcriptional start site in the rat α1(1) procollagen promoter (Ritzenhaler et al, 1991). Although this element resembles NF-1 and AP2 binding sites, a 82 kD protein different to both interacts with this region of the α1(1) collagen gene (Ritzenhaler et al, 1993).

TGF-β stimulates the mouse α2(1) collagen gene 6-10 fold. The TGF-β responsive element in this gene is located at position -300 and the transcription factor thought to mediate this response is CTF/NF-1. A 3bp mutation in the NF-1 binding site, which abolishes NF-1 binding, prevents stimulation of mouse α2(1) procollagen expression by TGF-β (Rossi et al, 1988). NF-1 binding to the mouse α2(1) collagen promoter is, however, detected in both TGF-β treated and untreated cells and it is therefore proposed that TGF-β may not neccessarily affect DNA binding, but rather the transcriptional activation function of NF-1 (de Crombrugghe et al, 1990). The exact

role of CTF/NF-1 in the responsiveness of the mouse α2(1) gene to TGF-β could not be confirmed by Chang and Goldberg, (1995). These authors implicate AP1 since overexpression of jun-B induced transcription while jun-B antisense RNA interfered with TGF-β stimulation. This controversy has not been solved as yet, however, evidence for CTF as the factor involved in relaying the signal from TGF-β and tumor necrosis factor-α (TNF-α) on mouse α2(1) collagen gene expression is provided by Alevizopoulous et al, (1996).

Similar controversy as to the transcription factors responsible for the TGF-β responsiveness of the human α2(1) collagen exists. The TGF-β-response element in the human α2(1) collagen gene was initially located between -376 and -108 upstream from the transcriptional start site (Kikuchi et al, 1992). Although an NF-1-like sequence is found at -311 to -298 it has been accepted that NF-1 appears not to be the transcription factor responsible for TGF-β stimulus of the human α2(1) collagen gene. The TGF-β-responsive element (TbRE) has subsequently been mapped to a 131 bp region (-323 to -186) of the COL1A2 promoter. This promoter region contains at least two cis-acting elements, one which has Sp1 recognition sequences reported to be required for promoter inducibility. A nuclear protein complex, of which Sp1 is a component, showed increased binding to the TbRE upon TGF-β stimulation using electrophoretic mobility shift assays (EMSA) and tyrosine dephosphorylation of this complex may be required for transactivation (Inagaki et al, 1994; Greenwel et al, 1995). This region of the promoter also regulates α2(1) collagen gene expression in fat storing cells and dermal fibroblast and this activity can be stimulated by TGF-β (Inagaki et al, 1995a). Three GC-boxes resembling Sp1 binding sites are located in the 131 bp region (-323 to -186). These GC-boxes are located between -303 and -271 and are essential for α2(1) procollagen promoter activity (Tamaki et al, 1996). Contrary to the finding that an Sp1-containing complex may be responsible for the TGF-β stimulus of the human α2(1) procollagen gene, Chung et al, (1996) report that an AP-1 binding sequence overlapping the 3' end of the TbRE is essential for the TGF-β

stimulus. These authors show that deletion of the Sp1 containing sequences does not interfere with TGF-β responsiveness, while deletion of sequences -265 to -241 (including an AP-1 sequence and flanking NF-κB sequence) inhibits TGF-β stimulus. Neither NF-1 or NF-κB appear to be able to bind this region as NF-1 and NF-κB oligonucleotides failed to compete for complex formation in EMSAs (Chung et al, 1996). As a counter arguement in support of Sp1 as the transcription factor responsible for TGF-β responsiveness, Greenwel et al, (1997) present evidence excluding AP-1 participation and confirming Sp1 requirement. This was achieved by using the pharmacological agents, mithramycin and curcumin, which inhibit the activity of Sp1 and AP-1 respectively. The inhibition of Sp1 activity inhibited TGF-β stimulation of α2(1) procollagen gene expression, while AP-1 inhibition did not. Further exclusion of AP1 as the TGF-β responsive factor was provided by EMSA showing that JunB and c-Jun antibodies failed to recognise proteins in the TbRE-bound complex. AP-1 can, however, bind its recognition sequence in the absence of the Sp1 sequences, indicating that these two factors may bind the TbRE in a mutually exclusive manner, with Sp1 binding being favoured (Greenwel et al, 1997). Although the evidence in the case of Sp1 is certainly very convincing, it is possible that within a given cell type and under differing experimental conditions, DNA elements can be used selectively by different transcription factors.

The stimulation of type I collagen gene expression by TGF-β can also be suppressed by other exogenous factors. Tumour necrosis factor α (TNF-α) and interferon-γ (IFN-γ) both suppress stimulation of the human type I collagen gene expression by TGF-β. TNF-α has a marked effect on suppression of α2(1) procollagen promoter activity, while IFN-γ only has a minimal effect at a transcriptional level (Kahari et al, 1990). TNF-α treatment of cells has also been shown to result in the increased binding of a trans-acting factor which binds a negative element adjacent to the TGF-β-response element (TbRE) (Inagaki et al, 1995b). It has therefore been postulated that TNF-α opposes TGF-β stimulation of the human α2(1) collagen gene via overlapping nuclear

signalling pathways. A recent study showed that TGF-β-induced α2(1) collagen gene expression can also be inhibited by the basic fibroblast growth factor (bFGF). bFGF is capable of suppressing COL1A2 promoter activity in both normal and scleroderma fibroblast, suggesting a transcriptional inhibitory mechanism for bFGF (Ichiki et al, 1997).

To summarise section 1.2.3; A complex set of transcription factors interact with the proximal promoters of the α1(1) and α2(1) procollagen genes. These factors and the *cis*-elements they interact with are essential in the regulation of these genes under different physiologic and pathologic conditions. A summary of the *cis*-elements and their corresponding trans-acting factors is shown in fig 1.1.

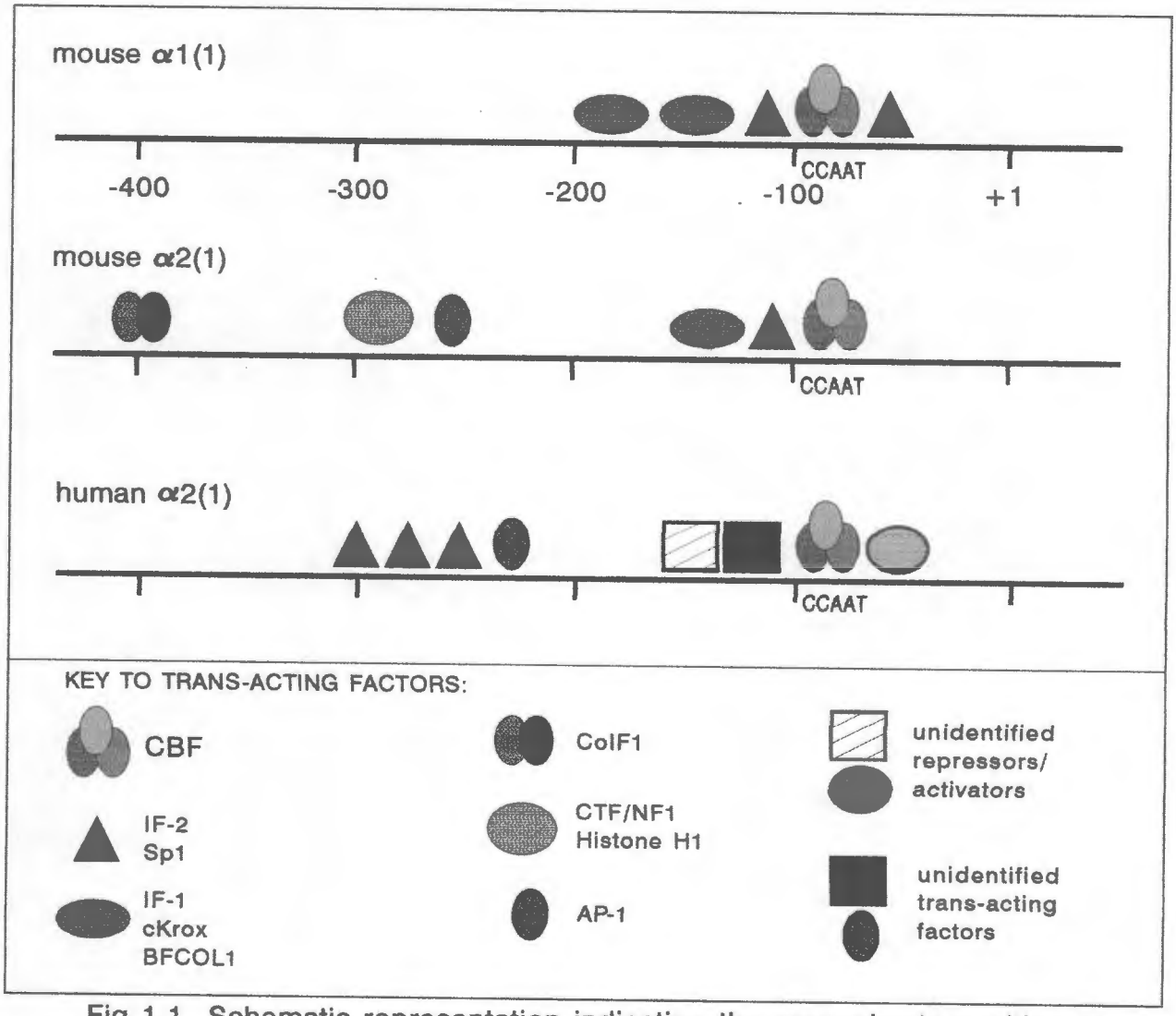


Fig 1.1 Schematic representation indicating the approximate position of trans-acting factor interactions with the proximal promoters of the type I collagen genes. Note: IF-2 is reported to bind the mouse promoters.

1.2.4 Cell- and tissue-specific expression of the type I collagen genes

1.2.4.1 The Alpha 2(1) procollagen gene

Type I collagen gene expression is confined predominantly to fibroblast and osteoblastic cell lines in certain tissues, and is clearly maintained by cell- and tissue-specific regulatory mechanisms. A number of cis-regulatory elements located in the proximal promoter, 5' flanking sequences, the first intron and far upstream sequences of the α2(1) procollagen gene are thought to play a role in the tissue- and cell-specific expression of the gene. Early studies showed that tissue specific expression of the type I collagen genes can be determined at both transcriptional and post-transcriptional levels (Focht and Adams, 1984). Expression of the -2000 to +54 promoter of the mouse α2(1) promoter linked to a reporter gene in transgenic mice showed high levels of expression in many tissues. This expression was similar to that of the endogenous gene (Khillian et al, 1986). Using deletion constructs, Schmidt et al, (1986) also showed that the -2000 to +54 segment of the mouse α2(1) procollagen gene is sufficient for the cell specific expression of the promoter and that two segments (-979 to -502) and (-346 to -104) are required for optimal expression of the chimeric gene. Subsequently, the minimal sequences required for the cell and tissue-specific expression of the α2(1) procollagen gene have been localised to the -350 to +54 proximal promoter albeit at levels lower than that of the endogenous gene (Boast et al, 1990, Niederreither et al, 1992). Although unique cell-specific promoter elements have not been defined as yet, a tandem repeat of the -315 and -284 sequence linked to the mininal α2(1) collagen promoter (-41 to +54) showed considerably higher activity in type I collagen producing cell lines (Goldberg et al, 1990). Cell-specific expression of the human α2(1) procollagen promoter has also been associated with trans-acting factor switching within the -107 to +54 region of the promoter (Parker et al, 1992).

A transcriptional enhancer is located in the first intron of the mouse α2(1) procollagen gene. The enhancer (+418 to +1524) functions in an orientation-independent manner

to stimulate both the α2(1) procollagen promoter and a heterologous simian virus 40 early promoter (Rossi and de Crombrugghe, 1987; de Crombrugghe et al, 1990). It should be noted that this intronic sequence has, however, not been shown to act as a transcriptional enhancer in transgenic mice (Goldberg et al, 1992). Two cis-elements have been located in the first intron and shown to bind a collagen intron binding factor (CIBF-1) and affinity purified Sp1. CIBF-1 binding occurs at a 5'-TGTTTTAA-3' motif at +992, while the Sp1 binding sequence is located in a 'GT box' at +871 to +880 (Pogulis and Freytag, 1993). Contrary to the enhancing effect of the first intron in the mouse α2(1) collagen gene, this region of the human α2(1) procollagen gene, has an inhibitory effect. The human first intron contains two AP-1 motifs and a GT stretch. These sequences are absent in the mouse first intron and the low sequence similarity between the human and mouse α2(1) first introns could account for the contradictary results (Sherwood et al, 1990). A transcriptional enhancer is also located in the far upstream region of the mouse α2(1) collagen promoter. Sequences between -13.5 to -19.5 kb have been shown to stimulate the cell-specific expression of the 350 bp α2(1) proximal promoter. Interestingly, the same DNA segment was able to activate the cell-specific expression of the α1(1) collagen promoter (Bou-Gharios et al, 1996). A number of other collagen genes contain activator and repressor elements in intronic sequences. Examples of these include the α2(XI), α1(II), α1(X), α1(1) genes, where sequences within the first intron contain positive regulatory elements (Tsumaki et al, 1996, Krebsbach et al, 1996, Beier et al, 1997, section 1.2.4.2) and the α2(IV) gene, where sequences within the third intron has an inhibitory effect on gene expression (Haniel et al, 1995).

1.2.4.2 The Alpha 1(1) collagen gene

Tissue-specific expression of the human α1(1) collagen gene is effectively achieved with the -300 to +100 proximal promoter region (Bornstein et al, 1987; Rossouw et al, 1987; Sokolov et al, 1995). Similar experiments with the mouse α1(1) collagen gene identified 220 bp of the proximal promoter as being sufficient to direct tissue-specific expression of the mouse gene (Rippe et al, 1989). In combination with the proximal

promoter, other 5' flanking regions in the α1(1) promoter have also been identified as important in expression of the gene. One of these, a major tissue-specific enhancer, is located in the 5' flanking region of the human α1(1) collagen promoter (-2300 to -444) and confers high-level, position-independent expression of the human α1(1) collagen gene in bone, tail and skin (Slack et al, 1991). Subsequent investigations indicated that sequences up to -804 of the human α1(1) collagen gene could drive expression of a CAT reporter gene (Thompson et al, 1991). Deletion of the sequences between -625 and -442 resulted in significant decreased expression, while deletion of the -804 to -609 region resulted in elevated activity indicating that tissue-specificity of the human gene is determined by both positive and negative cis-regulatory elements in the promoter (Simkevich et al, 1992).

Bedalov et al, (1994) reported that 3.5 kb of the rat promoter fused to a CAT reporter gene had similar expression patterns as that of the endogenous gene. Sequences located upstream in the rat α1(1) collagen promoter have potential osteoblast- and odontoblast-specific stimulatory elements (-3521 to -1672) (Pavlin et al, 1992). In support of this, the sequences between -2.3 and -1.7 kb have been shown to be required for COL1A1 expression in bones and teeth, while sequences between -3.5 and -1.7 are required for expression in tendon (Begdanovic, 1994). Seperate cis-elements in the mouse α1(1) collagen upstream promoter can also direct cell-specific expression in transgenic mice. Rossert et al, (1995) identified three cell-specific elements in the 3.2 kb mouse α1(1) collagen promoter. Each of these elements confers α1(1) expression in different collagen producing cell types. The minimal sequence required for high level expression in osteoblasts is a 117 bp segment (-1656 to -1540) fused to the proximal 220 bp promoter. A 28 bp region (-1656 to -1628) is proposed to bind cell-specific proteins for expression in osteoblasts while the -1575 to -1540 region may bind ubiquitous proteins involved in high level expression of the gene (Rossert et al, 1996). A similar study showed that two different *cis*-regulatory elements direct cell-specific expression of the α1(1) procollagen gene in stellate, skin and tendon cells. While the 0.44 kb 5' promoter region (including the first intron) was

sufficient for expression in stellate cells, high level expression in skin and tendon required the -2.3 to -0.44 kb region (Houglum et al, 1995). Recently it has been shown that a 1.9 kb fragment of the α1(1) gene containing 0.48 kb of the promoter and most of the first intron and exon is sufficient to drive expression of both the COL1A1 and COL1A2 gene in transgenic mice (Ala-Kokko et al, 1996).

Certain evidence exists for the importance of the first intron in optimal tissue and cell-specific expression of the α1(1) collagen gene. There is, however, considerable controversy regarding the exact function of the first intron as the results obtained in transfection studies and transgenic mice are not consistent. Initial experiments indicating a role for the first intron stemmed from the development of Mov13 mice in which retroviral insertion into the first intron resulted in the complete inhibition of the α1(1) collagen gene, leading to the subsequent death of these mice (Schnieke et al, 1983, Harbers et al, 1984). The exact mechanism of this inhibition remains unclear, and Chan et al, (1991) excluded DNA methylation as being responsible for the inhibition and proposed that alterations in chromatin structure may be responsible. Interestingly, Kratachwil et al, (1989) report that although the α1(1) collagen gene is inhibited in the fibroblasts of Mov13 mutants, expression of the gene is detected in odontoblasts This suggested that retroviral insertion in the first intron interferes with tissue-specific transcriptional control of the α1(1) gene. Both positive and negative *cis*-elements are located in the first intron of the human gene (Bornstein et al, 1987, 1988; Bornstein and McKay, 1988; Liska et al, 1990; Boast et al, 1990; Simkevich et al, 1992; Liska et al, 1992). A highly conserved AP-1 binding site has been identified in the first intron and is thought to play a role in α1(1) collagen gene expression (Liska et al, 1990; Katia et al, 1992; Maata et al, 1993). The AP-1 element can act either as a positive or a negative element in a cell-specific manner (Katai et al, 1992). Although an initial study on the inhibition of human α1(1) collagen gene expression by Ras indicated that the first intron had no role in this down-regulation (Slack et al, 1992), a later investigation showed that loss of AP-1 binding at the intronic site, could in part, account for the inhibition of α1(1) collagen gene expression observed in these cells

(Slack et al, 1995). Additional cis-elements in the first intron include two Sp1 binding motifs (+820 to +1093), a viral core enhancer element (+999 to +1007) and a DNase1 footprint (+951 to +978) (Bornstein et al, 1987; Rossouw et al, 1987; Liska et al, 1992).

While a number of studies emphasize the role of the first intron as that of a positive regulatory element (Bornstein et al, 1987, 1988; Slack et al, 1991; Simkevich et al, 1992; Liska et al, 1994), others have found that intronic sequences are not essential for the tissue-specific expression of the gene (Olsen et al, 1991; Bornstein and McKay, 1988; Sherwood and Bornstein, 1990; Boast et al, 1990; Sokolov et al, 1993, 1995; Houglum et al, 1995). A thorough attempt to assess the literature on the nature of the α1(1) first intron has been made by Bornstein, (1996). The author proposes a number of possible causes for the conflicting data, one of particular interest to this thesis questions the placement of functionally equivalent regulatory elements in different species. Indeed, the results presented in section 3 of this thesis, suggest that subtle differences in the human and rodent α2(1) procollagen promoters may result in significant differences in promoter activity. These observations therefore suggest that cross-species experiments should be viewed with care.

It is also likely that co-operative interactions between ubiquitous promoter-binding factors and trans-acting factors and *cis*-elements not identified as yet are essential for the tissue- and cell-specific expression of the type I collagen genes. As such, the search for other regulatory sequences in the type I collagen genes are under way. A recent study reported that a region in the mouse α1(1) collagen gene contains the binding sites for two upstream stimulating factors, USF-1 and USF-2 (Rippe et al, 1997).

1.3 REPRESSION OF THE TYPE I COLLAGEN GENES

While transcriptional activation mechanisms are well documented for most genes, including the type I collagen genes, considerably less is known about the processes involved in transcriptional repression (see Johnson, 1995). There are, however, models that convincingly argue for the presence of proteins which inhibit gene transcription.

Certainly, transcriptional repression is as important an event as activation for a number of reasons, all which are pertinent to the type I collagen genes. These include; (a) negative regulation of genes may be required in certain cell and tissue types, (b) some genes are developmentally controlled, hence repressor mechanisms play a key role in developmental expression, (c) genes which can be induced in response to a stimulus or stimuli need to be "turned off" when the stimulus is removed, eg collagen synthesis in wound healing.

The following section will review repression of the type I collagen genes under two subsections; (1) Mechanisms by which repression can be achieved and (2) Events that down-regulate type I collagen gene expression by utilising the mechanisms described in (1).

1.3.1 Mechanisms of achieving repression

1.3.1.1 Repression by trans-acting factors

Essentially three broad mechanisms of transcriptional repression involving trans-acting factors exist, these include the inhibition of binding of activator proteins, blocking of activation of the transcription initiation complex and direct binding of a trans-acting repressor which leads to silencing of genes (reviewed by Levine and Manley, 1989; Renkawitz, 1990; Clark and Docherty, 1993). Repressor mechanisms of the kind where activators and repressors compete for the same control element account for a diversity of transcriptional control. Examples of competition between transcription factors resulting in repression have been reported for several genes. These usually involve overlapping or closely linked sequences within the promoter of the gene. For example, a 91 kD protein which binds the GC rich sequences in the EGF-receptor and β-actin promoters, inhibits transcription of these genes by preventing activators from binding (Kageyame and Pastan, 1989). A repressor of the sperm histone H2B-1 gene, the CCAAT-displacement protein (CDP) acts as a competitor for interaction between the CCAAT binding protein and DNA (Barberis et al, 1987). CDP is also implicated in blocking the binding of the CCAAT binding factor, CP-1, to the myelomonocytic specific gp91-phox promoter (Skalnik et al, 1991; Lievens et al, 1995). Similar results were obtained for the bZIP factor E4BP4 which overlaps in DNA binding-site

specificity with the transcriptional activators CREB and ATF (Cowell and Hurst, 1994).

It can be speculated that the inhibitory factors, IF-1 and IF-2 identified as inhibitors of the type I collagen genes do so in a competitive manner in view of the fact that a number of other positive trans-acting factors bind to the same DNA element (section 1.2.3.1.3). Data presented by Parker et al, (1992) (and section 2 of this thesis), suggest that two trans-acting factors, complex II (potential repressor) and complex III (potential activator) interact with the same DNA element in the human α2(1) procollagen promoter. It is possible that competition for binding to the recognition sequence may occur in cell lines where both these factors are present. A negative DNA element located between -173 and -155 in the human α2(1) collagen promoter also binds a potential transcriptional repressor (Ihn et al, 1996). Negative regulatory elements are also located in the α1(1) procollagen promoter between -804 to -609 (Simkevich et al, 1992).

1.3.1.2 Repression via methylation

Hypermethylation of genes are generally associated with the inactivation of these genes, while under-methylation is associated with transcriptionally active genes (reviewed by Boyes and Bird, 1991). A number of studies indicate that hypermethylation of the type I collagen genes may result in its inactivation (Parker et al, 1982, 1986; Smith and Marsilio, 1988). It was initially thought that retroviral insertion into the intronic sequence of the COL1A1 gene in Mov13 mice resulting in hypermethylation of the gene was responsible for inactivation of the gene (Jahner and Jaenisch, 1985). Chan et al, (1992) however, reported that even in the absence of methylation, COL1A1 promoter activity was suppressed in a cell-specific manner in Mov13 cells, and speculated that the insertion of the retroviral DNA in the first intron probably prevented the formation of correct methylation patterns and chromatin structure resulting in inactivation of the gene. DNA methylation has also been implicated in the inhibition of α2(1) collagen gene expression in W8 cells.

Hypermethylation within the promoter and first exon of the gene is observed in these cells and methylation of two sequences (-345 to +58 and -218 to +58) specifically inhibited expression of the gene (Guenette et al, 1992). The exact mechanism by which methylation inhibits transcription is still poorly understood. Chromatin structure, however, plays an important role in this process, since methylated DNA forms chromatin which is insensitive to DNase1 treatment (Keshet et al, 1986). Several proteins which bind methylated DNA have been characterised and are thought to be involved in the repression of methylated genes (Lewis et al, 1992; Jost and Hofsteenge, 1992). Repression of the murine α1(1) collagen promoter by DNA methylation, however, appears to occur by a mechanism independent of methyl C-binding proteins (Rhodes et al, 1994). The inhibitory effect of DNA methylation in vertebrate cells therefore probably results from cooperativity between chromatin structure and methyl C-binding proteins.

1.3.2 Events resulting in down-regulation of type I collagen synthesis

1.3.2.1 Repression by exogenous factors

The cytokines TNF-α and IFN-γ can synergistically suppress the stimulatory effect of TGF-β on the type I collagen genes (section 1.1.3.3). TNF-α inhibits α2(1) collagen gene expression at a transcriptional level, while IFN-γ inhibition is postulated to function at a post-transcriptional level (Kahari et al, 1990). Cytokines such as Interleukin-1 and TNF-γ may have both inhibitory and stimulatory effects on type I collagen gene expression (reviewed by Slack et al, 1993). The major negative regulator of type I collagen gene expression appears to be IFN-γ. This inhibition occurs at a post-transcriptional level since no change in primary transcript formation is observed while steady state mRNA levels are reduced by 60-70% by 48 hours after exposure of cells to IFN-γ (Slack et al, 1993). Down-regulation of the types I and III collagen genes can also be achieved in a dose-dependent manner by the melanoma growth-stimulatory activity cytokine, [(MGSA)/GRO] (Unemori et al, 1993). Other exogenous factors such as 1,25-Dihydroxyvitamin D_3 (vitamin D) has both stimulatory and inhibitory

effects on type I collagen gene expression. A 41 bp fragment in the α1(1) procollagen promoter is reported to be involved in vitamin D mediated transcriptional repression in osteoblastic cells (-2256 to -2216). This sequence shares homology with vitamin D-reponsive elements identified in the osteocalcin gene (Pavlin et al, 1994).

1.3.2.2 Repression by transformation

Collagen synthesis is markedly reduced upon transformation by several agents (Bornstein and Sage, 1989; Schreir et al, 1988). Both chemically and virally transformed fibroblasts exhibit reduced type I collagen synthesis (Parker and Fitschen, 1980; Parker et al, 1989; Guenette et al, 1992). Transformation by oncogenes such as v-src, v-mos and v-fos all have an inhibitory effect on transcription of the type I collagen genes (Sandmeyer et al, 1981; Schmidt et al, 1985; Hoemann and Zarbl, 1990). The c-Myc oncoprotein specifically inhibits the expression of α1(1), α2(1) and α3(IV) collagen genes in mouse 3T3-L1 cells (Yang et al, 1991). Similarly, constitutive expression of the Fos oncoprotein inhibits the expression of α1(1) collagen gene in MC3T3-E1 cells (Kuroki et al, 1992). Cells that have been immortalised by mutant p53 or the Ras oncogene have reduced levels of both α1(1) and α2(1) collagen gene expression (Eizenberg and Oren, 1991; Slack et al, 1992).

1.4 SIGNAL TRANSDUCTION AND COLLAGEN EXPRESSION

1.4.1 The effect of signalling pathways on collagen gene expression

The fact that type I collagen gene expression can be up or down regulated by a number of oncoproteins and cytokines implies that cytoplasmic kinase cascades are associated with the regulation of the collagen genes. These responses vary dramatically in a cell-specific manner indicating that signal transduction by different stimuli may occur via cell-specific mechanisms. A stimulator of protein kinase C activity, phorbol 12-myristate 13-acetate (PMA) causes a decrease in collagen synthesis in murine osteoblasts (Harrison et al, 1990), mouse epidermal cells (Sobel et al, 1983), NIH3T3

mouse fibroblast (Rabin et al, 1986) and in human fetal skin fibroblast (Greenwel et al, 1997). Contrary to these findings, PMA has a stimulatory effect on α2(1) collagen gene expression in 3T3-L1 cells when compared to a PMA-non-responsive cell variant (Stuiver et al, 1991). Together these studies show that protein kinase C mediated signalling pathways have a regulatory effect on the expression of the type I collagen genes.

The response of cells to TGFβ is mediated via signal tranduction pathways which involve the activation of protein kinase C and other transmembrane serine/threonine kinases. These kinases in turn lead to the phosphorylation of other effector substrates such as Raf-1 kinase, myristorylated alanine rich C kinase substrate (MARCKS) and cytoplasmic SOS (reviewed by Massague and Weis-Garcia, 1996; Mufson, 1997; Hunter, 1997). The action of protein kinase C has also been implicated in the down-regulation of type II collagen gene transcription in chondrocytes by TGF-β and fibroblast growth factor (FGF) (Bradham et al, 1994).

It is thus possible that a cascade of events resulting from the activation of kinases such as protein kinase C may play a key role in the regulation of expression of the type I collagen genes via modification of transcription factors. Activated protein kinase C and other kinases such as mitogen-acitivated protein kinase (MAP kinase) can translocate to the nucleus where they phosphorylate a number of transcription factors (Boulikas, 1995). There are three main levels via which phosphorylation events can modulate transcription factor activity. These include entry of transcription factors into the nucleus, modulation of the DNA binding activity (positively or negatively) and, finally, phosphorylation can affect the transactivation activity of transcription factors (reviewed by Hunter and Karin, 1992; Sassone-Corsi, 1994; Karin and Hunter, 1995). A well characterised example where phosphorylation events lead to nuclear import is that of NF-κB. Phosphorylation of the inhibitor of NF-κB, IκB, results in the dissociation of the NF-κB-IκB complex followed by rapid translocation of NF-κB to the nucleus (Ghosh and Baltimore, 1990). The regulation of DNA binding activity by

phosphorylation has been reported for numerous transcription factors. Most reports, however, implicate inhibition of DNA binding in response to phosphorylation at specific sites in a transcription factor. These include c-Jun, Max, Oct-1 (Hunter and Karin, 1992), Ets, ERF (an Ets domain protein) (Sgouras et al, 1995) and CREB (Mahonety et al, 1992). While direct phosphorylation of the interferon consensus binding protein (ICSBP), a transcriptional repressor, prevents DNA binding, phosporylation of a specific tyrosine allows ICSBP to bind its' target DNA via association with other transcription factors (interferon regulator factors, IRF-1 and 2) (Sharf et al, 1997). One of the best documented examples of the role of protein phosphorylation in the regulation of transcription is that of the components of AP-1, Jun and Fos. Phosphorylation causes an increase in the transcriptional activities of the c-fos and c-jun genes. The newly synthesised Jun and Fos proteins combine to form the AP-1 complex which on phosphorylation of the Jun activation domain enhances the transcriptional activity of AP-1 (reviewed by Hunter and Karin, 1992).

Although well documented examples of stimulation of DNA binding activity by phosphorylation are rare, a recent report shows that phosphorylation of ATF-1 enhances DNA binding and the subsequent activation of the Na,K-ATPase α1 subunit gene promoter. Dephosphorylation of nuclear extracts results in the loss of ATF-1 DNA binding at a ATF/CRE binding site in the promoter (Kobayashi et al, 1997). In section 3.2.5 of this thesis we report that dephosphorylation and inhibition of kinase activity also result in reduced binding of transcription factors to the human α2(1) procollagen promoter. Signal transduction pathways associated with Raf and MAP kinase have been implicated in the regulation of rat α1(1) collagen expression. MAP kinase and Raf kinase activities are associated with the stimulation and inhibition of the α1(1) collagen promoter respectively. These kinase activities are mapped to the 5'-untranslated region of the gene, including the proximal Sp1, NF-1 and the TGF-β activating element (Davis et al, 1996).

The inhibition of protein phosphatases 1 and 2A result in suppression of type I collagen gene regulation. This implies that dephosphorylation events are also required

for the activation of the type I collagen genes (Westermarck et al, 1995; Wang and Raghow, 1996). It is possible that Sp1 may be one such factor, since dephosphorylation of Sp1 results in enhanced DNA binding and transcriptional activation of the acetyl-coA carboxylase gene (Daniel et al, 1996). This is in support of the proposal by Greenwel et al, (1995) that tyrosine dephosphorylation of nuclear proteins, possibly Sp1 mimic TGF-β stimulation of the human α2(1) procollagen gene. It is therefore highly probable that signal transduction pathways resulting in both the phosphorylation and dephosphorylation of transcription factors have a significant effect on the regulation of type I collagen gene transcription.

1.4.2 The effect of intact type I collagen on signalling pathways

An intact type I collagen heterotrimer and other components of the extracellular matrix (ECM) are important for cell adhesion and anchorage dependent cell growth. Gross alterations in ECM components can result in the progression and metastasis of tumours (reviewed by Lochter and Bissel, 1995). In cell lines representative of a tumorigenic state, transfection of components of type I collagen results in partial reversion of these cells to a 'normal' state. Transfection of the full length α2(1) collagen cDNA into chemically transformed cell line, W8, where no α2(1) collagen chains are synthesised, resulted in a decrease in colony formation in soft agar and an increase in cell adhesion (Lim et al, 1994). Similarly, Travers et al, (1996) reported that restoration of the α2(1) collagen chain in Ras-transformed mouse fibroblasts result in the suppression of malignant properties. The restoration of type I collagen in these cells was correlated with the down-regulation of a Ras-responsive gene, suggesting that type I collagen has a functional role in oncoprotein signalling pathways associated with tumorigenicity.

The mechanism by which type I collagen is involved in signal transduction occurs via the integrin family of proteins. ECM components such as type I collagen are ligands for members of the integrin family of heterodimers. The major integrin involved in type I collagen-mediated-signalling are α1β1, α2β1 and α1β3 (Sonnenberg, 1993,

Haas and Plow, 1994). These integrins are responsible for regulating ECM remodulation and α1β1 integrin receptor mediates signals which induce down regulation of type I collagen gene expression, while α2β1 integrin is involved in the upregulation of collagenase (MMP-1) expression (Langholz et al, 1995). Type I collagen-integrin mediated signal transduction can involve both the stimulation and inhibition of protein kinases. Fibrillar collagens specifically regulate integrin signalling that lead to the stimulation of cdk2 inhibitors such as p27 and p21 resulting in the inhibition of smooth muscle cell proliferation (Koyama et al, 1996).

Type I collagen itself may also have a feedback regulatory effect on its own expression. This proposal stems from experiments showing that the N-terminal propeptide of the α1(1) collagen chain is associated with inhibition of endogenously produced type I collagen (Fouser et al, 1991). The exact mechanism by which this inhibition occurs is unknown, but down-regulation of type I collagen synthesis by fibroblasts grown in collagen lattices has been shown to involve both transcriptional and post-transcriptional mechanisms (Eckes et al, 1993).

In conclusion, cellular control of and by type I collagen therefore involves a cascade of signal transduction pathways most of which still have to be identified. These signalling events may have an effect on transcription either in a direct manner by phosphorylation/dephosphorylation of specific transcription factors or in an indirect manner involving other kinase sensitive pathways.

2. ANALYSIS OF COLLAGEN SYNTHESIS IN DIFFERENTIATED CELL LINES

CHAPTER 2:

ANALYSIS OF COLLAGEN SYNTHESIS IN DIFFERENTIATED CELL LINES

2.1 INTRODUCTION

Cellular and tissue differentiation is characterised by the expression of a vast array of specialised genes encoding cell-specific proteins. One such protein is type I collagen, the synthesis of which varies widely in different tissues and cell types. These variations occur during processes such as embryogenesis, development, pathologic conditions and in the adult organism in a tissue and/or cell-specific manner (Parker et al, 1989; Ramirez et al, 1989; Slack et al, 1991; Sokolov et al, 1995). The expression of type I collagen is confined mostly to cells of mesenchymal origin, while it is not produced by cells derived from the ectoderm, endoderm or hematopoetic cells (Bornstein and Sage, 1989; Ramirez et al, 1990; Prockop and Kivirikko, 1995).

Studies to date have concentrated mostly on the transcriptional mechanisms involved in type I collagen synthesis and both transcriptional and post-transcriptional mechanisms play a role in the cell and tissue-specific expression of the type I collagen genes (Focht and Adam, 1984; Simkevich et al, 1992; Slack et al, 1993). Trans-acting factors have been implicated in the regulation of both the $\alpha 1(1)$ and $\alpha 2(1)$ collagen genes (Maity et al, 1988; Karsenty and de Crombrugghe,

1991; Hatamochi et al, 1988; 1993; Tamaki et al, 1995, Chen et al, 1997). The minimal promoter region (-350 to +54) has been found to be all that is required for the correct expression of the α2(1) collagen gene (Neiderreither et al, 1992). Subsequently, most of the transcription factors identified in the regulation of type I collagen gene expression were shown to bind within this region of the promoter. In the mouse α2(1) procollagen promoter at least three sites within the minimal promoter have been identified as important for protein binding and promoter activity (de Crombrugghe et al, 1990 and Neiderreither et al, 1992). These sites include an inverted CCAAT-box between -84 and -80, a CAGA box between -250 and -247 and an NF-1-like site between -312 and -300. The transcription factors which bind to these sites are the CCAAT binding factor (CBF), the CAGA-binding factor and possibly NF-1, respectively (Oikarinen et al, 1987 and Hatamochi et al, 1988). NF-1 binding is thought to activate the mouse α2(1) collagen gene via a TGF-β response element. A reporter construct containing the -2000 to +54 region of the α2(1) procollagen promoter, displays increased expression upon TGF-β treatment. When the NF-1 binding site was included in a TGF-β unresponsive SV40 promoter, inducibility of the promoter was observed (Rossi et al, 1988 and de Crombrugghe et al, 1990). Inagaki et al, (1995) however, have shown that although there is high sequence homology between the mouse and human α2(1) procollagen promoters, TGF-β stimulation of the human α2(1) collagen gene promoter is thought to be achieved by the increased affinity of an Sp1-DNA-protein complex binding to the TbRE in the human α2(1) procollagen promoter. They implicated Sp1 in mediating the TGF-β effect and not NF-1 as thought to be the case in the mouse promoter reported by Rossi et al, (1988). Chung et al, (1996), however, exclude Sp1 binding sites as a requirement for the TGF-β response in the human α2(1) procollagen promoter. They argue that TGF-β responsiveness in the COL1A2 promoter occurs via AP-1 binding between -265 and -241. Subsequently, additional support for Sp1 as the transcription factor

involved in TGF-β responsiveness in the human α2(1) procollagen promoter has been provided (Greenwel et al, 1997; section 1.2.3.3).

The best characterised transcription factor involved in the regulation of the type I collagen gene is the mouse CCAAT binding protein (CBF). Hatamochi et al (1988), identified the mouse CBF as a member of a family of fairly ubiquitous proteins which bind CCAAT motifs in promoters of eukaryotic genes. The mouse CBF comprises three different components (A,B,and C) as previously described (Maity et al, 1992, Vuorio et al, 1995, section 1.2.3.1.1).

Most of the studies on the α2(1) procollagen gene have focused on the mouse proximal promoter (-350 to +54) while the factors binding within the corresponding region of the human α2(1) promoter are poorly characterised. Recent studies by Inagaki et al, (1994, 1995) and Tamaki et al, (1995) propose GC-rich regions containing Sp1 binding sites as regulatory elements in the human α2(1) procollagen promoter, while Chung et al (1996) reported that AP-1 binding is also essential for regulation of the gene. We have embarked on a study aimed at elucidating the factors which interact with the proximal COL1A2 promoter. A previous study has identified at least two factors binding within the region between -107 and -60 (including the inverted CCAAT box) of the COL1A2 promoter in a collagen producing cell line, CT-1 (Parker et al, 1992). An additional DNA-protein complex was observed in a cell line, SVWI-38, where no α2(1) collagen synthesis occurs. These differences implied negative regulation of the α2(1) procollagen gene in the SVWI-38 cell line.

Since type I collagen synthesis varied so widely in the two cell types, it was decided to explore the possibility of a correlation between α2(1) collagen synthesis and transcription factor binding to the proximal promoter of the gene in nine different cultured cell lines. To address this issue, a cross section of cell types

consisting of cells synthesising the α2(1) chain of type I collagen and others where collagen synthesis is repressed was selected. The ability of these cells to produce type I collagen was examined by investigating collagen synthesis, steady state mRNA levels and primary transcript formation. This was followed by analysis of transcription factor binding to the proximal promoter of the α2(1) procollagen gene in the different cell lines and seeking a correlation between transcription factor binding and the ability of the cell line to synthesise the α2(1) chain of type I collagen.

2.2 RESULTS

The findings by Parker et al, (1992) showed that α2(1) collagen synthesis is associated with differences in trans-acting factor binding to the proximal promoter of the gene. The collagen producing cell line (CT-1) contains predominantly complex I and III proteins, while the SVWI-38 cell line, where type I collagen expression is abolished, contains three DNA-protein complexes, I, II and III (fig 2.1). The aim of this study was therefore to determine whether differential transcrtiption factor binding may be a more general cell-specific mechanism of control. Nine human cell lines displaying a wide range of type I collagen synthetic abilities were used in this investigation. The cells types ranged from normal lung fibroblasts, transformed lung fibroblasts, skin fibroblasts, fibrosarcoma, hepatocellular carcinoma, lymphoblastoid and breast epithelial cells. The experiments performed in this study included; (i) measuring type I collagen protein synthesis by labelling cells with ^{3}H-proline and analysing secreted proteins; (ii) determining the steady state levels of α2(1) mRNA by Northern hybridisation analysis and (iii) determining primary transcript formation by nuclear run-on transcription assays. (iv) Nuclear proteins extracted from these cell lines were used in EMSA's using the -107 to -60 fragment of the COL1A2 promoter.

2.2.1 Analysis of type I collagen synthesis

Cells were labelled with ^{3}H-proline, proteins secreted into the medium were extracted and the pepsin-resistant proteins analysed on SDS 7% polyacrylamide gels (section 6.5.1). The amount of pepsin-resistant collagenous material obtained from each cell line is shown in fig 2.2.A. As can be seen from the SDS-polyacrylamide gels (fig 2.2.B) the cells of fibroblast origin; WI-38 (lung fibroblast), CT-1 (WI38 fibroblasts transformed by γ-irradiation) and FG_o (skin fibroblasts), all synthesise both the α1(1) and α2(1) chains of type I collagen.

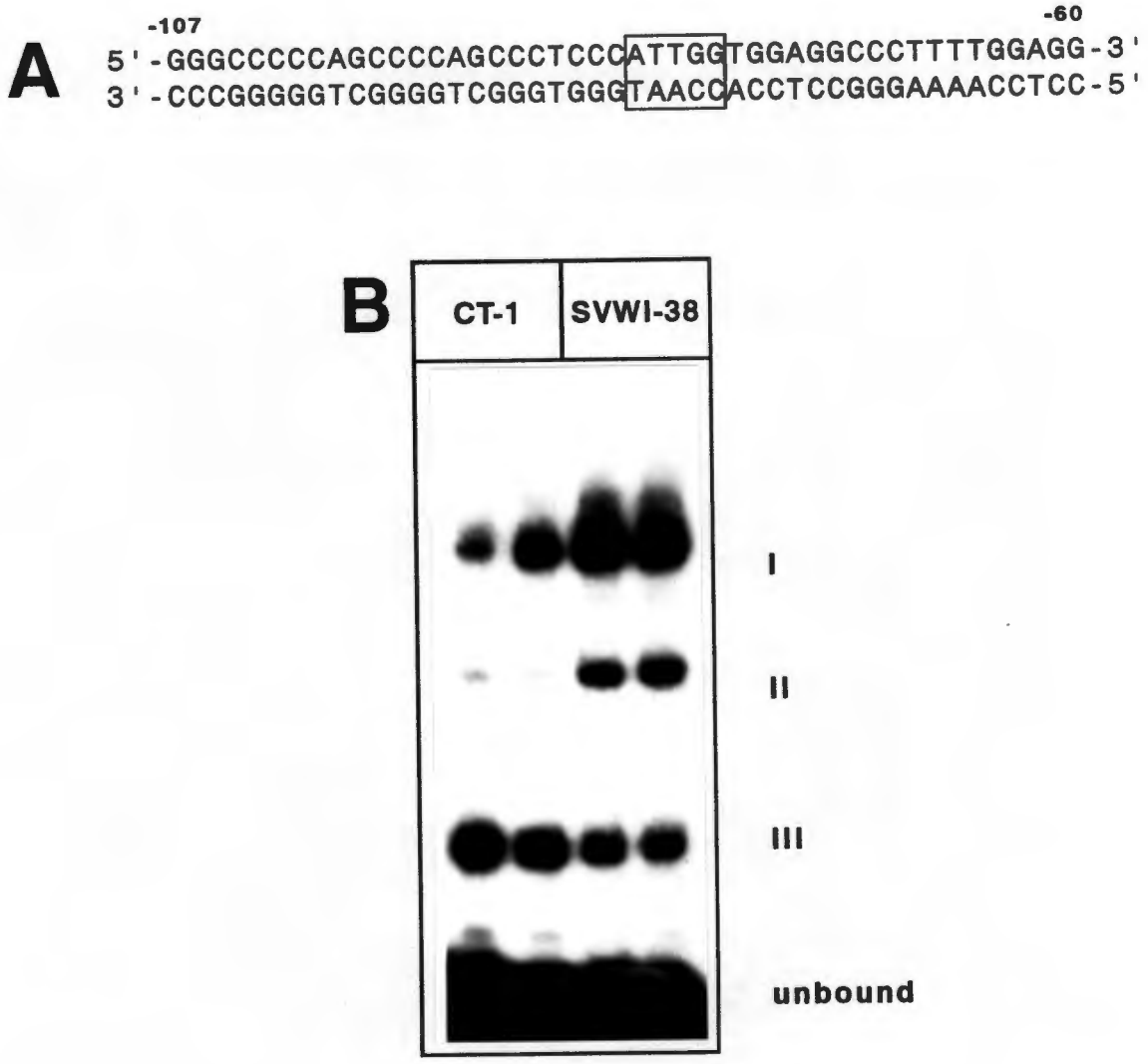


Fig 2.1 Trancription factor binding to the proximal promoter of the human α2(1) procollagen gene. **(A)** Wild type α2(1) promoter sequence between -107 to -60 from transcription initiation site. The inverted CCAAT box is boxed. **(B)** Electrophoretic mobility shift assays using the radioactively labelled DNA fragment showed in (A) as a probe. DNA-protein complexes were analysed in duplicate using 4μg crude nuclear protein from CT-1 or SVWI-38 cells incubated with 1 ng of ^{32}P-labelled DNA (10^4 cpm) and electrophoresed on non-denaturing 5% polyacrylamide gels (section 6.4.3). Dried gels were exposed to X-ray film for 16 hrs. DNA-protein complexes are indicated as I, II and III.

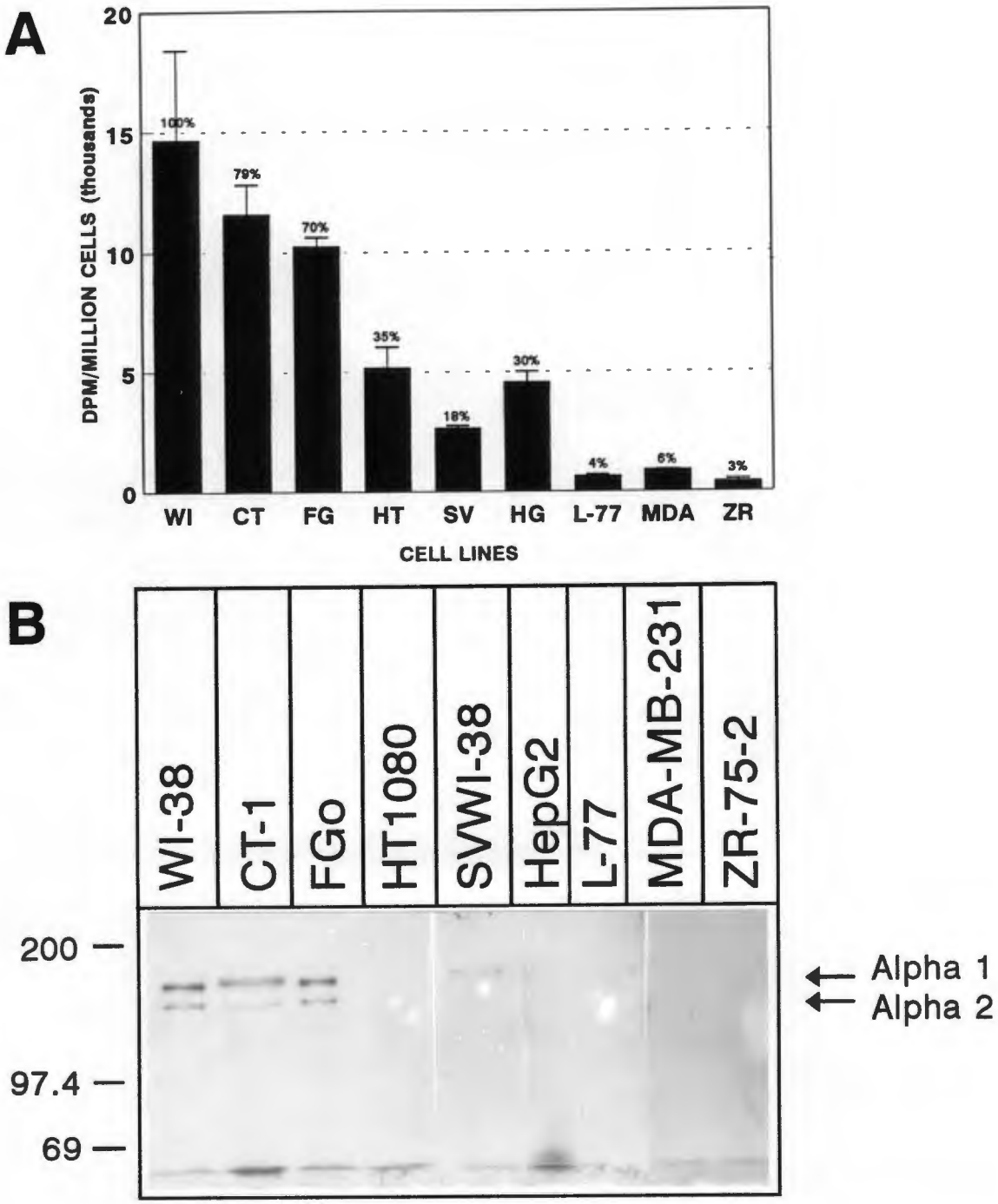


Fig 2.2 Type I collagen synthesis in differentiated cell lines. **(A)** Cells were labelled with ^{3}H-proline, secreted proteins were extracted and treated with pepsin (section 6.1.3). Pepsin-resistant material was precipitated with TCA and radioactive counts determined. Pepsin-resisitant collagenous material were corrected for cell number and is shown for the various cell types. Results are of the mean ± S.D. **(B)** Radioactively labelled pepsin-resistant collagenous material, corrected for cell number was resolved on SDS 7% polyacrylamide gels. Dried gels were exposed to X-ray film for 48 hours. Molecular weight markers are indicated on the left and the position of the α1(1) and α2(1) procollagen chains are indicated on the right hand side of the figure.

The HT1080 (fibrosarcoma), HepG2 (hepatocellular carcinoma), L-77 (lymphoblasts), MDA-MB-231 and ZR-72-2 (breast cancer epithelial cells) contained no detectable α2(1) or α1(1) collagen, while the SVWI-38 (WI-38 transformed with simian virus-40) contained no α2(1) chains but reduced amounts of α1(1) chains (fig 2.2.B). The α1(1) chain in SVWI-38 cells had an altered electrophoretic mobility due to overmodification of these chains as observed by two-dimensional cyanogen bromide peptided analysis (Parker et al, 1989). Of note, is that the HT1080 cells (fibrosarcoma) did not synthesise any significant amounts of type I collagen.

The cell lines included in this study, showed a varied degree of type I collagen synthesis and were thus suitable for assessing the cell-specific mechanisms of type I collagen synthesis. Since our interest was the elucidation of the transcriptional regulatory mechanism(s) responsible for the observed differences in α2(1) collagen synthesis in these cells, the α2(1) collagen mRNA levels were determined.

2.2.2 Alpha2(1) collagen mRNA levels in differentiated cell lines.

2.2.2.1 Northern Hybridisation Analysis

To determine steady-state mRNA levels, cytoplasmic RNA was isolated from the various cell lines as described by Chomczinsky and Sacchi, (1987). The purity of the RNA was determined spectrophotometrically and 5μg total RNA from each cell line was electrophoresed on 1% agarose-formaldehyde gels (fig 2.3.A). The RNA was transfered to nylon membranes as described in section 6.6.2.1. and hybridised to radioactively labelled α2(1) collagen (Myers, et al 1981) and β-actin probes (Gunning et al, 1983). Results obtained by Northern blot analysis (fig 2.3.B) followed the same trend as the results obtained in the collagen synthesis

experiments in the different cell lines (fig 2.2). This indicated that down-regulation of the α2(1) collagen gene may occur at a transcriptional level in these cell lines. To obtain a direct correlation with transcription, the rate of newly formed type I collagen mRNA transcripts in the different cell lines were examined using nuclear run-on transcription assays.

2.2.2.2 Nuclear run-on transcription assays

The nuclear run-on transcription assay was used to identify the formation of newly transcribed RNA and is therefore a direct indication of transcriptionally active genes. Radioactively labelled, newly transcribed RNA was hybridised to cloned gene fragments to assess the transcriptional activity of the type I collagen genes in the various cell types. The results indicate that differences in α2(1) collagen synthesis occur at a transcriptional level (fig 2.4). Cells of fibroblast origin displayed considerable type I collagen transcripts, while the other cell types showed little to no transcripts. Although no α2(1) collagen protein and mRNA were detected in the HT1080 (fibrosarcoma) cell line (fig 2.3.B), high levels of nascent α1(1) and α2(1) collagen transcripts were detected in this cell line (fig 2.4). A similar finding has been reported in HeLa cells, where type I collagen is transcribed but steady state mRNA or protein are not detected (Furth et al, 1991). The SVWI-38 cell line, which showed no α2(1) chain synthesis or steady state mRNA, contained very low levels of nascent α2(1) run-on transcripts (fig 2.4.A and B). This result is not completely suprising, as it may be possible that the gene may not be entirely silenced in this cell line, but could be in a partially active (or partially repressed) state.

Analysis of type I collagen gene expression, in particular the α2(1) chain, clearly indicated that the levels of type I collagen varied significantly in the different cell lines used in this study. A mechanism that could account for these variations in expression may include the interaction of trans-acting factors with the promoter of

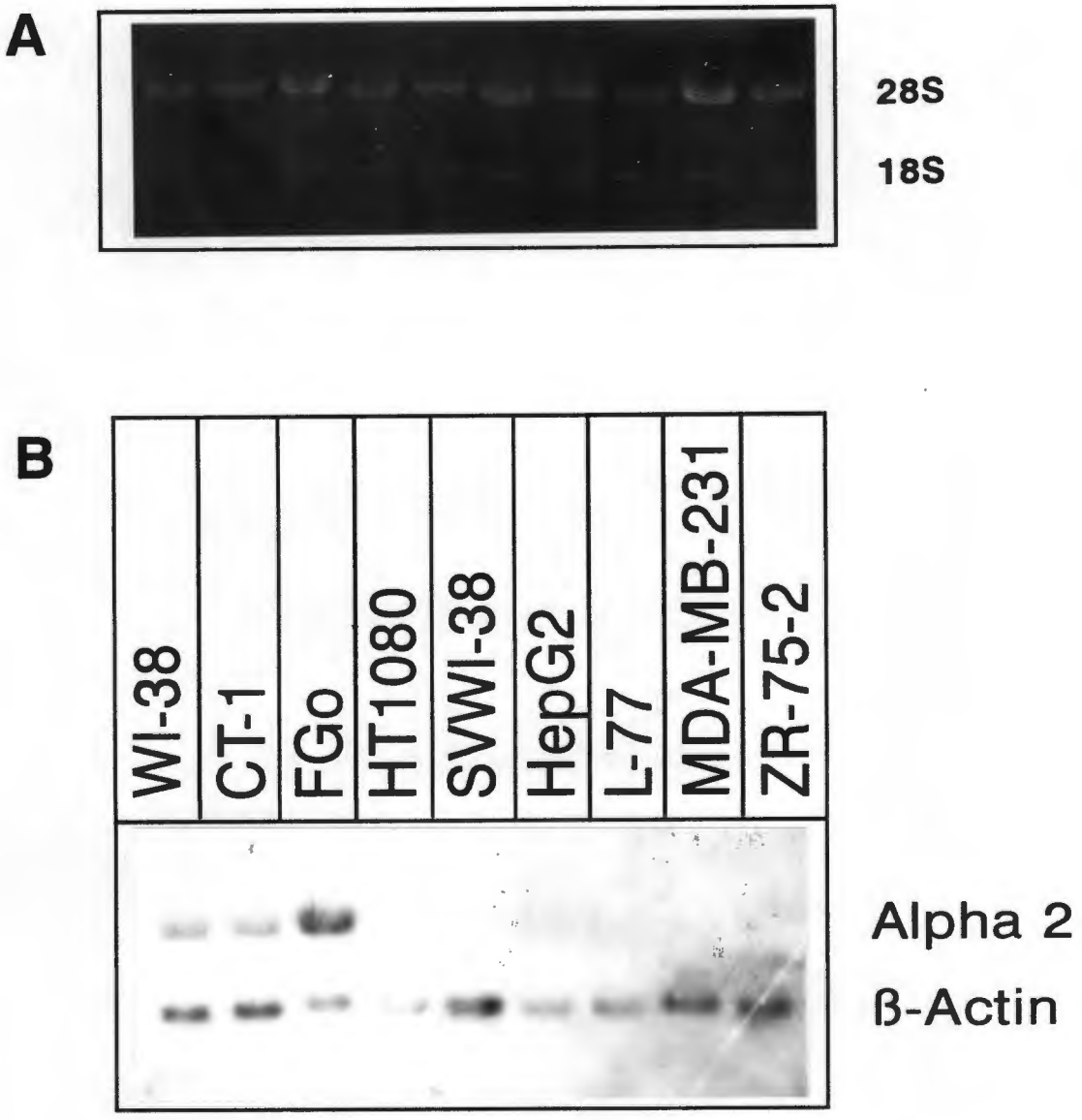


Fig 2.3 Steady state mRNA levels in the different cell lines. **(A)** Cytoplasmic RNA was extracted from the different cell lines by the method of Chomzcinsky and Sacchi, (1987). After determining the purity of the RNA spectrophotometrically, 5μg of total RNA was loaded on 1% agarose formaldehyde gels. The position of the 28S and 18S RNA is indicated on the figure. **(B)** RNA from the gel shown in (A) was transfered to nylon membranes, UV cross-linked and the membranes hybridised to α2(1) collagen specific and a β-actin probes to correct for loading errors. Membranes were exposed to X-ray film for 16 hrs and the position of the α2(1) procollagen and β-actin mRNA bands are indicated.

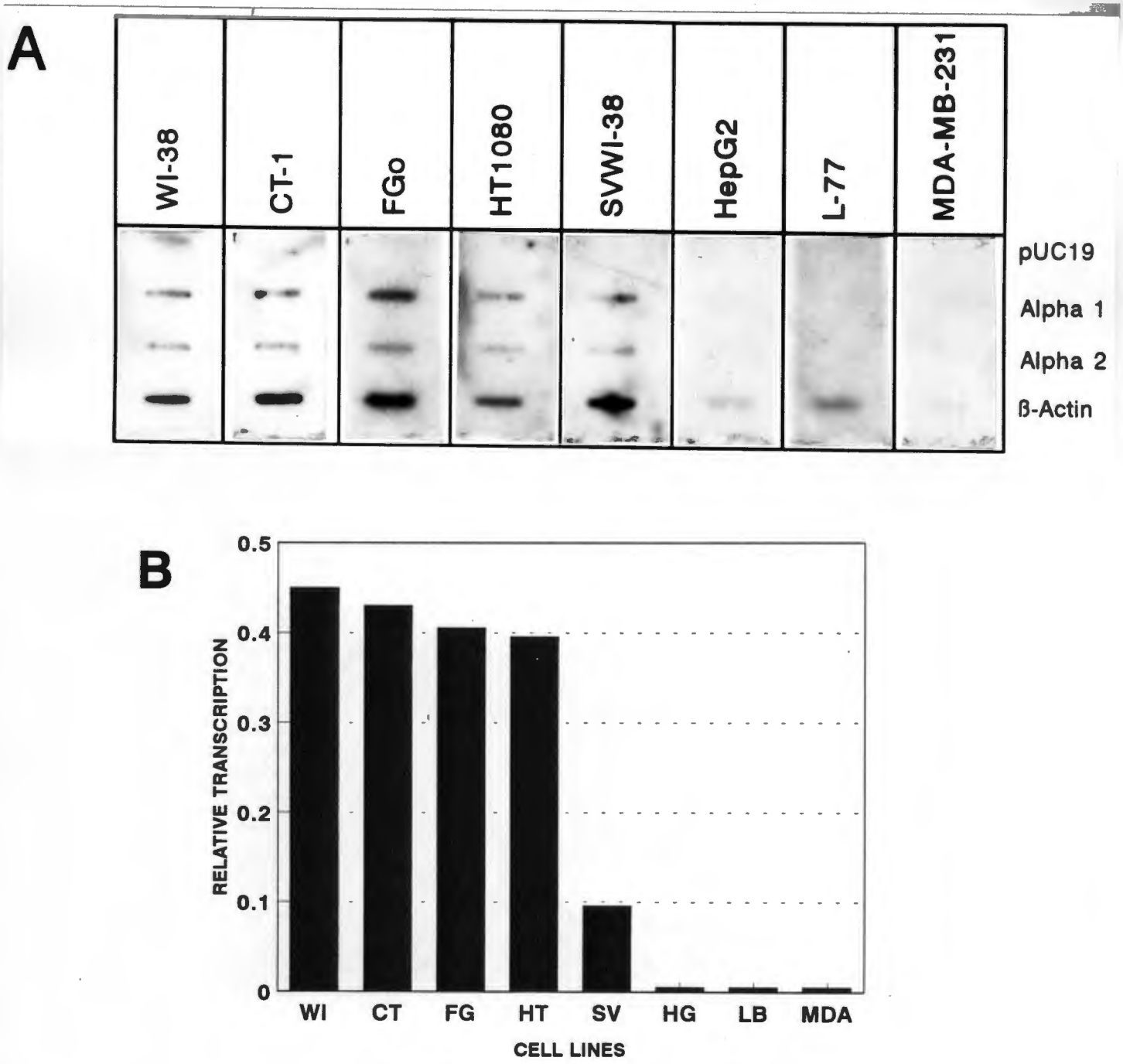


Fig 2.4 Nuclear run-on transcription assays of newly transcibed RNA from different cell lines. **(A)** ^{32}P- labelled nuclear run-on transcripts were prepared as described in section 6.6.2.2 and hybridised to cloned α1(1), α2(1) and β-actin DNA immobilised onto nitrocellulose membranes. Membranes were exposed to X-ray film for 16 hrs. The vector pUC-19 was used as a negative control and β-actin was used as an internal control. **(B)** Bar graph showing α2(1) procollagen trancripts expressed relative to the amount of β-actin transcripts. Although the SVWI-38 cell line, expresses α2(1) procollagen RNA, it does so at approximately 20% of the parental WI-38 cell line.

the gene. It was therefore decided to determine if the transcription factors which bind the proximal promoter region of the human α2(1) procollagen gene, do so in a cell-specific manner. The eventual aim being to correlate the data obtained thus far with transcription factor binding to the α2(1) procollagen promoter in the different cell lines.

2.2.3 Transcription factor binding to the human α2(1) procollagen promoter.

The results obtained in the nuclear run-on experiments indicated that the differences in collagen synthesis in these cell lines reside mainly at the level of transcription. Parker et al (1992), showed that the absence of α2(1) procollagen gene expression in SVWI-38 cells is associated with differential transcription factor binding to the -107 to -60 region of the human α2(1) procollagen promoter (fig 2.1). This region of the promoter contains the binding site for the CCAAT binding factor (G/CBE) and the recently identified adjacent site, the CME (collagen modulating element). The CME is described by Collins, (1993) and Collins et al, (1997). These studies showed that the sequences within this region of the α2(1) procollagen promoter are critical for the expression and regulation of the gene.

The G/CBE and CME are adjacent to each other (fig 2.5) and bind to distinct DNA-binding proteins. The G/CBE (GGAGG/CCAAT Binding Element) binds the ubiquitous CBF while the CME bind an as yet unidentified protein or proteins. These interactions appear to be responsible for differences in transcriptional activity of the α2(1) procollagen gene (Collins et al, 1997).

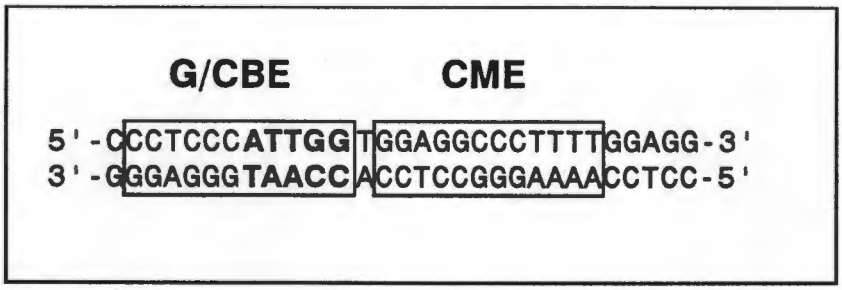


Fig 2.5 Transcription factor binding sites in the proximal promoter of the human α2(1) procollagen gene. The G/CBE contains an inverted CCAAT box and binds the complex I proteins. The adjacent element, the CME, binds both the complex II and III proteins (see fig 2.6).

It was therefore logical to determine the presence of these transcription factors in the range of cell lines used above, in an attempt to explain the observed differences in α2(1) collagen synthesis. Nuclear proteins were isolated from the different cell lines by the method of Dignam et al (1983). Binding of nuclear proteins to the -107 to -60 COL1A2 promoter fragment was analysed by electrophoretic mobility shift assays (EMSA) as described in section 6.4.3. Cells of fibroblast origin (WI-38, CT-1, FG_o, HT1080), where α2(1) collagen mRNA transcripts were detected, show predominantly two protein factors binding to this region of the promoter (complex I and complex III). In the cell lines where α2(1) collagen synthesis is not detected (SVWI-38, HepG2, L-77, MDA, ZR), a third DNA-protein complex, (complex II) is detected (fig 2.6.A). The complex I protein(s) are fairly ubiquitous, found in varying amounts in all the cell types and is thought to be a member of the CCAAT binding family of proteins (evidence for this is provided in chapter 3). In the L-77, MDA and ZR cell lines, complex II completely replaced complex III, whereas in the SVWI-38 and HepG2 cell lines, both complexes II and III were present.

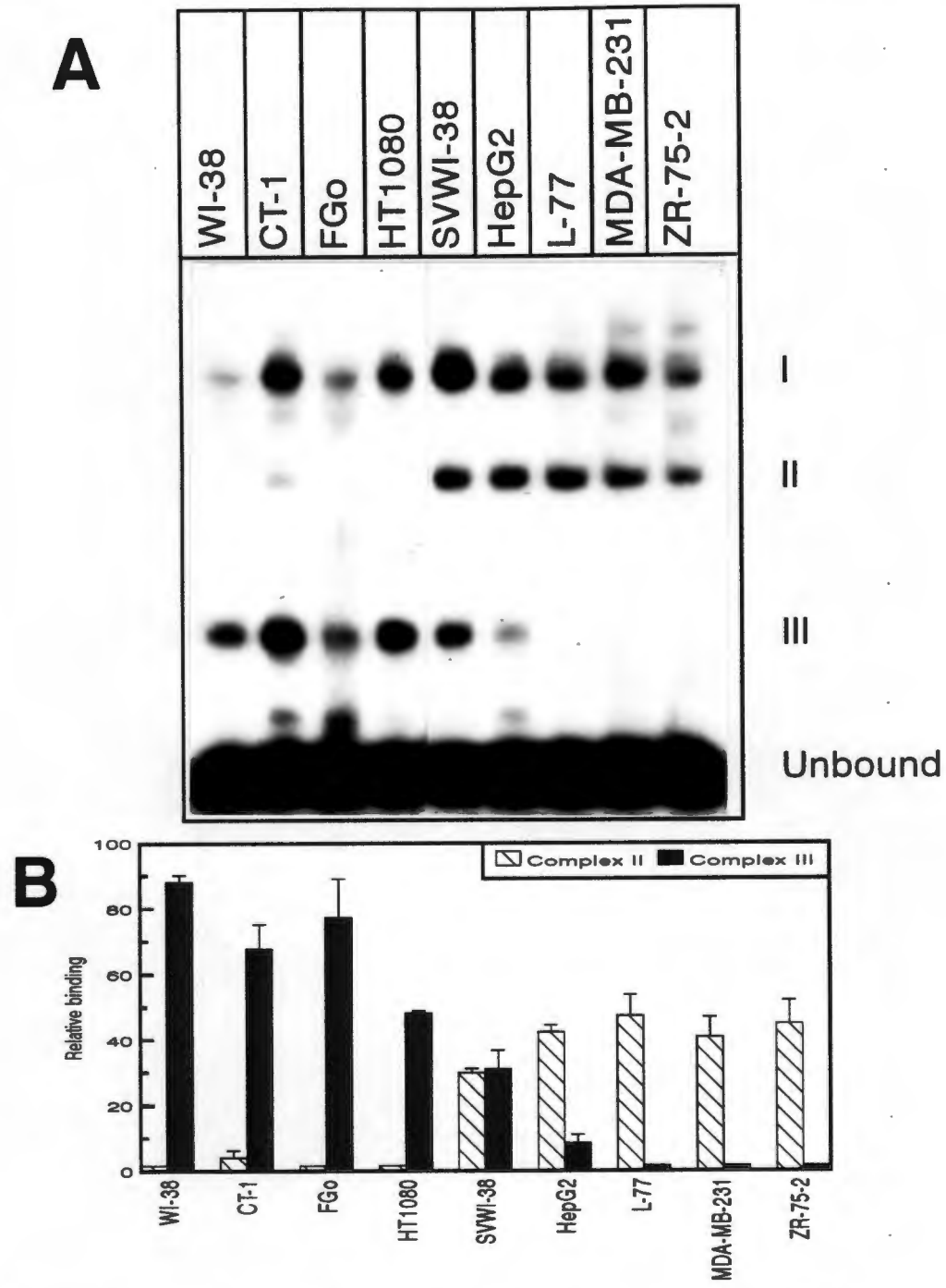


Fig 2.6 DNA-protein complex formation in differentiated cell lines. **(A)** 4μg of nuclear protein isolated from the indicated cell lines were incubated with ^{32}P labelled -107 to -60 α2(1) procollagen promoter fragment. 2μg poly dIdC.poly dIdC was included in the reaction. DNA-protein complexes were fractionated on 5% non-denaturing polyacrylamide gels, the gels dried and exposed to X-ray film for 16 hrs. The collagen producing cells showed predominantly complexes, I and III. Cells where collagen synthesis is not detected, contains an additional DNA-protein complex, II. **(B)** The intensity of the complex II and III bands was quantitated by densitometric scanning the autoradiographs of triplicate experiments and the relative intensity of these complexes in each cell line is shown in the bar chart. Cell lines are arranged such that the high collagen producers are on the left, progressing to the non-producers on the right. Results are the mean ± S.D. of four independent expreriments.

Based on these results it was proposed that regulation of human α2(1) collagen synthesis occurs via the following mechanisms:

1. In α2(1) collagen expressing cells: DNA-protein complexes I and III are involved in the co-activation of the COL1A2 gene.
2. In cell types not expressing α2(1) collagen: DNA-protein complex II is a potential repressor of COL1A2. The absence of complex III may also be significant in the repression of the gene.

The CT-1 cell line showed trace amounts of complex II binding to the α2(1) promoter which is in line with the observed 20% reduction in collagen synthesis in this cell line, when compared to that of the parental WI-38 cell line. The presence of trace amounts of complex II proteins in CT-1 cells may be responsible for the slight reduction in transcription observed. In the fibrosarcoma cell line (HT1080), where RNA transcripts were detected, both complex I and III proteins were present. This is consistent with other cells where type I collagen synthesis occurs and is in agreement with the hypothesis that complex I and III binding is associated with α2(1) collagen transcription.

In SVWI-38 cells, all three complexes, I,II and III were present. The presence of trace amounts of α2(1) procollagen RNA transcripts seen in nuclear run-on transcription assays in these cells may be as a result of the presence of complex I and some complex III proteins. There is therefore a direct relationship between α2(1) collagen expression and trans-acting factor interactions with the proximal promoter of the gene.

The hepatocellular carcinoma derived cell line, HepG2, contains all three complexes. Complex III, however was found in reduced amounts compared to the amount present in the collagen producing cells. Since we have proposed that complex III proteins may in part be involved in the activation of α2(1) procollagen gene expression, the reasons as to the presence of this complex in HepG2 cells still

remains unclear at the moment since this cell line contains no detectable α2(1) collagen transcripts. It is, however, possible that the abundant presence of complex II may mask any possible activation that complex III may have.

Quantitation of the complex II and III proteins in the various cell lines shows a direct relationship between the levels of these complexes and collagen synthesis (fig 2.6.B).

Direct Relationship:

Complex III proteins are present in cells where α2(1) collagen synthesis occurs, whereas the presence of complex II and the absence of complex III proteins are found predominantly in cells where no α2(1) collagen synthesis is detected.

This relationship holds true for all the cell lines used in this study, with the possible exception of SVWI-38 where the ratio for complex II:III proteins were 1:1. As explained in section 2.2.2.2, the α2(1) collagen gene was not completely silenced in this cell line, hence the detection of both DNA-protein complexes.

2.2.4 DNA binding site analysis

Methylation interference assays and mutation analysis showed that the different complexes binding the proximal promoter of the α2(1) procollagen gene, bind to distinct but adjacent elements (Parker et al, 1992 and Collins et al, 1997). EMSA's using an excess of cold oligonucleotides of the G/CBE and CME as competitors for protein binding were performed as described in section 6.4.3.1. SVWI-38 nuclear extracts were used as they contained all three DNA-protein complexes. The G/CBE oligonucleotide specifically competed out complex I formation (fig

2.7.A), while the CME oligonucleotide competed out both complex II and III formation (fig 2.7.B). These results showed that the complex II and III proteins bind the same DNA consensus sequence, (5'-GGAGGCCCTTTT-3'). Complex I formation is not affected by this oligonucleotide, showing distinct formation of complexes I, II and III. The binding of complex II and III proteins appeared to occur with similar affinity to the CME, since a densitometric scan of this competition shows that the competition for both complexes is virtually identical (fig 2.7.C).

2.2.5 Binding Kinetics of complexes I, II and III.

Since competition experiments indicated that complex II and III formation are similarly competed for by unlabelled CME oligonucleotide, an investigation into the binding kinetics of the proximal α2(1) procollagen promoter binding proteins was performed. Increasing amounts of radioactively labelled proximal α2(1) procollagen promoter fragment was incubated with a constant amount of nuclear extracts containing either partially purified complex I, II or III proteins. EMSA's were performed as described in section 6.4.3. The radioactivity in the DNA-protein complexes was determined using a phosphorimager. Binding curves of bound counts vs DNA concentration were generated and the dissociation constants determined. Fifty percent maximal binding was taken as the apparent K_d. Complex I bound the labelled probe with an apparent K_d of 2.9nM (fig 2.8.A), while the apparent K_d for complexes II and III were approximately 4.2 nM and 3.5 nM respectively (fig 2.8.B and C). These dissociation constants are within close range of each other, indicating that complexes II/III bind the CME with similar affinities, although complex III proteins has a slightly higher affinity for binding than complex II proteins.

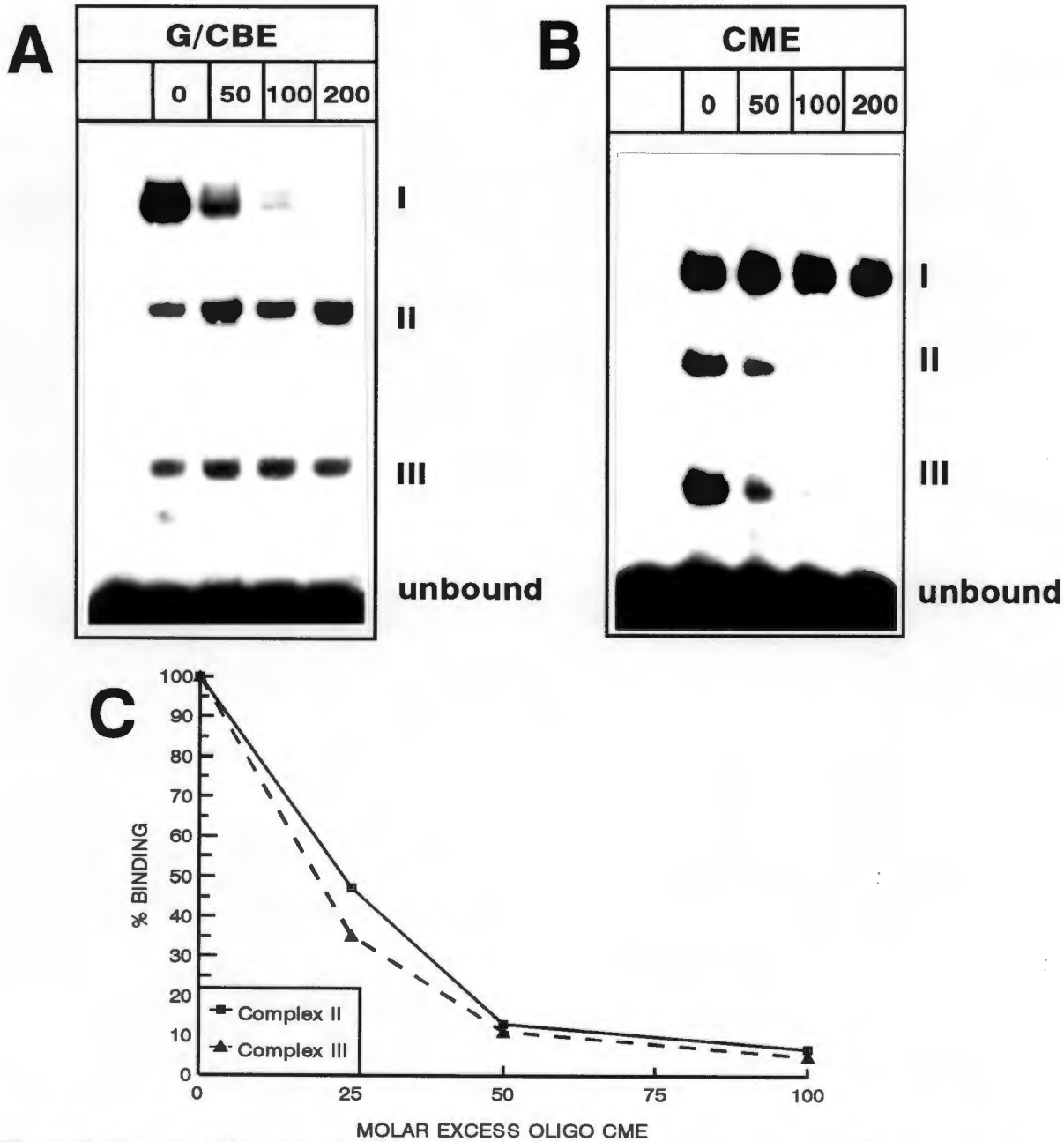


Fig 2.7 Competition of complexes I, II and III binding by molar excess of double stranded G/CBE and CME oligonucleotides. **(A)** An increasing molar excess of concatermerised double stranded G/CBE oligonucleotide was incubated with 4μg of SVWI-38 nuclear extract for 10 minutes prior to the addition of the -107 to -60 bp ^{32}P-labelled α2(1) procollagen promoter probe. DNA-protein complexes were analysed on non-denaturing 5% polyacrylamide gels, the gels dried and exposed to X-ray film for 16 hrs. **(B)** Increasing molar excess of double stranded CME oligonucleotide was incubated with SVWI-38 nuclear extracts as indicated in (A). **(C)** Autoradiographs were scanned and the percentage binding of complexes II and III with increase molar excess of oligonucleotide CME was plotted. Results are representative of three independent experiments.

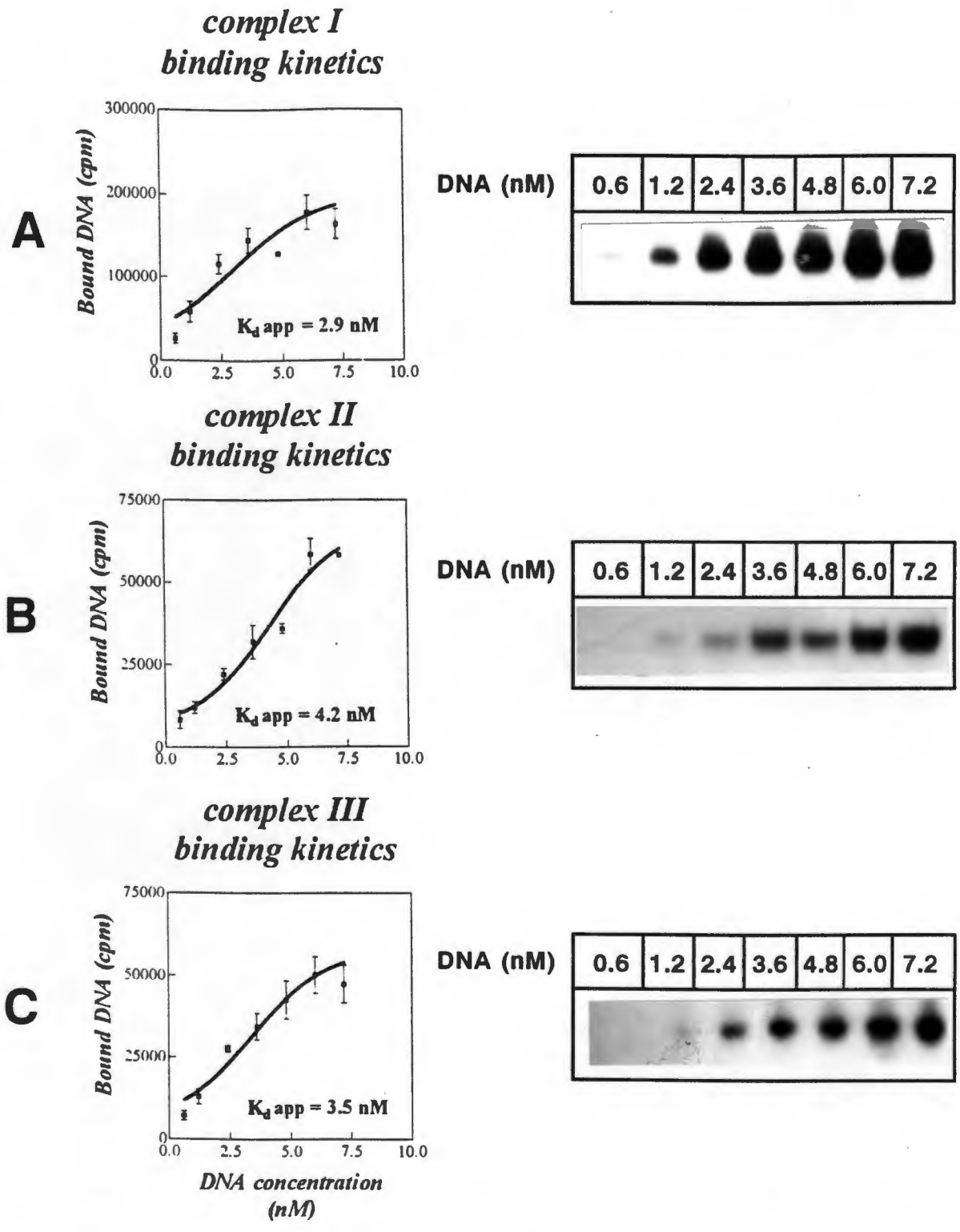


Fig 2.8 DNA binding affinity of complexes I, II and III. EMSA's were performed using increasing concentrations of labelled α2(1) procollagen proximal promoter fragment (-107 to -60) with 4 μg of partially purified complex I **(A)**, II **(B)** or III **(C)** proteins (section 6.4.3). The reaction mixtures were fractionated on nondenaturing 5% polyacrylamide gels. The dried gels were exposed to X-ray film for 16 hours. Bound counts were determined using a phosphorimager and plotted against the total DNA concentration. Fifty percent maximal binding was taken as the apparent K_d.

2.4 DISCUSSION

Steady-state levels of type I collagen mRNA is a major factor in the determination of type I collagen synthesis. In this study we showed a clear correlation between the ability of different cell lines to synthesise collagen and transcription factor binding to the proximal promoter of the human α2(1) procollagen gene. This correlation implies that cell-specific regulation of the human α2(1) collagen gene is mediated via the interaction of transcription factor binding to the proximal promoter. A way in which this regulation could be achieved, may be cell-type dependent and be due to trans-acting factor switching or the relative abundance of these factors in vivo. Interestingly, a recent study reported that nuclear proteins found in both collagen expressing and non-expressing cell types, exhibit cell-specific in vivo DNA-protein interactions at the mouse proximal α1(1) and α2(1) procollagen promoter (Chen et al, 1997).

The DNA elements which play a role in tran-acting factor interactions at the human proximal α2(1) procollagen gene have been characterised by Parker et al, (1989) and Collins et al, (1997) and are located at positon -92 to -67. The upstream element, the G/CBE, contains an inverted CCAAT box and forms a single DNA-protein complex, complex I. This element is identical to the CBF binding site in the mouse Col1A2 promoter (Hatamochi et al, 1988) and it is therefore likely that the complex I proteins are related to or identical to the mouse CBF. The second element, the collagen modulating element (CME), forms two different DNA-protein complexes, II and III.

This investigation showed that complexes II and III were associated with differences in α2(1) collagen gene transcriptional activity in a variety of cell lines. All the cell lines assayed contained equivalent amounts of complex I proteins, whereas the amount of complex II and III proteins varied from cell line

to cell line. It is proposed that complexes I and III are present in cells where the α2(1) procollagen gene is expressed and may be partly responsible for the activation of the α2(1) procollagen gene (fig 2.9.A).

The complex I proteins appear to be fairly ubiquitous, possibly a CCAAT-binding factor, while complex II and III proteins are as yet uncharacterised. Complex I and III proteins could co-operate with each other and with the transcription initiation machinery to activate the COL1A2 gene in a cell-specific manner. The cell-specific expression of a number of genes has been attributed to co-operative interactions of transcription factors. Tronche, (1989) and Wuarin et al, (1990), showed that cell-specific expression of the albumin gene is associated with the co-operative interaction of the CCAAT-binding factor (NF-Y) and a cell-specific factor, HNF-1 to adjacent elements within the albumin promoter. Hernandez-Munain and Krangel, (1994) showed that intact and appropriately positioned (adjacent) binding sites for c-Myb and the T-cell-specific CBF are both necessary and sufficient for enhancer activity in the T-cell receptor (TCR) gene. These two factors co-operate functionally to mediate the T-cell- specific expression of the TCRδ gene. Although the HT1080 cell line used in our study contained no detectable type I collagen protein or steady-state mRNA, newly formed RNA transcripts were present as shown by nuclear run-on experiments. This cell line contained both the complex I and III proteins in EMSA's, which is in agreement with the hypothesis that complex I and III proteins are involved in the activation of the COL1A2 gene. It can therefore be predicted that the -107 to +54 COL1A2 promoter-CAT construct should have activity in HT1080 cells. Preliminary results indicated that the construct had activity in these cells, albeit at levels lower than that detected in CT-1 cells but higher than that in non-collagen producing cell lines such as SVWI-38, HepG2, MDA-MB-231 and ZR-75-2 (data not shown). Similar findings have been reported in HeLa cells by

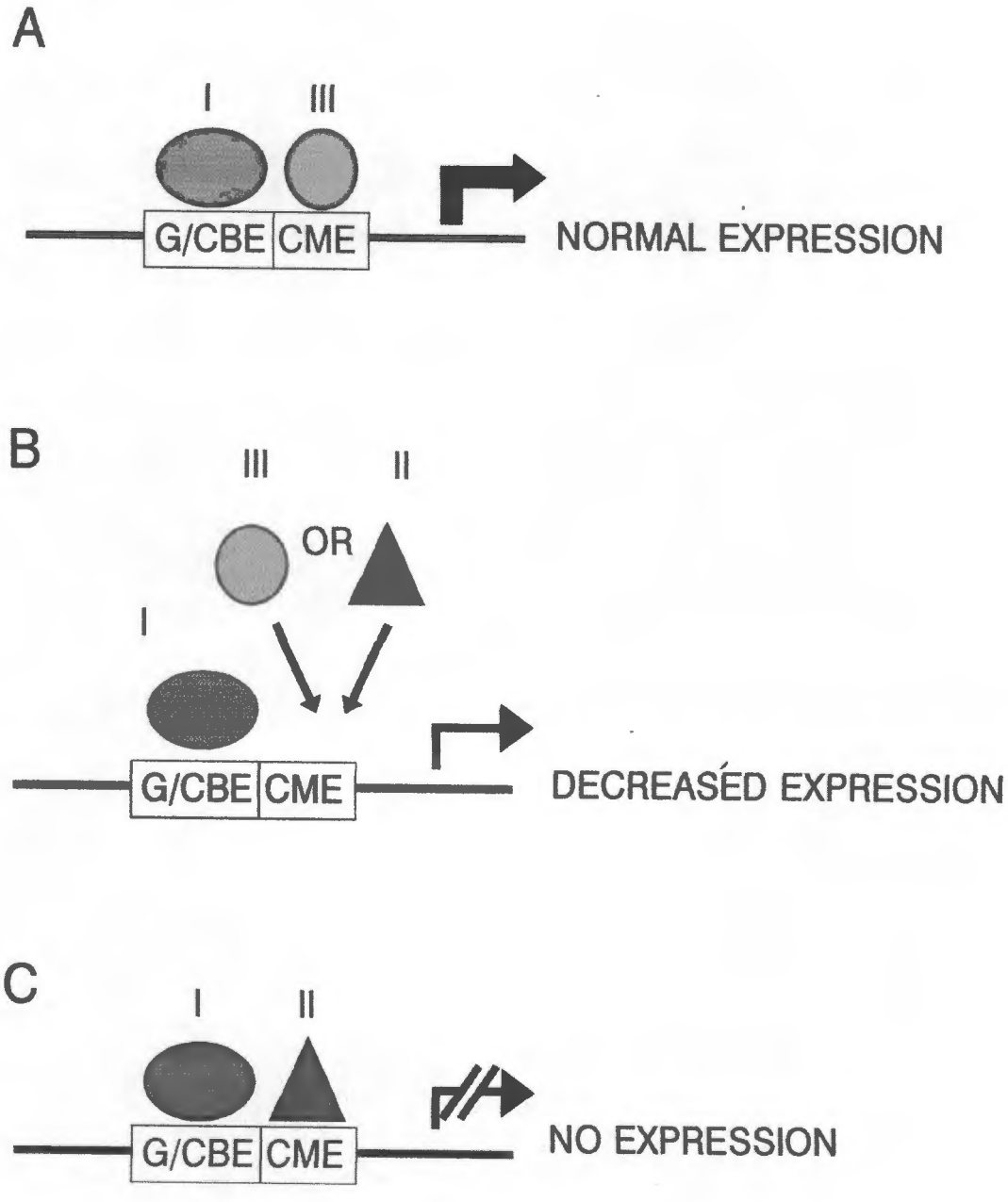


Fig 2.9 Model for the regulation of the human α2(1) procollagen gene. **(A)** In collagen producing cells, the complex I and III proteins interact with their binding sites, the G/CBE and CME respectively to activate gene transcription. **(B)** In some cases, the relative levels of complex II and III proteins may serve to modulate the extent to which the gene may be active. **(C)** In cells where collagen synthesis is not detected, complex III proteins may be completely replaced by complex II proteins, resulting in gene inactivation.

Furth et al, (1991) where the type I collagen genes are transcribed, but a post-initiation block prevents the formation of mature type I collagen mRNA.

In cells where very little or no α2(1) collagen is expressed, the presence of a different DNA-protein complex, complex II is detected. This complex is not detected in cells where the gene is active, and it is on the basis of this finding that it has been proposed that the complex II proteins may be involved in the cell-specific repression of the COL1A2 gene (fig 2.9.C). The complex II proteins could bind the CME to inactivate the expression of the COL1A2 gene in one of the following ways; (a) complex II proteins interact with the G/CBE binding proteins and in this way blocks activation of the gene, (b) The complex II proteins may be able to inactivate the gene independently, without any interaction with the G/CBE binding proteins. A repressor of the CCAAT binding factor, the CCAAT-displacement protein (CDP) has been shown to displace the CCAAT binding protein from the sea urchin histone H2B-1 promoter (Barberis et al, 1987). Similar findings have been reported for the repressor activity of CDP in HL-60 myeloid cells (Lievens et al, 1995). Walsh and Schimmel (1987) showed that myoblast and myotube nuclear extracts form a DNA-protein complex on the α-actin promoter, which is specific for the myotube cell-type, while a different DNA-protein complex is detected in non-myocyte cell types.

Although the SVWI-38 and HepG2 cell lines contained no α2(1) collagen, both cell lines had detectable levels of both complex II and III proteins. A study by Inagaki et al, (1987) also reported that type I collagen mRNA is not detected in HepG2 cells. Nuclear run-on transcription assays showed that SVWI-38 produced trace amounts of the COL1A2 transcripts which suggests that the relative ratios of complex II and III proteins may be important in the level of COL1A2 expression. The CT-1 cell line, which produces 80% of the type I collagen produced by its parent cell line, WI-38, also showed trace amounts of the complex II proteins. The

presence of trace amounts of complex II could account for the partial inactivation of the gene in CT-1 cells. It would therefore appear that a given cell may have both complex II and III proteins, but that the relative abundance of the different complexes will determine the transcriptional activity of the gene (fig 2.9.B) A similar situation exist in the control of the human α-globin gene where factors which may be functionally limiting in nuclear extracts have been shown to play a role in the regulation of the gene in that two adjacent sites (α-IRP and α-CP1) bind constitutively expressed transcription factors, while another factor (α-CP2), binding upstream of these sites is present in functionally limiting amounts (Barnhart et al, 1988; Kim et al, 1990) .

Our results show that the complex II and III proteins bind the same element (CME) in the COL1A2 promoter. The relationship between these proteins is unclear at the moment, but the possibility that complex III may be a degradation product of complex II has been excluded. It is possible that they are two distinct transcription factors with common DNA-binding domains or they may be products of the same gene arising by alternate splicing and each having a specific function. Examples of the latter are described by Foulkes and Sassone-Corsi, (1992).

At the onset of this study, our aim was to determine whether a correlation between cell-specific α2(1) collagen expression and trans-acting factor binding to the promoter of this gene existed. The results obtained in this study clearly showed that there is a direct relationship between expression of the gene and trans-acting factor binding to the -107 to -60 region of the promoter. This relationship appears to occur in a cell-specific manner, where the relative abundance of the trans-acting factors may be important in the regulation of the gene. The trans-acting factors described here may play a role in cell-specific expression of α2(1) collagen in combination with other transcription factors, such as the transcription initiation machinery and factors which interact upstream of the CCAAT box.

3. SPECIES-SPECIFIC EXPRESSION OF THE α2(1) PROCOLLAGEN GENE

CHAPTER 3:

SPECIES-SPECIFIC EXPRESSION OF THE α2(1) PROCOLLAGEN GENE.

3.1 INTRODUCTION

The transcriptional mechanisms involved in the cell and tissue specific regulation of the mouse α2(1) procollagen gene are well documented. Chapter 2 reported that the cell-specific regulation of the human COL1A2 gene may in part be controlled by trans-acting factor switching. These factors bind to adjacent elements in the proximal promoter of the gene. Collins et al, (1997) defined the exact sequences required for binding of these factors, where at least one of the three protein factors that bind the promoter is a member of the CCAAT-binding family. This conclusion is drawn from the observation that this factor binds the inverted CCAAT box in the COL1A2 promoter.

The CCAAT-binding proteins consist of a family of proteins (section 1.2.3.1.1) and methylation interference analysis showed that the DNA sequence interacting with the human CCAAT-binding protein includes an upstream GGAGG box as well as the inverted CCAAT sequence (G/CBE) (Parker et al, 1992). It would appear that the protein factor/s which interact with the human G/CBE and the mouse CBF are similar. Experiments aimed at determining the interaction of mouse nuclear proteins with the human -107 to -60 α2(1) collagen promoter

region led to an investigation of the species-specific regulation of the gene. Species specificity of the osteocalcin gene has been shown by Heinrichs et al, (1993) to involve trans-acting factor interactions with two proximal elements in the promoter of the gene. One of the elements involved in this species-specific regulation is the CCAAT sequence. These studies suggest that a single base mismatch between the human and rat osteocalcin gene around the CCAAT box could be responsible for species-specific regulation of the gene.

As more data regarding the TGF-β response in type I collagen gene expression is being generated, it is becoming clear that the mechanism by which this response is mediated differs between the human and mouse. While Rossi et al, (1988) proposed NF-1 as the factor involved in the TGF-β response in the mouse promoter, Inagaki et al, (1994) proposed that Sp1 may be the factor involved in the human promoter. Contrary to this, an additional study suggested AP-1 as the transcription factor responsible for the TGF-β response in the human promoter (Chung et al, 1996). This has subsequently been shown not to be the case by Greenwel et al, (1997), who provide data in support of their original claim that Sp1 is the TGF-β responsive factor. There is, however, general agreement that the TGF-β response is mediated in a species-specific manner, possibly due to sequence divergence between the human and mouse promoters, allowing the interaction of different trans-acting factors. Collectively, these results and those presented in this chapter suggest that subtle differences at a nucleotide level between the human and rodent α2(1) procollagen promoters may in part be responsible for species-specific regulation of the gene.

Protein phosphorylation is one of the most versatile post-translational modification events widely used by cells in the control of biological processes. Signal tranduction pathways are mediated via the activation of a variety of kinases including, MAP kinases and protein kinase C, which can translocate to the nucleus

where they phosphorylate transcription regulators in a cell cycle-dependent manner or in response to various stimuli. A number of transcription factors such as p53, c-Myc, Oct-1 and SV-40 T-antigen that stimulate DNA replication, are known to be controlled by phosphorylation which alters their DNA-binding activities and interactions with other transcription factors (reviewed by Boulikas, 1995). Protein kinase C-dependent mechanisms have been implicated in the regulation of collagen gene expression (Harrison et al, 1990; Stuiver et al, 1991; Greenwel et al, 1997). This kind of regulation can be mediated via transcription factors such as AP-1, known to be under the control of phosphorylation events (Hunter and Karin, 1995) and recently proposed to have a possible regulatory role in α2(1) procollagen gene expression (Chung et al, 1996). Tyrosine deposhorylation of nuclear proteins (possibly that of Sp1) has also been reported to mimic TGF-β responsiveness of the human α2(1) procollagen gene (Greenwel et al, 1995). Collectively these findings prompted a study aimed at determining the effect of phosphorylation/dephosphorylation on transcription factor binding to the proximal α2(1) procollagen promoter and the ultimate expression of the gene.

This chapter describes experiments which allowed the determination of species-specific mechanisms and phosphorylation events in the expression of the human α2(1) procollagen gene.

3.2 RESULTS

3.2.1 Binding of rodent nuclear proteins to the human α2(1) procollagen promoter.

Nuclear extracts from rodent cells with differing collagen synthetic activities were prepared by the method of Dignam et al, (1983). The 3T3 and 3T6 fibroblasts are rodent cells where α2(1) collagen is produced. An oncogenic ras transformed rat cell line, Rasrat-1 had been shown by Slack et al, (1992) to express very little type 1 collagen and a mouse macrophage cell line ($p388D_1$) where no α2(1) collagen was detected are representative of cells where the gene is inhibited. Nuclear proteins from these cell lines were incubated with the -107 to -60 fragment of the human promoter (fig 3.1). Surprisingly, all four cell lines contained DNA-protein complexes with migration patterns like that of complex I and II proteins. This is in contrast to the results obtained with human nuclear extracts, where a difference in trans-acting factor interactions between collagen expressing and non-expressing cells were observed (section 2.2.3). The proposal that complex II proteins are associated with the repression of the human α2(1) procollagen gene does not appear to hold true when using rodent nuclear extracts. This observation led to an investigation regarding the interaction of trans-acting factors with both the human and rodent α2(1) procollagen proximal promoters.

3.2.2 Analysis of cross-species promoter-binding activities

To determine whether the DNA-protein complexes on the rodent promoter occured on the same sequences as the human gene, the mouse α2(1) collagen promoter containing plasmid, pR40, (Rossi and de Crombrugghe, 1987) was digested with the restriction enzymes HindIII and Sma1 to release a 160 bp fragment (fig 3.2.A). This 160 bp HindIII/Sma1 fragment was cloned in the vector pUC19 (fig 3.2.B and C) and the sequence of the cloned fragment confirmed (fig 3.3.A).

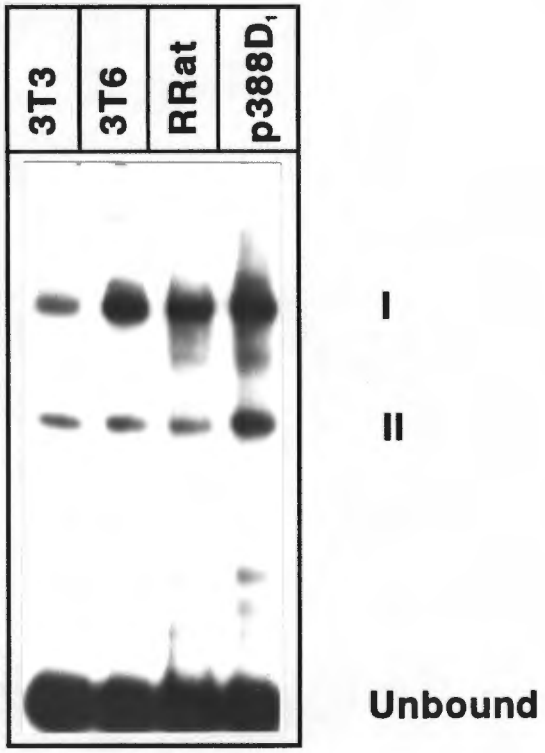


Fig 3.1 Interactions of mouse nuclear proteins with the human proximal α2(1) procollagen promoter (-107 to -60). Nuclear proteins from collagen expressing rodent cells, 3T3, 3T6 and rodent cells not expressing α2(1) collagen, Rasrat-1 (RRat) and mouse macrophages, p388D$_1$ were incubated with 1 ng of labelled human -107 to -60 α2(1) collagen gene fragment. The DNA-protein complexes were electrophoresed on non-denaturing 5% polyacrylamide gels at 4°C in 0.5X TBE and 150 V for 2-3 hours. The dried gel was exposed to X-ray film for 16 hours. The positions of complexes I and II are indicated.

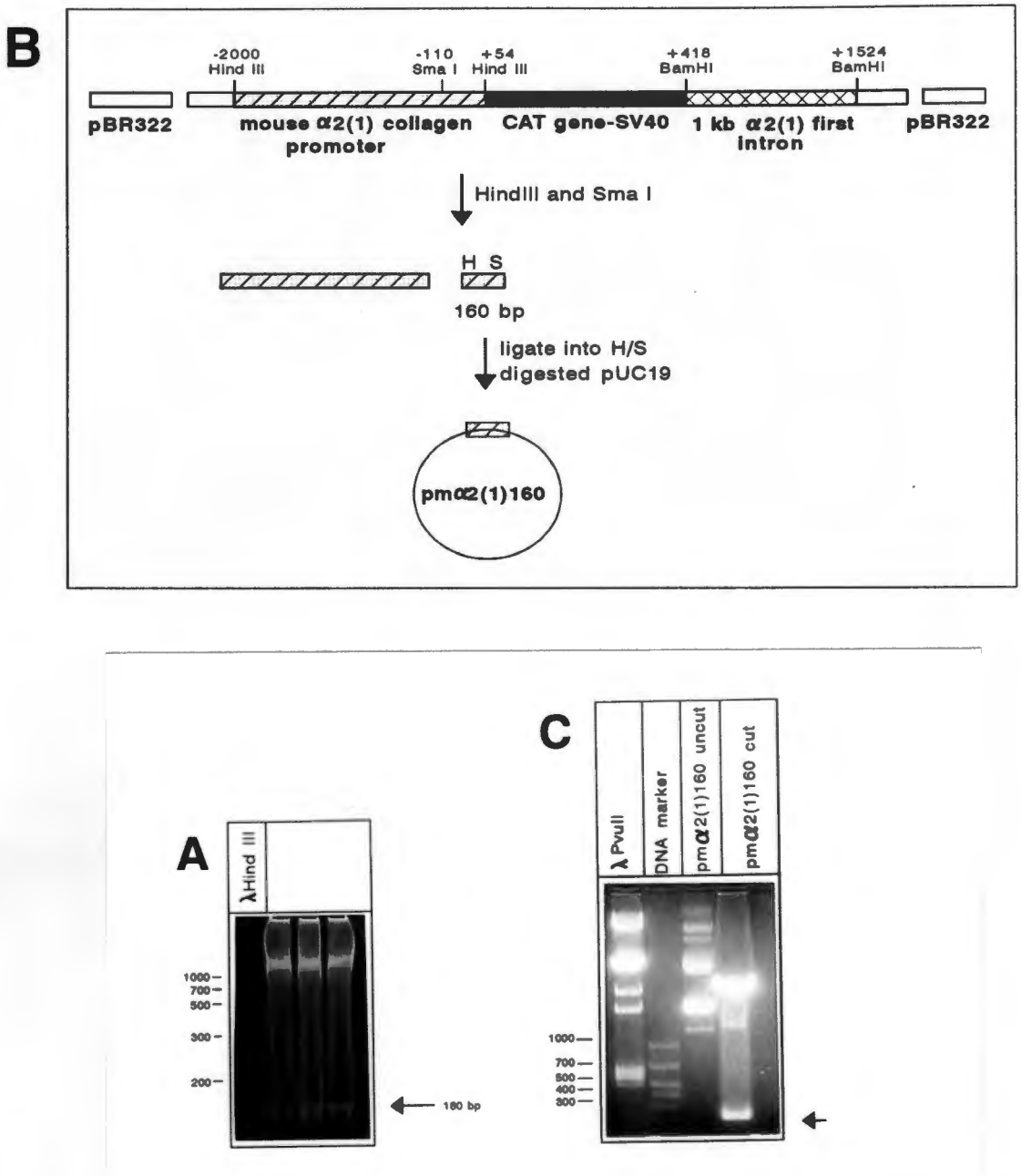


Fig 3.2 Cloning of the mouse -107 to +54 α2(1) procollagen promoter fragment. **(A)** The 160 bp fragment spanning -107 to +54 of the mouse α2(1) collagen promoter was released from the plasmid pR40 (Rossi and de Crombrugghe, 1987). The digested DNA was electrophoresed on 7% polyacrylamide gels and the 160 bp fragment eluted from the gel slices as described in section 6.3.2.1. **(B)** The released HindIII/Sma1 fragment was cloned into HindIII/Sma1 digested pUC19. A diagramatic representation of the cloning events is depicted and the resultant plasmid pmα2(1)160 is shown. **(C)** After cloning of the mouse promoter fragment, enzymatic digestion of pmα2(1)160 was used to confirm the presence of the cloned fragment. HindIII and EcoR1 double digest released the correct sized fragment as seen on 1% agarose gels.

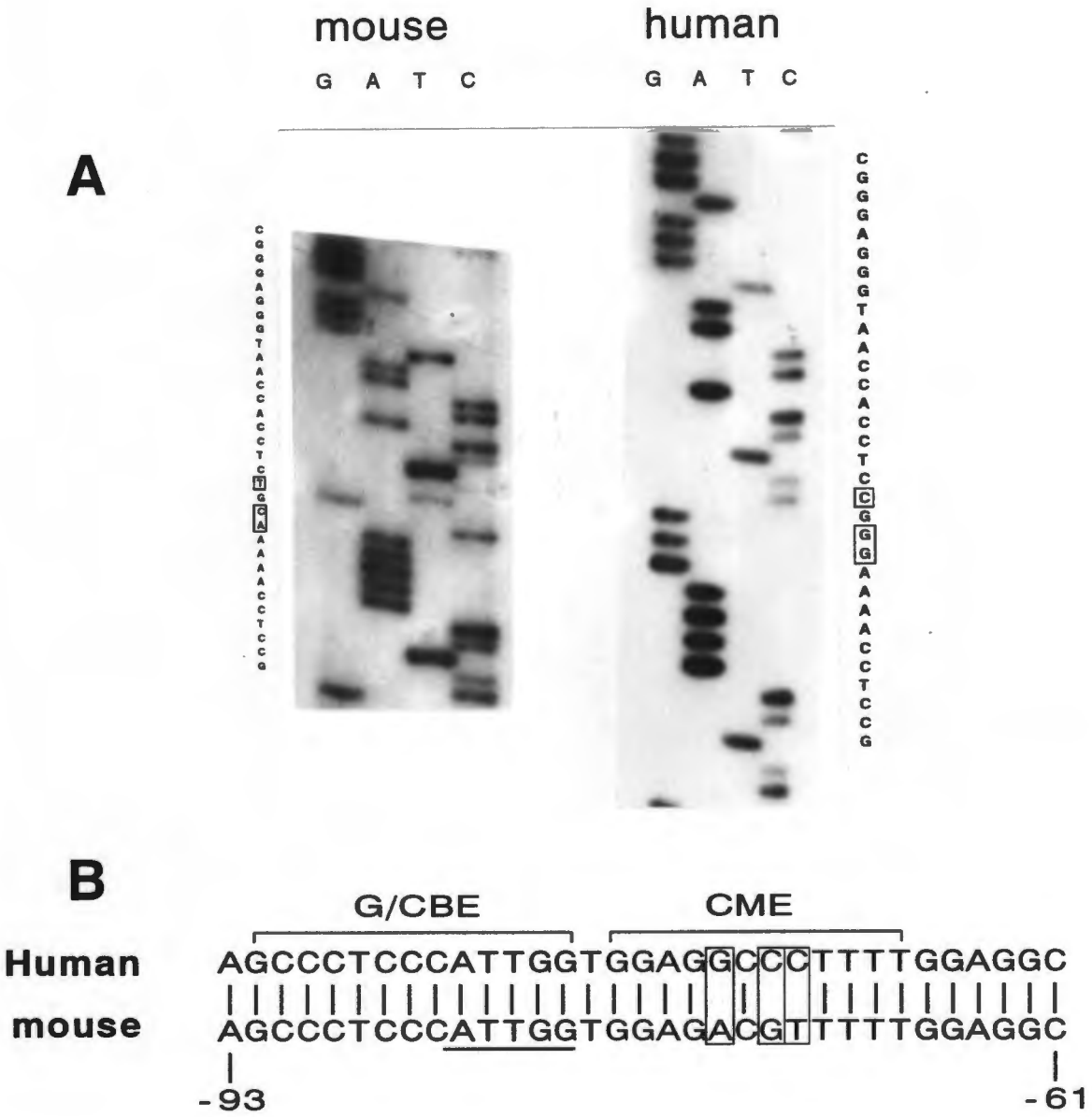


Fig 3.3 Comparison of the human and mouse α2(1) procollagen proximal promoter sequences. (A) DNA sequence analysis (section 6.8) was performed using the sequenase protocol (Amersham) to confirm the presence of the 3 bp difference between the human and rodent α2(1) procollagen promoters in the region immediately downstream of the CCAAT box. **(B)** Alignment of the -90 to -60 region of the human and mouse α2(1) procollagen promoters, indicating the 3 bp difference between the two promoters. The G/CBE and CME sequences are indicated.

Comparison of the equivalent region of the human α2(1) procollagen promoter revealed a three base pair mismatch between the human and mouse promoters in the region corresponding to the CME. The complex I binding site in the human and rodent α2(1) procollagen promoters, however, was conserved (fig 3.3B). Since the mismatch occured at the core of the CME, it is conceivable that complex II/III formation may not occur on the mouse α2(1) procollagen promoter. To test this theory, the rodent -107 to +54 fragment was used in EMSA's using nuclear extracts from collagen expressing and non-expressing rodent cells (fig 3.4). In all the cell lines tested, only one protein complex, the CCAAT-binding factor (CBF) was detected on the rodent promoter. This result together with the data in fig 3.1, indicated that although the complex II proteins were present in rodent nuclear extracts, it is unable to bind the Col1A2 promoter.

In order to assess the binding of human factors to the mouse promoter, nuclear extracts from representative α2(1) collagen expressing and non-expressing human and rodent cells were used in EMSA's to detect DNA binding activities of trans-acting factors to equivalent regions of the human and rodent promoters. CT-1 and 3T3 nuclear extracts were used as representatives of cells which express the human and rodent α2(1) procollagen genes respectively, while SVWI-38 and mouse macrophages ($p388D_1$), represent cells where the gene is not expressed. EMSA's were performed using either the human or mouse α2(1) procollagen proximal promoters as a probe. Human nuclear extracts from α2(1) collagen expressing cells, CT-1, showed binding of complex I and III proteins to the human probe. Extracts from SVWI-38 cells where α2(1) synthesis is inhibited contained the additional protein complex, complex II. Nuclear extracts from rodent cells 3T3 and $p388D_1$, contained both complex I and II proteins which bound the human probe (fig 3.5.A).

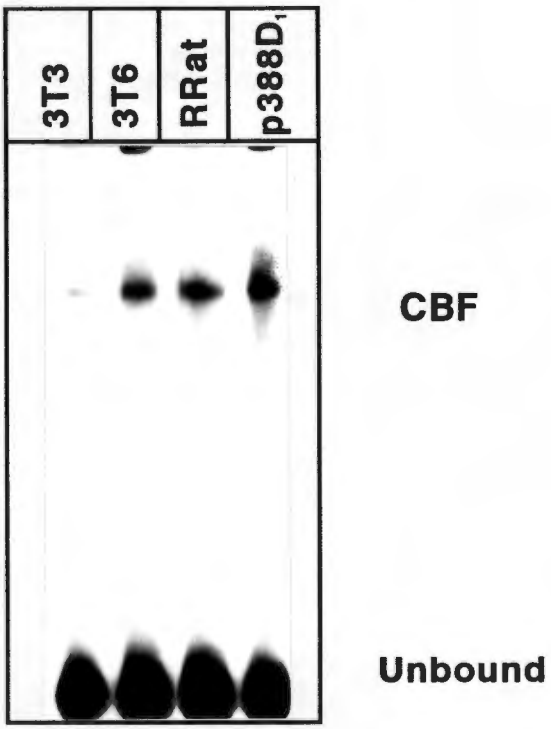


Fig 3.4 DNA-protein complex formation between rodent nuclear extracts and the mouse α2(1) procollagen proximal promoter. Nuclear extracts from α2(1) collagen expressing cells, 3T3 and 3T6 as well as that from non-expressing cells, rasrat-1 (RRat) and p388D$_1$ were incubated with 1 ng of labelled mouse promoter fragment (as described in fig 3.1). DNA-protein complexes were electrophoresed on non-denaturing 5% polyacrylamide gels as described in section 6.4.3. Dried gels were exposed to X-ray film for 16 hours. The position of the mouse CCAAT binding factor, CBF is indicated.

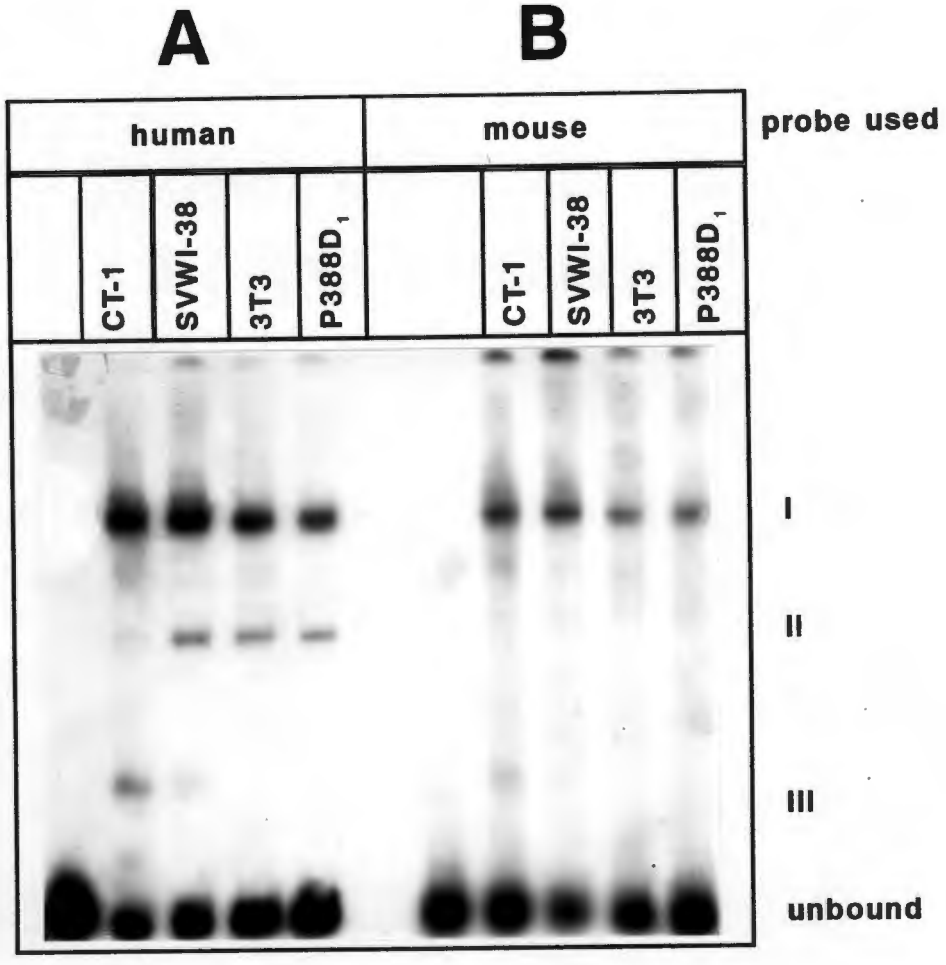


Fig 3.5 Cross-species DNA-protein interactions on the human and mouse α2(1) procollagen promoters. **(A)** Nuclear extracts from representative human and rodent collagen expressing and non-expressing cells were isolated as described in section 6.4.1. CT-1 and 3T3 are representative of human and rodent collagen expressing cells respectively, while SVWI-38 and $p388D_1$ represent the non-expressing cells. 4μg of nuclear proteins were incubated with the human -107 to +54 promoter fragment. DNA-protein complexes were separated on non-denaturing 5% polyacrylamide gels and dried gels were exposed to X-ray film for 16 hours. The positions of complexes I, II and III are indicated. **(B)** The same of amount of nuclear proteins described in (A) were incubated with the mouse -107 to +54 promoter fragment. EMSA's were performed as described in (A). A single DNA-protein complex with a similar migration pattern to that of complex I was observed.

When the same nuclear extracts were incubated with the equivalent region of the mouse α2(1) procollagen promoter, complex I was the only complex observed in all the cell lines (fig 3.5.B). These results are in agreement with those shown in fig 3.4 where the human complex II and III proteins were not detected on the rodent -107 to +54 promoter fragment. The 3 bp mismatch is therefore sufficient to abolish the formation of complexes II and III with the mouse promoter although complex II proteins are present in these cells.

Since the binding of complex II and III proteins appear to be affected by the 3 bp mismatch between the human and rodent α2(1) procollagen promoters in the CME region of the gene, it is possible that this mismatch may confer species-specific regulation of the gene. In order to determine whether this difference would be significant for species-specific regulation, this region of both human and mouse promoters were cloned into a CAT expression vector for measurement of promoter activities.

3.2.3 Human and Rodent promoter activities

The -107 to +54 α2(1) procollagen promoter-CAT constructs were co-transfected with CMVβgal as a control for transfection efficiency into CT-1 and 3T3 cells. The mouse ColCAT construct had a significantly lower activity than the human COLCAT construct in the CT-1 cell line (fig 3.6.A). This was a significant difference with $p<0.01$, Mann-Whitney *U* test. Similar results were obtained in four independent experiments. The difference between the human and mouse promoter activities in CT-1 cells varied from between 4 to 7 fold in different experiments. The only detectable difference between these two promoters in terms of trans-acting factor binding sites (complexes I, II and III) is the 3 bp mismatch in the CME region. Although other differences between the human and mouse proximal promoters further downstream of -60 do exist, these appear not to

involve DNA-protein interactions (except for the basal transcription factors which we do not detect in our assay system). It would therefore appear that the 3 bp mismatch in the CME is sufficient to confer not only species-specific binding of trans-acting factors, but also a species-specific difference in the activities of the α2(1) procollagen promoters. Interestingly, transfection of both promoter-CAT constructs into the mouse 3T3 cell line, showed no significant difference in activity between the two promoters. This may be an indication of basal expression of the gene in this cell line. Also, the activity of the human promoter-CAT construct is significantly lower in 3T3 cells than in CT-1 cells (fig 3.6.B). This result is consistent with the finding of Parker et al, (1992) that the human promoter-CAT construct has decreased activity in SVWI-38 cells which contain complex II proteins. It also supports the hypothesis that complex I, together with complex III enables a stronger activity. The presence of complex II and the absence of complex III proteins in 3T3 cells could therefore result in the decreased activity of the human promoter-CAT construct in these cells. This is in agreement with the model that complex II proteins may be involved in the inhibition of α2(1) procollagen gene expression and that complex III proteins may, on the other hand, be involved in co-activation of the α2(1) procollagen gene.

3.2.4 Analysis of the human CCAAT-binding factor

3.2.4.1 Competition experiments using known CCAAT binding sequences

Oligonucleotides containing known CCAAT binding motifs for different members of the CCAAT-binding protein family were used in EMSA's to determine the nature of the complex I proteins and their relation to the mouse CBF. The competitor oligonucleotides used in EMSA's with CT-1 nuclear extracts are shown in fig 3.7.A. Complex I formation was effectively competed out by the wildtype G/CBE, mouse CCAAT-binding element (CBE) and the NF-Y oligonucleotide while complex III formation was not affected (fig 3.7.B).

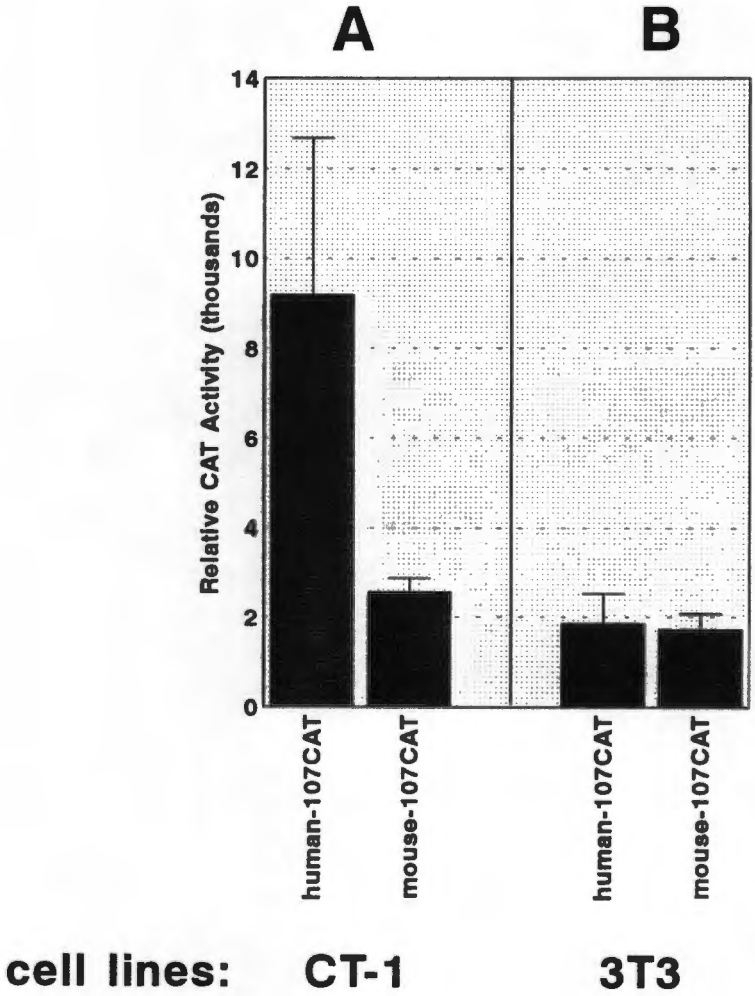


Fig 3.6 Transcriptional activities of the human and mouse proximal α2(1) procollagen promoters. 10 μg of the indicated promoter constructs in the p8CAT vector was transfected into CT-1(A) and 3T3 (B) cells using the $CaPO_4$ precipitation method as described in section 6.9. The plasmid CMVβgal was included in the transfections as an internal control to measure DNA uptake into the cells. CAT activity relative to β-galactosidase activity is indicated. Results are the mean ± S.D. of 3 experiments.

Complex I formation was however not competed out by the CCAAT/enhancer binding consensus sequence (C/EBP) (fig 3.7.B). Similar results were obtained with SVWI-38 nuclear extracts, showing that complex I formation was competed out by the G/CBE, mouse CBE and to a marginally lesser extent by the NF-Y binding sequence The C/EBP oligonucleotide had no effect on complex I formation as previously observed (fig 3.8). Competition for complex I formation by these oligonucleotides resulted in an increase in the intensity of complex II binding. The reason for this is unclear, but it can be speculated that the loss of complex I binding to the labelled DNA fragment may decrease steric hindrance, allowing more effective binding of complex II. This may also be an *in vitro* artifact and *in vivo* binding may be of a different nature. The sequences immediately upstream of the G/CBE and the mouse CBE are conserved between the human and mouse α2(1) procollagen promoters. There are, however, differences between the nucleotides flanking the COL1A2 CCAAT box and the sequences flanking the CCAAT-binding sequence of NF-Y (fig 3.9). Since it has previously been shown that sequences upstream of the CCAAT box in the G/CBE are essential for complex I formation (Collins, 1993, Collins et al, 1997), this could explain why the NF-Y binding sequence was a slightly weaker competitor.

CBF binding on the mouse -107 to +54 promoter was also competed out by the human G/CBE in a manner similar to that observed by competition for complex I binding by the mouse CBF sequence. The human CCAAT-binding sequence, however was not as effective a competitor for CBF binding as was the wild type mouse CCAAT sequence (fig 3.10). These results indicate that the interaction of rodent and human CCAAT-binding proteins to each of the human and mouse promoters are similar, but not identical.

3.2.4.2 Supershift assays using mouse anti-CBF antibodies

Antibodies against the mouse CCAAT binding protein were used in EMSA's in order to gain more insight into the nature of the human complex I proteins. The mouse anti-CBF antibodies used were directed against the A and B-subunits of CBF (a gift from H. Eberspaecher and B de Crombrugghe, University of Texas).

A

NF-Y	5'-AACATTTTTCTGATTGGTTAAAAGTTG-3' 3'-TTGTAAAAAGACTAACCAATTTTCAAC-5'
C/EBP	5'-AATTCAATTGGGCAATCAGG-3' 3'-GTTAACCCGTTAGTCCTTAA-5'
G/CBE	5'-GATCCGCCCTCCCATTGGTG-3' 3'-GCGGGAGGGTAACCACCTAG-5'
CBE	-110 TO +54, murine Col1A2 gene

B

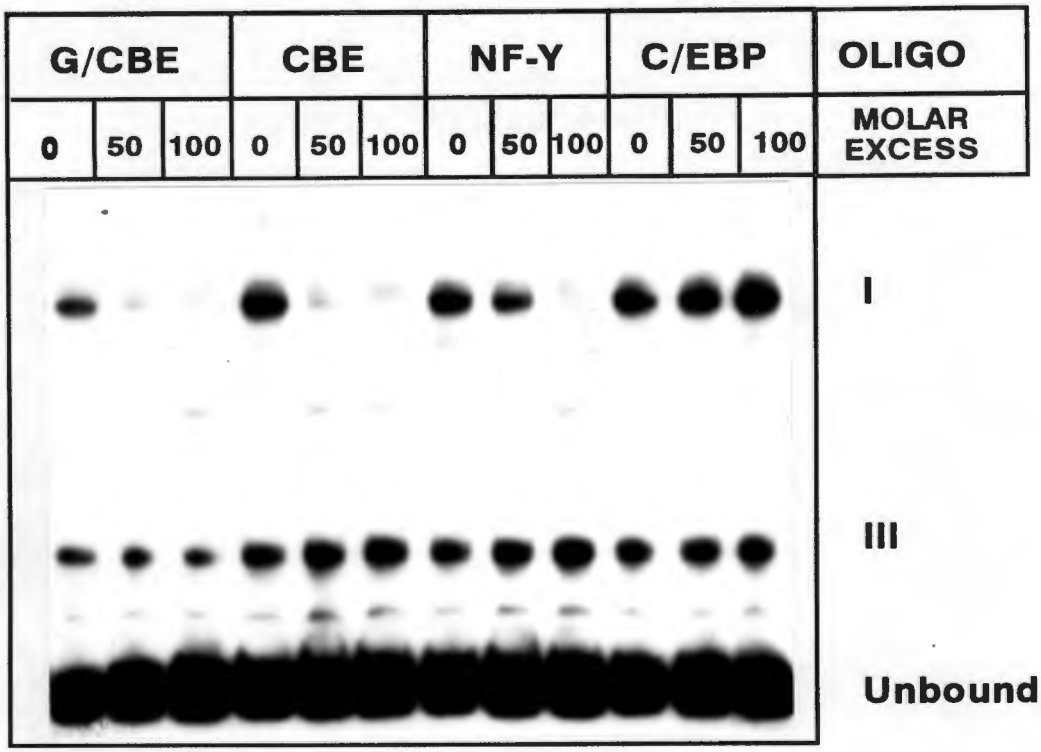


Fig 3.7 Competition of CT-1 extracts with known CCAAT-binding motifs. **(A)** Double-stranded oligonucleotides with known CCAAT-binding motifs were prepared as described in section 6.4.3.1. **(B)** Increasing molar excess of double-stranded oligonucleotides were incubated with CT-1 nuclear extracts for 10 mins prior to the addition of labelled COL1A2 -107 to -60 promoter fragment. The DNA-protein complexes were electrophoresed on non-denaturing 5% polyacrylamide gels as described in section 6.4.3. Gels were dried and exposed to X-ray film for 16 hrs. The positions of complexes I and III are indicated.

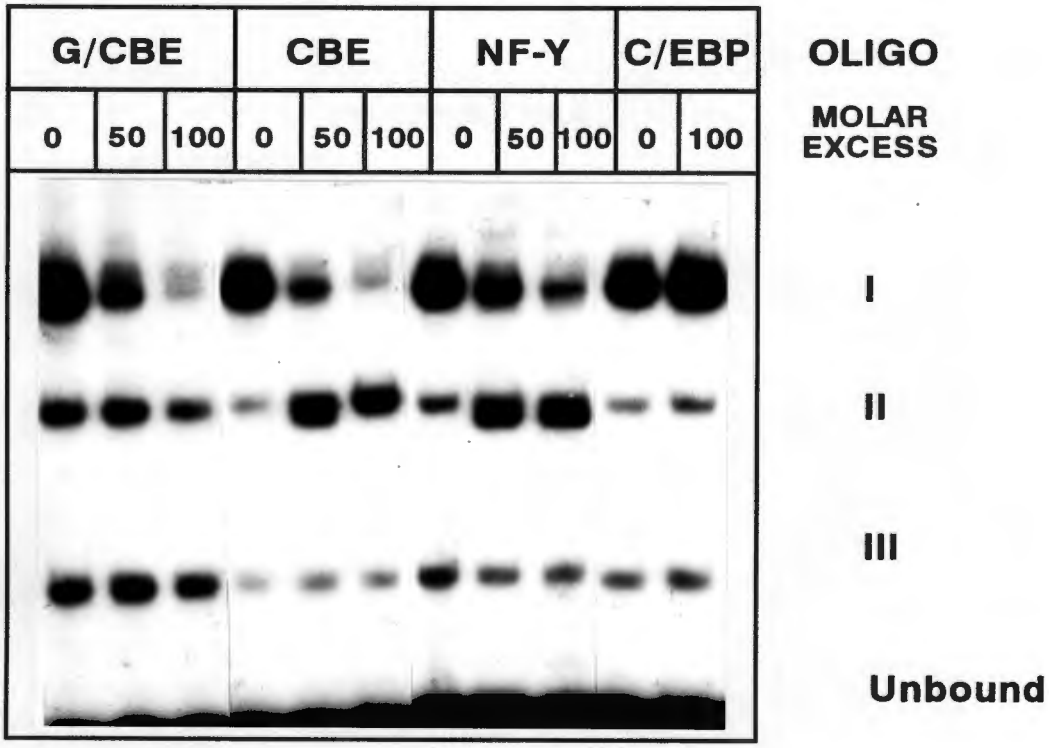


Fig 3.8 Competition of SVWI38 nuclear extracts with known CCAAT-binding motifs. Increasing molar excess of double stranded oligonucleotides were incubated with SVWI-38 nuclear extracts for 10 mins prior to the addition of labelled COL1A2 -107 to -60 promoter fragment. The DNA-protein complexes were electrophoresed on non-denaturing 5% polyacrylamide gels as described in section 6.4.3. Gels were dried and exposed to X-ray film for 16 hrs. The positions of complexes I, II and III are indicated.

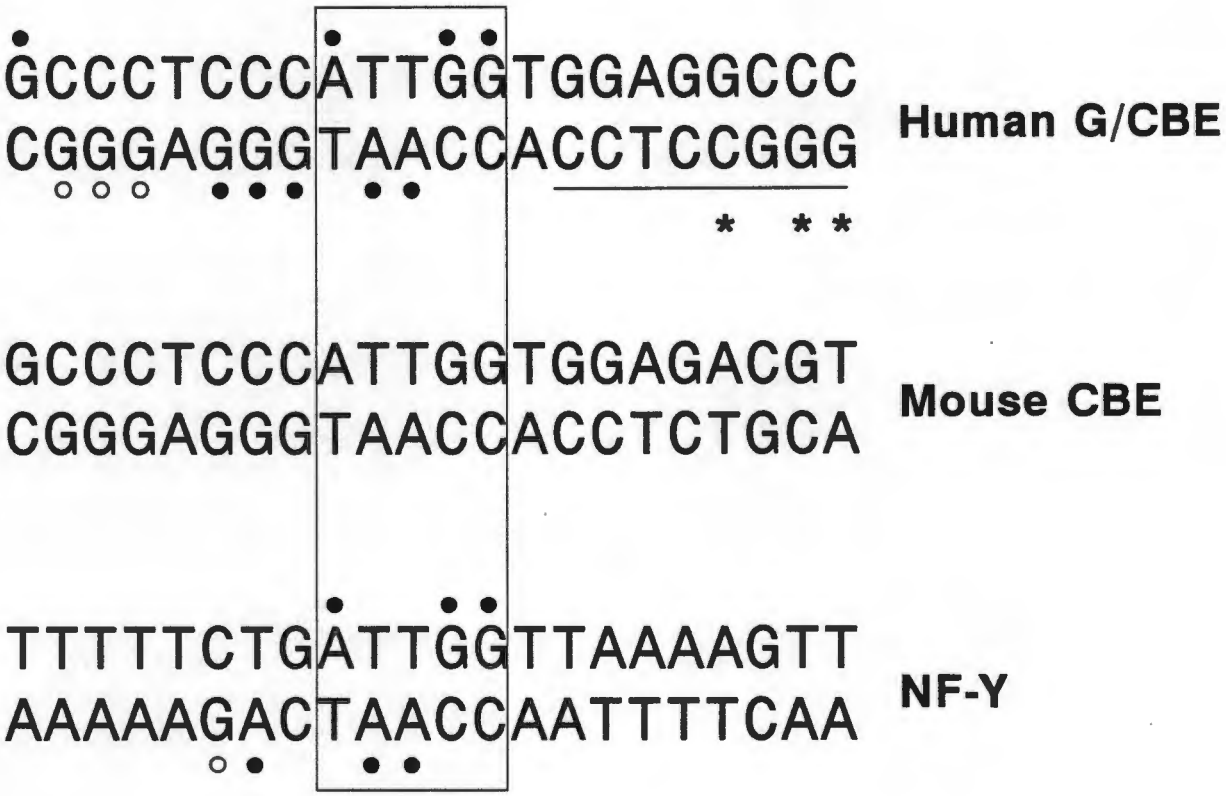


Fig 3.9 Comparison of sequences of known CCAAT box elements. Alignment of nucleotides -110 to -61 of the human CCAAT element with the mouse α2(1) collagen and NF-Y CCAAT elements. A high sequence homology is found between the human and mouse CCAAT box elements in the sequences 5' to the CCAAT box, while a 3 bp mismatch (indicated by asterisks) is found 3' of the CCAAT box. The inverted CCAAT boxes are boxed while the human CME is underlined. Full and partial methylation of purines in the G/CBE (Collins et al, 1997) and NF-Y (Chodosh et al, 1988) are indicated by (•) and (o) respectively.

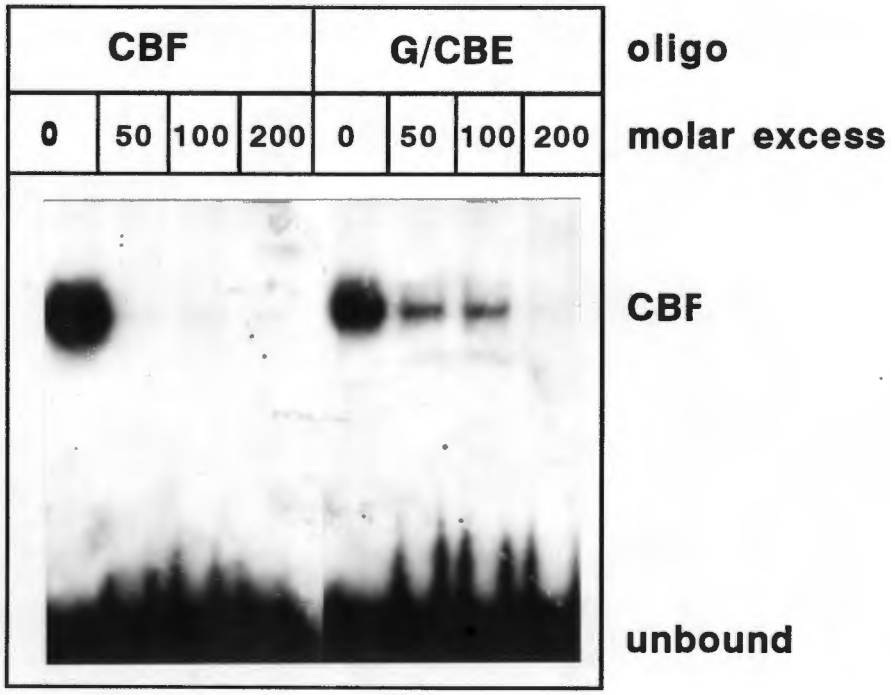


Fig 3.10 Competition of mouse CBF binding by human G/CBE. Increasing molar excess of double-stranded oligonucleotides were incubated with 3T3 nuclear extracts for 10 mins prior to the addition of labelled mouse Col1A2 -107 to +54 promoter fragment. The DNA-protein complexes were electrophoresed on non-denaturing 5% polyacrylamide gels as described in section 6.4.3. Gels were dried and exposed to X-ray film for 16 hrs. The position of the mouse CBF is indicated.

In EMSA's using increasing amounts of anti-CBF-B antibody, a supershifting of complex I proteins was observed, whereas complex III formation remained unaffected in CT-1 cells (fig 3.11.A). Complex II binding also remained unaffected with increasing amounts of antibody as observed with longer exposure of the gel to X-ray film (not shown). The supershifted complex is an indication that the anti-CBF-B antibody, although interacting with the B-subunit, does not interfere with complex I binding to its recognition sequence. Similar results were obtained when using the antibody in combination with mouse 3T3 nuclear extracts before addition of the mouse proximal promoter fragment (fig 3.11.B). Experiments with the anti-CBF-A antibody produced no supershifted complexes in both CT-1 nuclear extracts with the human promoter fragment and 3T3 nuclear extracts with the mouse promoter fragment (fig 3.11.C and D). This antibody has, however, been reported to result in the supershift of the CBF-DNA complex (Hasegawa et al, 1996), and it is possible that the antibody dilution used by us was too low to detect the desired effect.

The competition experiments and antibody studies collectively provide support for the suggestion that the G/CBE binding proteins (complex I) and the mouse CBF are related to each other.

3.2.5 The role of phosphorylation in trans-acting factor binding to the α2(1) procollagen promoter.

Different stimuli may affect gene expression due to the activation of protein kinases. It is highly probable that the binding activities and function of the proximal α2(1) procollagen promoter-binding proteins are regulated by phosphorylation. To determine whether protein phosphorylation is required in the binding of complex I, II and III proteins, EMSA's were performed using phosphatase treated nuclear extracts and nuclear proteins extracted from cells treated with protein kinase inhibitors.

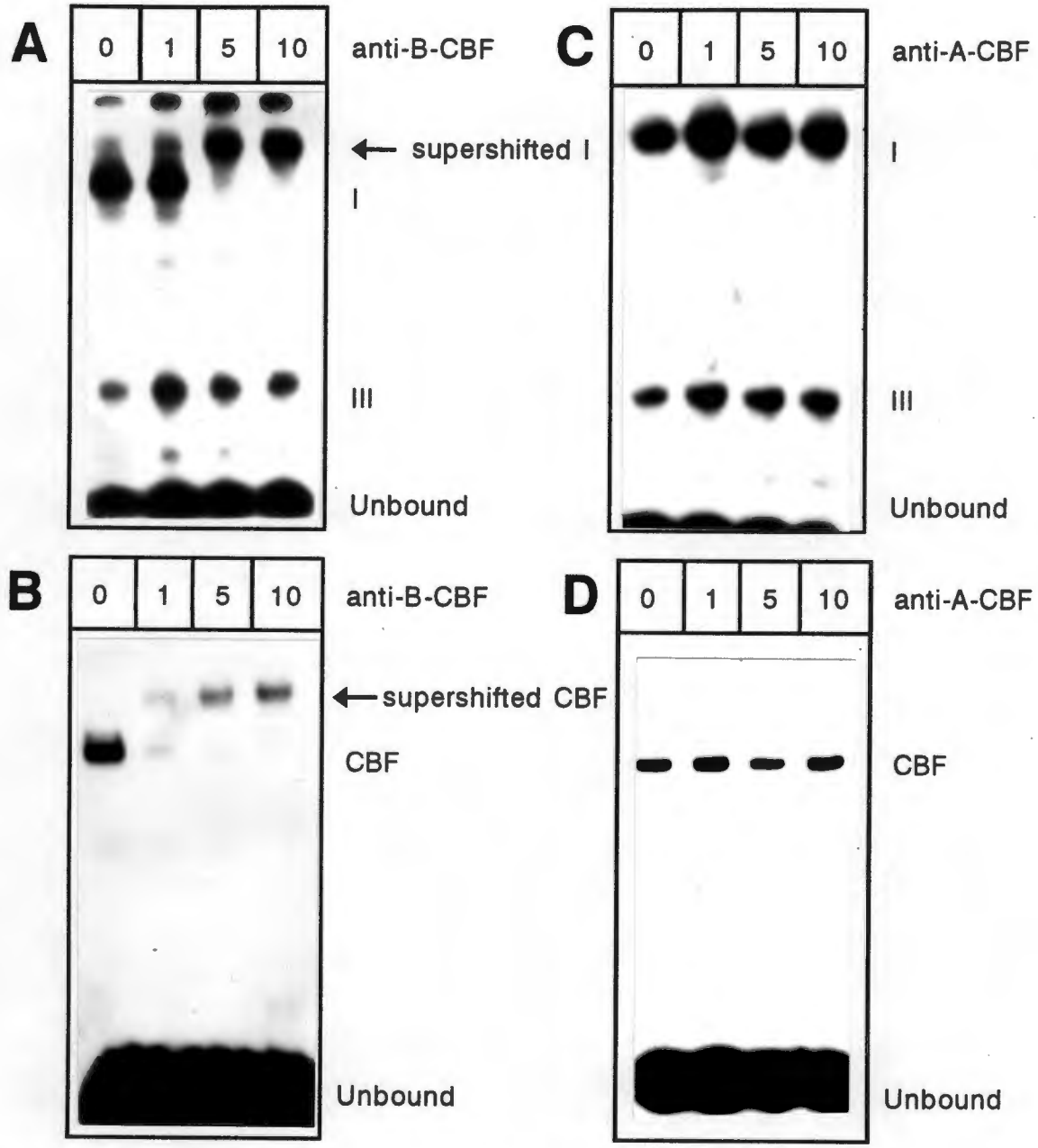


Fig 3.11 Supershift assays with mouse anti-CBF antibodies. **(A)** Increasing amounts (1,5,10µl) of a 1/100 dilution of rabbit anti-mouse CBF-B antibody was incubated with CT-1 nuclear extracts for 30 minutes on ice before addition of labelled human -107 to -60 α2(1) promoter fragment. The DNA-protein complexes were electrophoresed on non-denaturing 5% polyacrylamide gels as described in section 6.4.3.2. Gels were dried and exposed to X-ray film for 16 hrs. The supershifted DNA-protein complex is indicated. **(B)** Incubation of the rabbit anti-mouse CBF-B antibody with 3T3 nuclear extracts for 30 minutes prior to the addition of mouse -107 to +54 α2(1) promoter fragment. EMSA's were performed as described in section 6.4.3.2. Supershifted CBF is indicated. **(C)** Incubation of rabbit anti-mouse CBF-A antibody with CT-1 nuclear extracts and the human proximal α2(1) procollagen promoter as described above. No supershifted complexes are detected. **(D)** Incubation of rabbit anti-mouse CBF-A antibody with 3T3 nuclear extracts and the mouse proximal α2(1) procollagen promoter as described above. Supershifting of CBF is not observed.

3.2.6.1 Phosphatase treatment of nuclear extracts

Two phosphatases, calf intestinal phosphatase (CIP) and the serine/threonine specific phosphatase PP2A, were used to dephosphorylate CT-1, U937 and mouse 3T3 nuclear extracts. The phophatase treated nuclear extracts were then used in EMSA's to determine DNA binding activity. In all the cell lines, CIP treatment interferes with the binding of complex I and CBF (fig 3.12). CIP treatment of U937 and CT-1 nuclear extracts resulted in a 70% and 67% decline in complex I binding respectively. Complex II and III formation remained largely unaffected by CIP treatment. The ser/thr phosphatase, PP2A, however had a significant effect on complex II and III formation. In U937 nuclear extracts, complex II formation is entirely abolished by PP2A treatment while in CT-1 nuclear extracts complex III formation is abolished (fig 3.12 A and B). It would therefore appear that complex II and III proteins require ser/thr phosphorylation to enable interaction with their recognition sequence, the CME. The binding of CBF to the mouse promoter was completely inhibited by treatment of 3T3 nuclear extracts with CIP, while PP2A had no effect (fig 3.12.C). These results are an indication that phosphorylation of these DNA binding proteins are required for DNA-protein interaction in both human and rodent nuclear extracts. In terms of its' phosphorylation status, therefore, the human G/CBE proteins and mouse CBF behave similarly.

3.2.5.2 Inhibition of protein kinases in cell culture

The kinase inhibitor, staurosporin was used to inhibit intracellular kinase activity. U937 cells were treated with 1 μM staurosporin for 3 hours, nuclear proteins were extracted and incubated with labelled -107 to -60 human α2(1) procollagen promoter fragment in EMSA's (fig 3.13.A). No effect was observed in this instance. Cells were therefore treated with various concentrations of the drug for 16 hours. At a concentration of 10 nM for 16 hours, a decline in interaction of all

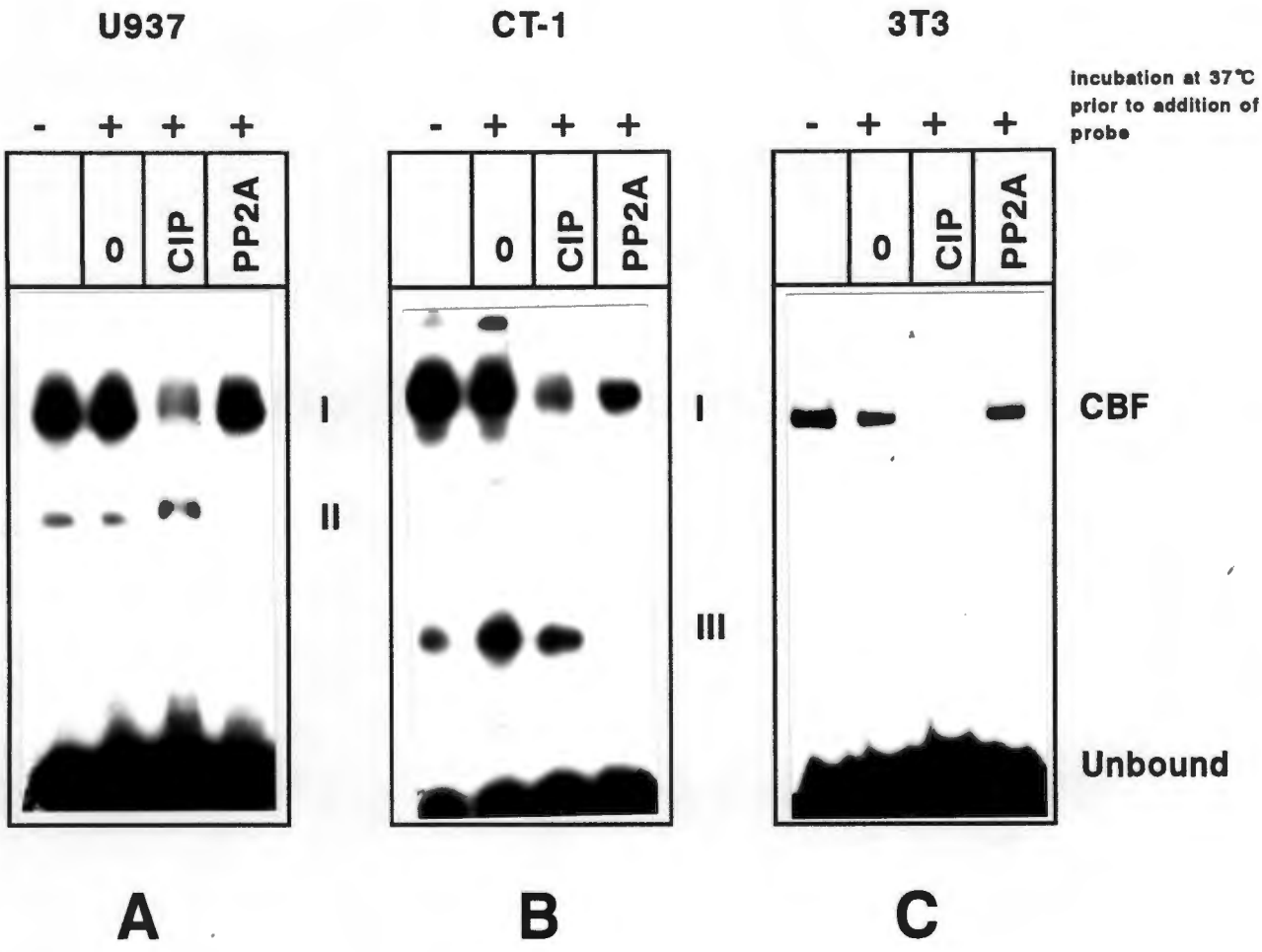


Fig 3.12 The effect of protein dephosphorylation on DNA-protein complex formation. **(A)** 1 unit of calf intestinal phosphatase (CIP) or 0.4 units of phosphatase PP2A were incubated with 4 μg of U937 nuclear extracts for 30 mins at 37°C. Labelled COL1A2 -107 to -60 promoter fragment was then added and incubated with the protein mixture for 30 mins on ice. DNA-protein complexes were electrophoresed on non-denaturing 5% polyacrylamide gels at 4°C in 0.5X TBE (section 6.4.3). Gels were dried and exposed to X-ray film for 16 hours. The positions of complexes I and II are indicated. **(B)** 1 unit of calf intestinal phosphatase (CIP) or 0.4 units of phosphatase PP2A were incubated with 4 μg of CT-1 nuclear extracts for 30 mins at 37°C. EMSA's with the human COL1A2 probe were performed as described in (A). The position of complexes I and III are indicated. **(C)** 1 unit of calf intestinal phosphatase (CIP) or 0.4 units of phosphatase, PP2A were incubated with 4 μg of 3T3 nuclear extracts for 30 mins at 37°C. EMSA's were performed as described in (A) except that the labelled probe used was the mouse Col1A2 -107 to +54 promoter fragment.

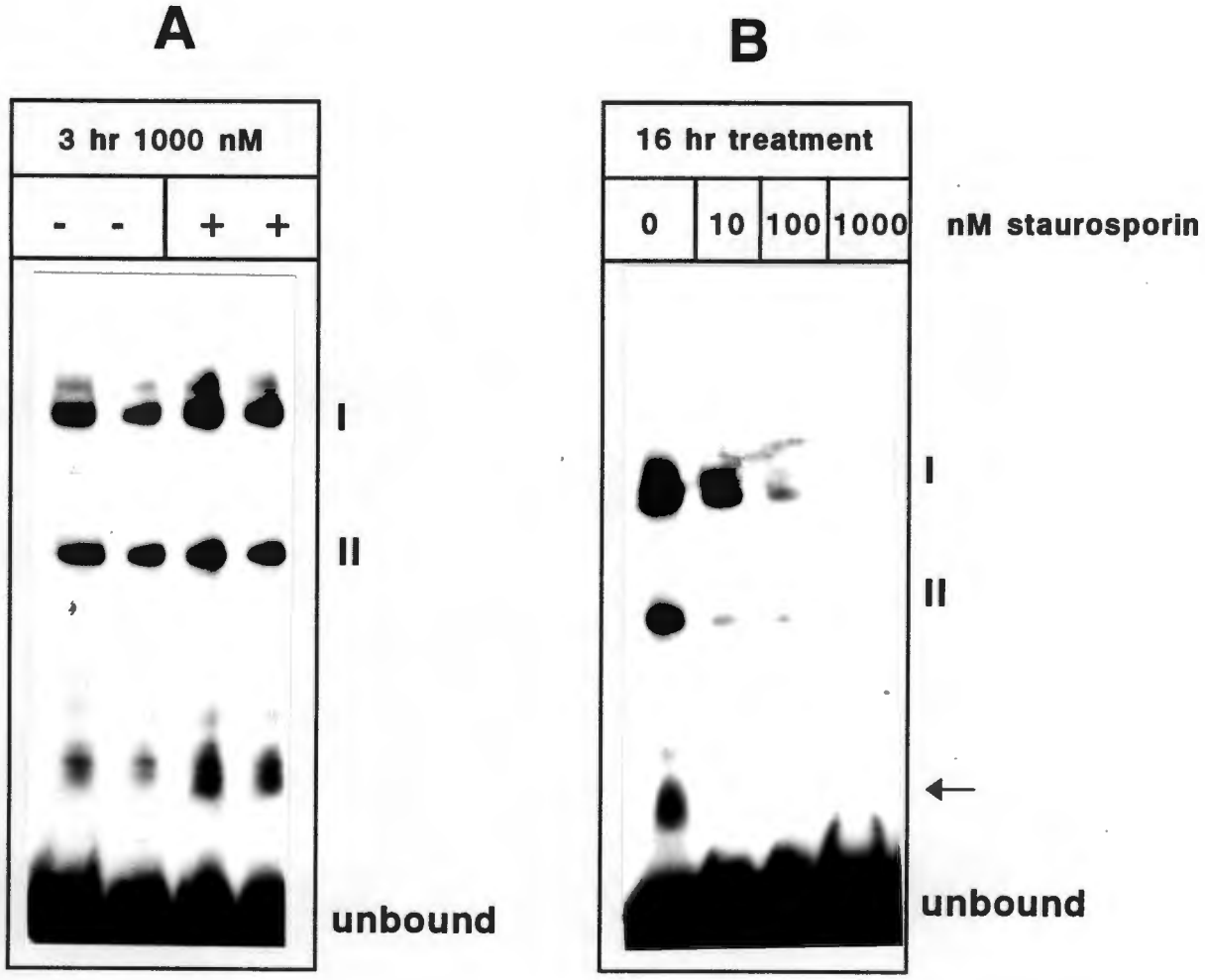


Fig 3.13 The effect of staurosporin on DNA-binding activity. **(A)** U937 cells were treated with 1μM staurosporin for 3 hours before nuclear proteins were harvested as described in section 6.4.1. 4 μg of nuclear proteins from untreated or treated cells were incubated with the -107 to -60 COL1A2 promoter fragment. DNA-protein complexes were electrophoresed on non-denaturing 5% polyacrylamide gels. Gels were dried and exposed to X-ray film for 16 hours. **(B)** U937 cells were treated with the indicated concentrations of staurosporin for 16 hours. Nuclear proteins from untreated or treated cells were used in EMSA's as described in (A). The fastest migrating band (indicated by an arrow) are possible degradation products of complexes I and II and its migration pattern is distinct from that of complex III. It is also not consistently observed.

complexes was observed, with no binding detected at 1000 nM (fig 3.13.B). This result was probably due to cell death, since a large percentage of dead cells were observed in culture at the higher concentrations (100 and 1000 nM) using the trypan blue exclusion test.

The optimum treatment time which resulted in minimal cell death (<15% as observed by the trypan blue exclusion test) for the U937 cell line, was found to be 8 hours. Cells were treated with varying concentrations of staurosporin for 8 hours and nuclear proteins extracted for use in EMSA's (fig 3.14.A). Binding studies showed that both complex I and II formation was significantly lower in nuclear extracts isolated from staurosporin treated U937 cells. The extent of the inhibition was similar for both complexes as determined by densitometric scanning of the autoradiographs (fig 3.14.B). These results are representative of three different experiments. Similar results were obtained with nuclear proteins extracted from CT-1 cells treated with staurosporin, albeit at concentrations lower than that used in the treatment of U937 cells. CT-1 cells showed a greater sensitivity to staurosporin treatment and concentrations greater than 1nM resulted in cell death within 30 minutes. Staurosporin treatment resulted in a loss in the binding of both complex I and complex III, although complex III binding was affected to a lesser extent (fig 3.15.A and B).

Staurosporin predominantly affects protein kinase C activity as well as that of ser/thr kinases and some tyrosine kinases (Rosenshine et al, 1994). To determine whether phosporylation of tyrosine residues are required for DNA-protein complex formation, cells were incubated with the tyrosine kinase inhibitor, genistein. Nuclear proteins from genistein treated cells showed no change in DNA-binding activities (fig 3.16.A). The finding that complex II/III binding is inhibited by treatment with the ser/thr phosphatase, PP2A and that the tyrosine kinase inhibitor, genistein had no effect on DNA binding activity, indicated that phosphorylation of these proteins more than likely occur at ser/thr residues. Treatment of cells with the phosphatase inhibitor, okadaic acid had no effect on

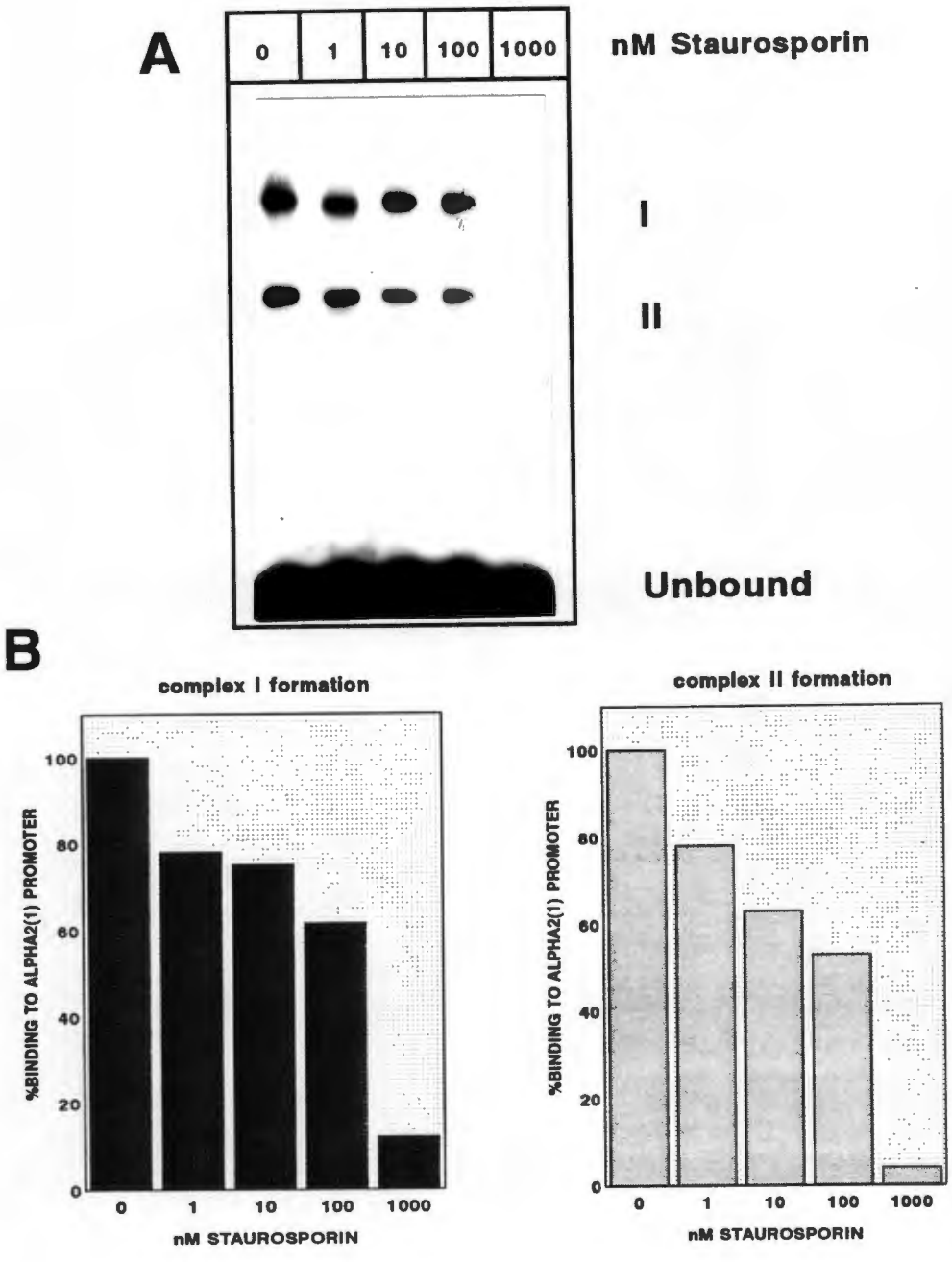


Fig 3.14 Effect of staurosporin treatment of U937 cells on DNA-binding activity. **(A)** U937 cells were treated with the indicated concentrations of staurosporin for 8 hours before extraction of nuclear proteins as described in section 6.4.1. 4 μg of nuclear proteins from untreated or treated cells were incubated with the -107 to -60 COL1A2 promoter fragment. DNA-protein complexes were electrophoresed on non-denaturing 5% polyacrylamide gels. Gels were dried and exposed to X-ray film for 16 hours. **(B)** Bar graphs of densitometric scanning of autoradiographs in fig 3.14.A.

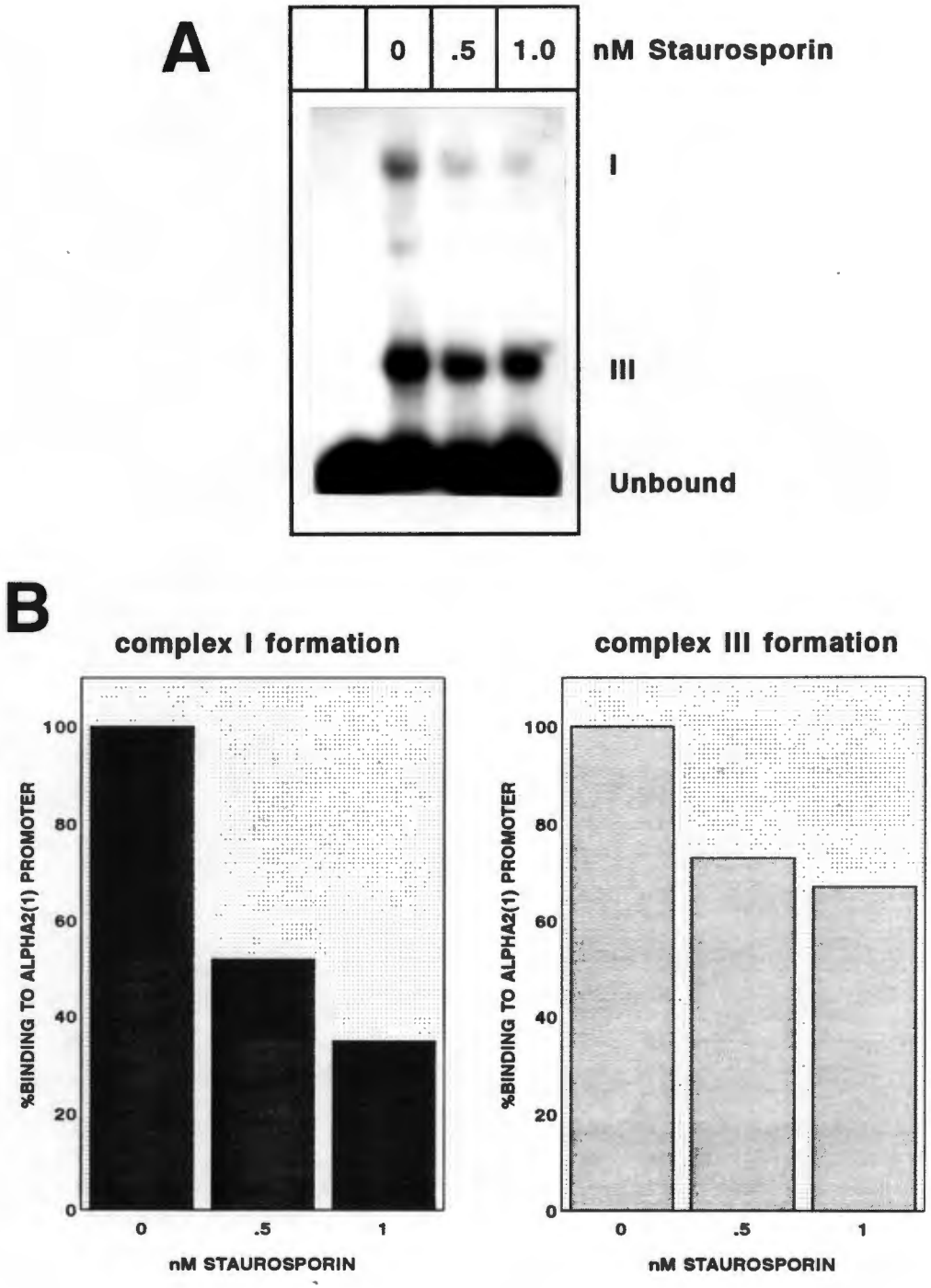


Fig 3.15 Effect of staurosporin treatment of CT-1 cells on DNA-binding activity. **(A)** CT-1 cells were treated with the indicated concentrations of staurosporin. Nuclear proteins were extracted as described in section 6.4.1 and 4μg protein was incubated with the -107 to -60 COL1A2 promoter fragment and DNA-protein complexes electrophoresed on non-denaturing 5% polyacylamide gels. Dried gels were exposed to X-ray film for 16 hours. **(B)** Bar graphs of densitometric scanning of autoradiographs in fig. 3.15.A.

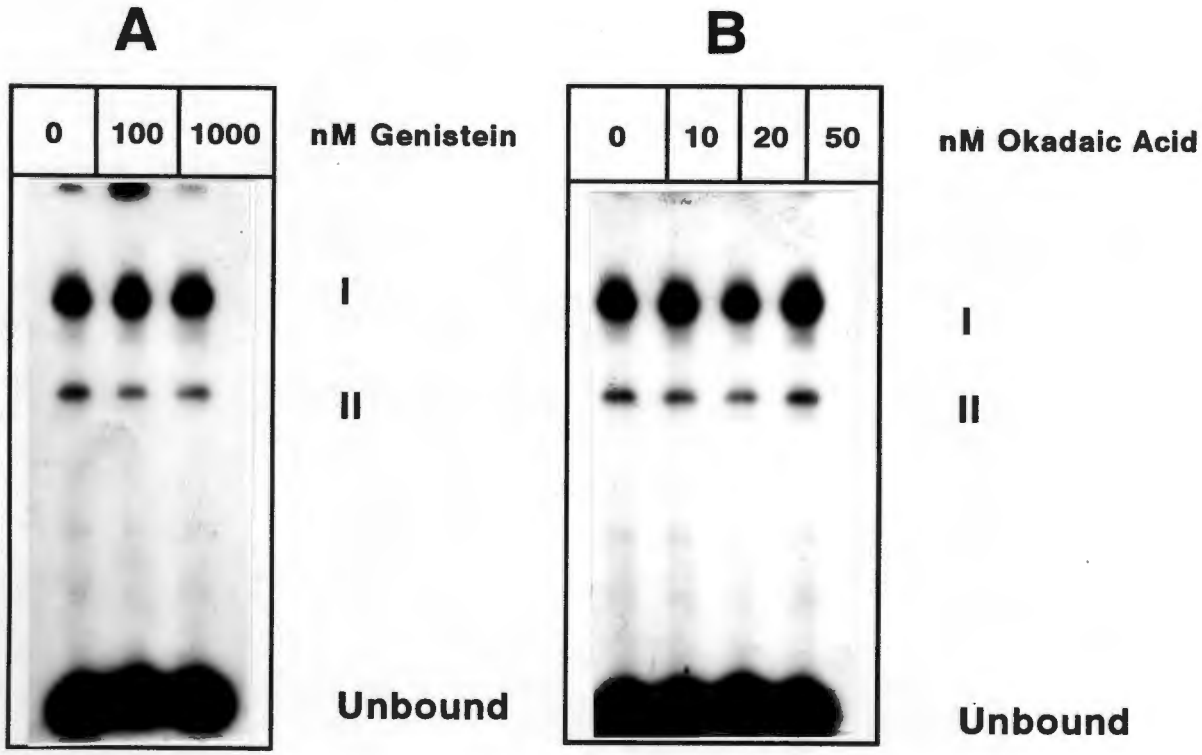


Fig 3.16 Effect of genistein and okadaic acid on DNA-binding activity. (A) U937 cells were treated with the indicated concentrations of genistein for 8 hours before extraction of nuclear proteins (section 6.4.1). 4 μg of nuclear proteins from untreated or treated cells were incubated with the human -107 to -60 COL1A2 promoter fragment. DNA-protein complexes were electrophoresed on non-denaturing 5% polyacrylamide gels. Gels were dried and exposed to X-ray film for 16 hours. **(B)** U937 cells were treated with the PP2A phosphatase inhibitor, okadaic acid, as indicated. Nuclear proteins were extracted and EMSA's performed as described in (A).

the DNA-binding activities of any of the complexes (fig 3.16.B). There may even be a slight increase in complex formation with increasing okadaic acid concentration. This would be consistent with the idea that phosphorylation is required for DNA binding activity. It is clear from the above experiments that phosphorylation has a significant effect on the binding of proteins to the proximal promoter of the α2(1) procollagen gene, since inhibiting with such events result in the loss of this activity.

3.2.5.3 Effect of kinase inhibition on promoter activity and mRNA levels

Since dephosphorylation of nuclear factors and inhibition of kinases by staurosporin resulted in a loss of complex formation at the -107 to -60 region of the α2(1) procollagen promoter, the effect of kinase inhibition on the activity of the proximal promoter was investigated. Proximal α2(1) procollagen promoter-CAT constructs were transfected into CT-1 cells as described in section 3.2.3. After transfection, the cells were treated with staurosporin at 0.5 and 1nM for 16 hours. The results obtained indicated a significant loss in α2(1) procollagen promoter activity in the presence of both 0.5 and 1nM staurosporin (fig 3.17). This therefore suggests that interference of kinase activation pathways result in the inhibition of the proximal α2(1) procollagen promoter activity. To determine the effect of kinase inhibition on the steady state α2(1) collagen mRNA levels, total RNA was extracted from CT-1 cells treated with staurosporin as described above (fig 3.18.A) and used in Northern blot analysis. Staurosporin treatment resulted in significant inhibition of endogenous α2(1) procollagen gene expression when compared to cells that have not been treated (fig 3.18.B). The filters were stripped of the α2(1) procollagen probe and reprobed with radioactively labelled β-actin cDNA. As observed with α2(1) procollagen, β-actin expression was also inhibited in CT-1 cells with increasing concentrations of staurosporin (fig 3.18.B). The inhibition is observed even though the loading of total RNA in lane 2 (0.5nM staurosporin) is more than that in lane 1 (no staurosporin).

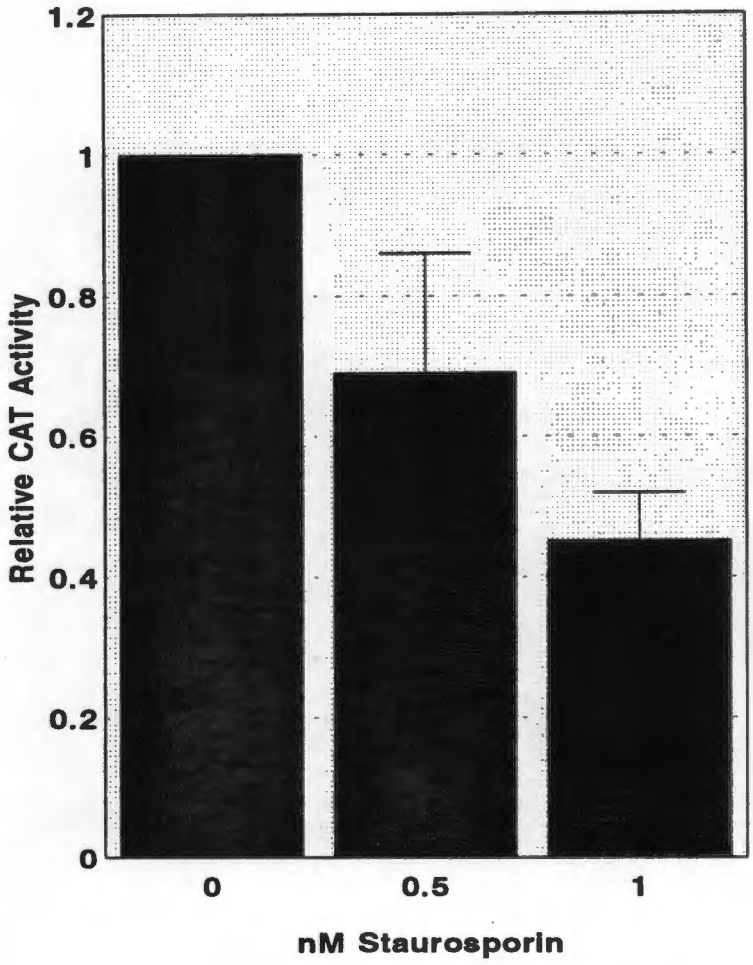


Fig 3.17 Effect of staurosporin on the human proximal α2(1) procollagen promoter activity. 10 μg of the human -107COL-CAT construct was transfected into CT-1 cells (as described in section 6.9), afterwhich the cells were treated for 16 hours with the indicated concentrations of staurosporin. Results are the mean ± S.D. of four independent experiments in which each construct was transfected onto triplicate 100 mm dishes of CT-1 cells. The activity in the absence of staurosporin was set at 1.0. Treatment with 0.5 nM and 1 nM staurosporin was significantly different from the control with $p<0.05$ and $p<0.001$ respectively (Mann-Whitney *U* test).

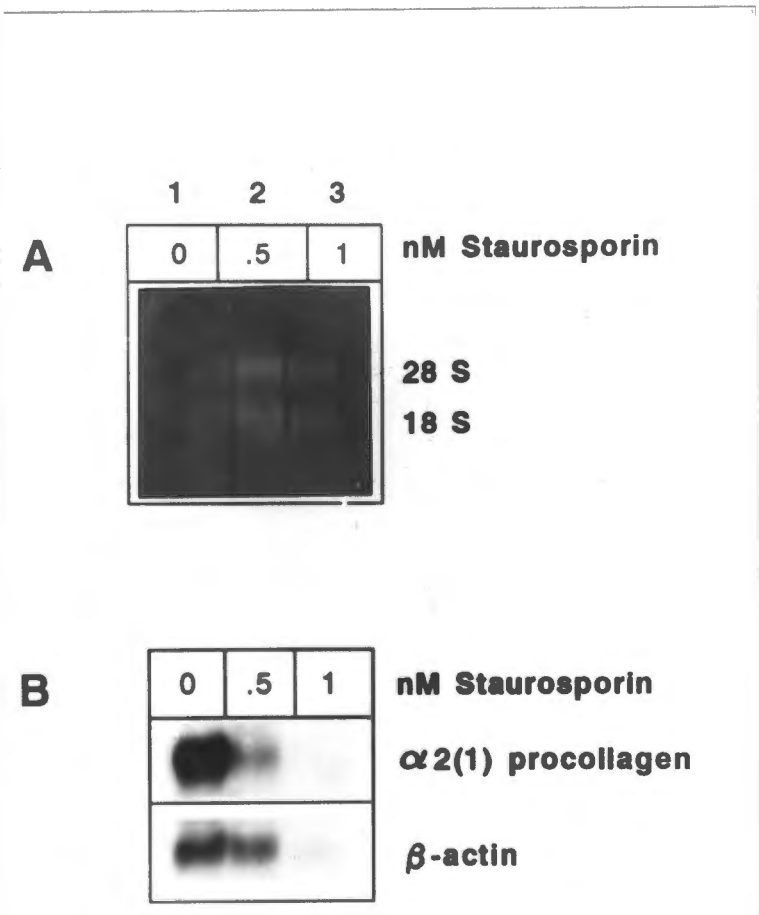


Fig 3.18 Effect of staurosporin on steady state α2(1) procollagen mRNA levels. (A) Total RNA from CT-1 cells exposed to 0.5nM and 1nM staurosporin were isolated and 5μg run on 1% agarose-formaldehyde gels. The gels were stained with ethidium bromide (1μg/ml) to visualise the quantities of 18S and 28S rRNA. **(B)** Northern blot analysis of the levels of α2(1) procollagen and β-actin mRNA in the absence and presence of staurosporin.

The results presented indicate that kinase activity, possibly that of protein kinase C, has a role in the regulation of the α2(1) procollagen gene. It was therefore decided to investigate the effect of phorbol 12-myristate 13-acetate (PMA), an inducer of protein kinase C activity, on α2(1) procollagen gene expression.

3.2.5.4 Effect of PMA on collagen gene expression in CT-1 cells

CT-1 cells grown to confluency were growth arrested in serum free DMEM for 16 hours, then treated with 100ng/ml PMA for 30 minutes and RNA isolated as described in section 6.6.1. The steady-state α2(1) procollagen mRNA levels were determined by Northern blot analysis (section 6.6.2) and found to differ significantly between PMA treated and untreated cells (fig 3.19.A). CT-1 cells treated with PMA showed a decrease in endogenous levels of α2(1) procollagen mRNA compared to expression in untreated cells. The decrease varied between 30 and 60% relative to that of untreated controls in different experiments (fig 3.19.B). Interestingly, a recent study reported paradoxical responses of endogenous α2(1) procollagen gene expression and a transfected -378 COL1A2/CAT construct to PMA treatment in primary human fetal skin fibroblast. While PMA downregulated the expression of the endogenous gene, upregulation of the transfected promoter was observed (Greenwel et al, 1997). These results therefore confirm that protein kinase C and possibly other signal transduction pathways have a role in the regulation of α2(1) procollagen gene expression.

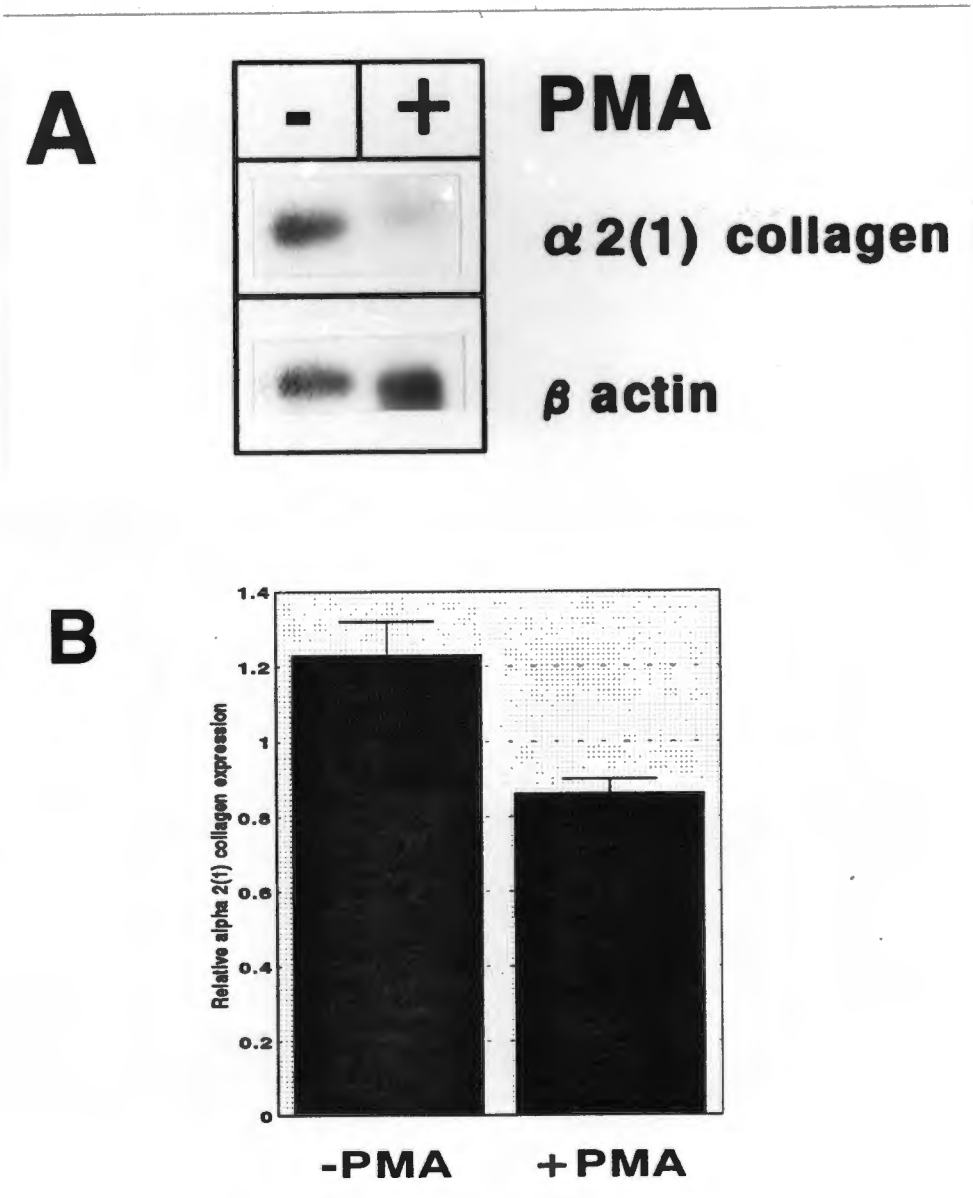


Fig 3.19 Effect of PMA on α2(1) procollagen gene expression in CT-1 cells. **(A)** CT-1 cells were exposed to PMA (100ng/ml) for 30 minutes after which total RNA was extracted using the method of Chomczinsky and Sacchi, (1987). 5μg of total RNA from untreated (-) and PMA treated (+) cells was analysed for α2(1) procollagen and β-actin mRNA by Northern Blotting (section 6.6.2). **(B)** Autoradiographs of four independent experiments were scanned using a densitometer and the ratio of α2(1) procollagen expression relative to that of β-actin is shown in the bar chart.

3.3 DISCUSSION

While DNA elements such as the CCAAT motif and a pyrimidine segment at -164 and -159 are functionally conserved between the human and mouse α2(1) procollagen promoters, the TGF-β responsive elements differ extensively between the two species. TGF-β induces NF-1 binding activity in the mouse promoter at -315 to -295 (Rossi et al, 1988), while Sp1 (Inagaki et al, 1994; Tamaki et al, 1995; Greenwel et al, 1997) or AP-1 (Chung et al, 1996) binding may be induced in the human promoter (section 1.2.3.3). Also, the Sp-1 binding sites at -300 and a TCCTCC motif (-133 to -118) appear to be specific for the human promoter (Ihn et al, 1996). These loci may in part be responsible for species-specific regulatory mechanisms of the human and mouse collagen genes.

The findings reported in chapter 2 indicated that cell-specific regulation of the human COL1A2 gene is associated with differences in trans-acting factor binding to the proximal promoter. One of the objectives of this section was therefore to determine whether the same association existed within collagen expressing and non-expressing rodent cells. The surprising result was that the correlation between trans-acting factors and collagen gene expression observed in human cells was not observed in rodent cells. This observation prompted a study into to the species-specific expression of the α2(1) procollagen gene. A 3 bp mismatch in the binding site of the complex II/III proteins between the human and rodent promoters was confirmed. Both the mouse and rat promoters contained the mismatch which was sufficient to interfere with complexes II and III formation on the rodent promoter. Although complex II proteins were present in rodent nuclear extracts, they were unable to bind the rodent promoter. It is possible that these proteins are involved in the regulation of other genes in rodent cell lines. A Genebank search for the CME (complex II and III binding sequence) showed that several genes contained the binding sequence. GGAGG (an integral part of the CME) boxes have been

reported to footprint in the promoters of a number of genes and often occur as repeats in these promoters (table 3.1).

Table 3.1 Gene promoters containing functional GGAGG sequences.

Gene	Location of GGAGG box	Reference
human α1(1) collagen	-477 to -332	Liska et al, 1992
human α2(1) collagen	-133 to -119	Ihn et al, 1996
grp78/BiP	-109 to -74	Roy and Lee, 1995
c-myc	-199 to -176	Suen and Hung, 1991 Bossone et al, 1992
EGF-r	-290 to -331	Johnson et al, 1988
α1(IV) collagen	-107 to +10	Piccolo et al, 1995
α1(XI) collagen	-199 to -36	Yoshioka et al, 1995
SPARC/osteonectin	-120 to -60	McVey et al, 1988
α-myosin heavy chain	-111 to -129	Flink & Morkin, 1990
bek/KGFR		Avivi et al, 1992
α2A-adrenergic receptor	-131 to -92	Handy &Gavras, 1992
AP-2		Imagwa et al, 1987
malT		Raibaud et al, 1988
JCV repeated sequence		Martin et al, 1985

The difference in trans-acting factor interaction in the human and mouse α2(1) procollagen promoters suggested species-specific transcriptional regulatory mechanisms. Measurement of promoter activity using human and mouse promoter-CAT constructs transfected into CT-1 cells, indicated that the rodent promoter had a lower activity than the human promoter. This result can be explained by the absence in the binding site for complex III proteins (potential activators of transcription) in the rodent promoter. The lower activity of the human promoter in 3T3 cells compared with CT-1 cells, can be explained by the presence of complex II (potential repressor) and the absence of complex III (potential activator). These findings provide further evidence for the role of the CME in regulation of the human α2(1) procollagen gene. Similar species-specific transcriptional regulation

has been reported for the Apolipoprotein A-I gene. A study by Sorci-Thomas and Kearns, (1991) reported that differences in apoA-I expression in human and two nonhuman primate liver cell lines can be linked with sequence divergence within a 494 bp region of the apoA-I gene. A subsequent study identified nucleotide differences in three sites (-189, -144 and -48) that significantly alter the binding affinity of nuclear factors as well as determining the species-specific regulation of the apoA-I gene (Sorci-Thomas and Kearns, 1995). The investigation presented in this chapter supports the idea that factors that regulate the expression of the α2(1) procollagen gene may be important determinants in species-specific regulation.

Competition experiments and antibody supershifted data suggest that the G/CBE binding proteins (complex I) is a member of the structurally related CCAAT box binding factors which include CBF, (Karsenty, et al, 1988, Hatamochi et al, 1988, Maity et al, 1988,1992), NF-Y (Chodosh et al, 1988), α-CP1 and CP-1 (Kim and Sheffery,1990). Although the antibody recognised both the human complex I and mouse CBF, there is the possibility that these two factors may be different. Collins, (1993) proposed that complex I is a 120 kD protein, while the mouse CBF is a heterotrimeric protein. It is possible that some areas within the two proteins have been conserved and that these are recognised by the anti-CBF antibody while others have been changed during evolution.

In an attempt to determine if phosphorylation is required for complex formation, phosphatase treatment of nuclear extracts resulted in the loss of DNA-binding activities. CIP affected complex I formation, while the ser/thr phosphatase, PP2A affected complexes II and III formation. The loss of DNA binding activities of dephosphorylated protein complexes was observed in both human and rodent nuclear extracts. Phosphorylation of the CCAAT binding factor is therefore required for binding to its consensus sequence. This is in contrast to studies by Heinrichs et al, (1993;1995) which showed that CCAAT-specific DNA-protein

interactions on the rat osteocalcin gene promoter is insensitive to phosphatase treatment. It should be noted that the phosphatase used in their study as well as the time of incubation of nuclear proteins with phosphatase differs from that used in this investigation. While Heinrichs et al, (1993) showed that the CCAAT-specific DNA-protein interactions were heat stable, the complex I, II and III proteins have been shown to be heat labile (Smith, 1989). Hatamochi et al, (1988) also showed that the mouse CBF-A subunit is heat sensitive. Heating of CBF-A for 5 minutes at 65°C resulted in a loss of CBF DNA binding activity. This provides more support that the human and rodent CCAAT-binding proteins are related.

Treatment of cells with the kinase inhibitor, staurosporin, also exerted a negative effect on the binding of trans-acting factors to the proximal α2(1) procollagen promoter, while a tyrosine kinase inhibitor had no effect. Staurosporin also had a significant inhibitory effect on α2(1) procollagen promoter activity and steady state mRNA levels. This inhibition can be associated with the loss of trans-acting factor interactions with the proximal promoter of the gene. The data therefore indicate that both ser/thr kinase and protein kinase C activities may be required for the activities of these trans-acting factors (summarised in fig 3.20) and possibly other transcription factors which regulate α2(1) procollagen gene expression. Protein phosphorylation is required in the regulation of a number of other genes. Sudbeck et al, (1994) showed that collagen-stimulated induction of keratinocyte collagenase is mediated via tyrosine kinase and protein kinase C activities. Post-translational modifications of the AP-1 components, Fos and Jun are modulated by various kinases (Binetruy et al, 1991, de Groot and Sassone-Corsi,1992, Derijard et al, 1994). Activators of protein kinase C may influence transcription by increasing the activity of AP-1 (Slack et al, 1993). This is of particular interest to the regulation of the human α2(1) procollagen promoter since Chung et al, (1996) proposed that AP-1 binds the -265 to -241 region in response to TGF-β stimilus.

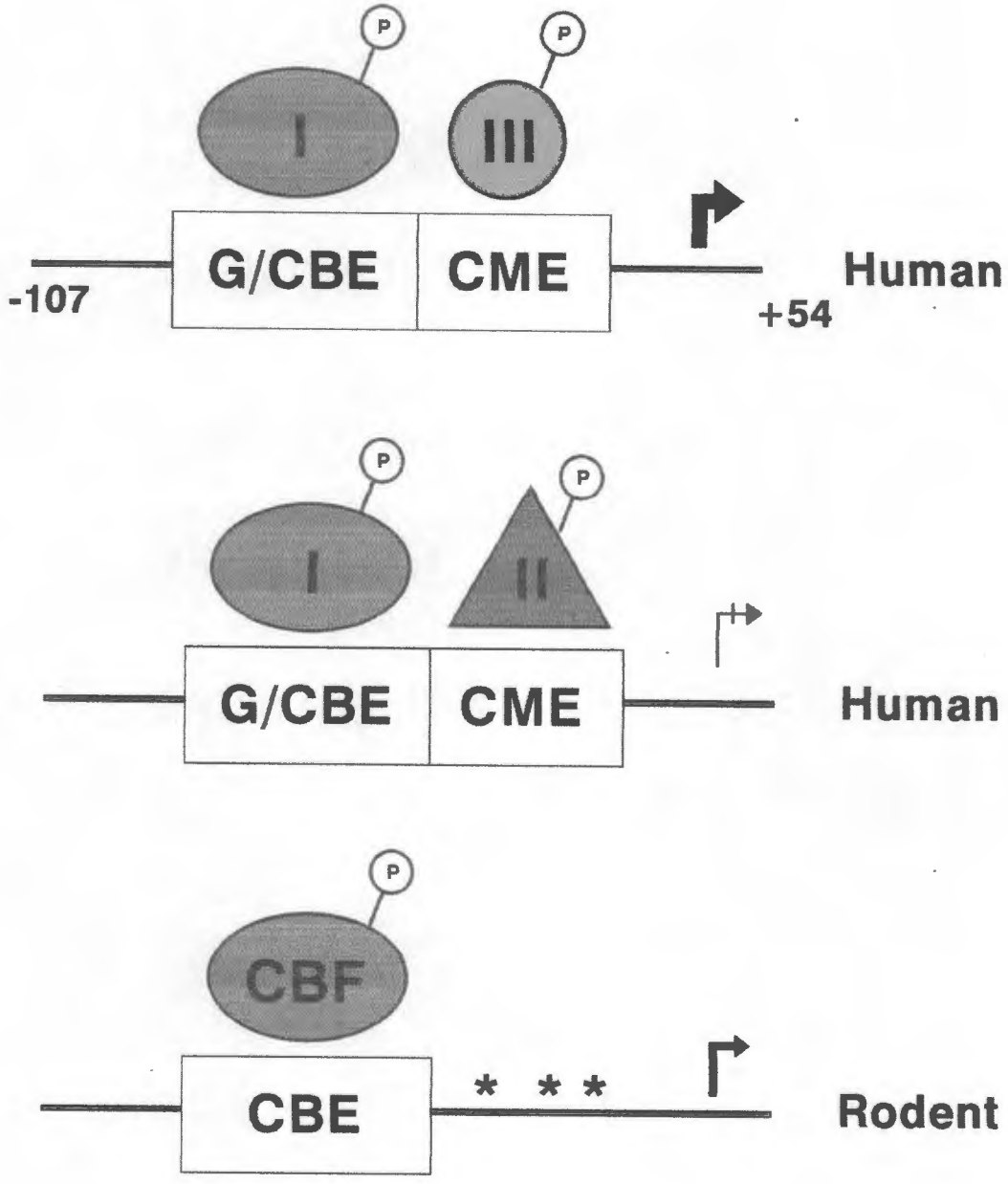


Fig 3.20 Model depicting species-specific differences and phosphorylation in trans-acting factors binding the α2(1) procollagen promoter. Phosphorylation of complexes I, II, III and CBF is required for binding to the proximal α2(1) procollagen promoter. The thickness of the arrow indicate the degree of promoter activity. The asterisks depicts the three base pair difference between the human and rodent promoters.

The present study also shows that the phorbol ester, phorbol 12-myristate 13-acetate (PMA), known to mediate mitogenic signal transduction, (Bell, 1986; Kikkawa and Nishizuka, 1986; Nishizuka, 1986) has a significant inhibitory effect on α2(1) procollagen gene expression. This is in agreement with other studies reporting that phorbol esters inhibit type I collagen gene transcription (Harrison et al, 1990; Rabin et al, 1986; Sobel et al, 1983; Greenwel et al, 1997). Contrary to these studies, Stuiver et al, (1991) reported that PMA had a stimulatory effect on collagen expression in 3T3-L1 cells. Collagen gene expression may therefore be mediated through different signalling pathways in a cell-specific manner. It is also clear that phosphorylation/dephosphorylation events have an effect on the DNA-binding activity of factors which interact both upstream and downstream of the CCAAT box in the α2(1) procollagen promoter, as well as the CCAAT-binding factor itself. These interactions may mediate both negative and positive influences on α2(1) procollagen gene transcription.

In summary, the work presented in this study indicate that:

1. Species-specific DNA-protein interactions in the proximal α2(1) procollagen promoter may be mediated via a 3 bp mismatch between the human and rodent promoter. The DNA sequences involved in this interaction includes the CCAAT box and an adjacent DNA element, the CME (in human).

2. The protein(s) associated with binding to the human CCAAT consensus sequence (G/CBE) is similar or related to the mouse CBF.

3. Phosphorylation via serine/threonine kinase and protein kinase C pathways are required for the binding of the CCAAT binding factor in both human and mouse nuclear extracts. Phosphorylation is also required for the binding of complex II and III proteins to the CME in the human α2(1) procollagen promoter and may induce conformational changes that would enable DNA-binding.

4. Inhibition of phosphorylation with staurosporin, a kinase inhibitor resulted in decreased activity of the proximal α2(1) procollagen promoter and inhibition of gene transcription.

5. Stimulation of protein kinase C activity resulted in a significant decrease in α2(1) procollagen mRNA levels in PMA treated cells. Stimulation of kinase activity and subsequent phosphorylation of complexes I, II and III are possibly terminal events in signalling pathways that result in either the activation or repression of the α2(1) procollagen gene.

4. CLONING THE GENES CODING FOR THE α2(1) PROCOLLAGEN PROMOTER-BINDING PROTEINS

CHAPTER 4:

CLONING THE GENES CODING FOR THE α2(1) PROCOLLAGEN PROMOTER-BINDING PROTEINS.

4.1 INTRODUCTION

The work presented in Chapters 2 and 3 identified DNA-binding proteins which bind the proximal promoter of the human α2(1) procollagen gene in a sequence-specific manner. At least two of these DNA-protein complexes (complex II and III) appear to be specific for the regulation of the human COL1A2 gene. Not only did this binding occur in a cell-specific manner, but regulation by these factors may also be species-specific. Both the species- and cell-specific trans-acting factor interactions are accompanied by differences in transcriptional activity of the α2(1) procollagen gene.

Since the complex II and III proteins appear to be novel in their binding to the human α2(1) procollagen promoter, it was important to characterise the role of these factors in the regulation of the COL1A2 gene. The classical way in which to identify transcription factors has been to purify the DNA-binding proteins using oligonucleotide affinity matrices. This approach was used successfully by a number of groups since it was first reported by Kadonaga and Tijian, (1986) in the purification of Sp1. This technique was also used in combination with ion-exchange chromatography in the identification of the mouse CCAAT binding proteins (Hatamochi et al, 1988).

Although the technique is highly successful, one of its limitations is the requirement for large amounts of starting material (cells or tissue), which makes the purification of transcription factors on a preparative scale difficult. To circumvent this problem, the technique of screening an expression cDNA library in *E.coli* with the recognition element of the DNA-binding protein has been developed by Singh et al, (1988). The strategy is modified from that developed for the isolation of genes using antibodies to screen recombinant libraries (Broome, 1978). It was used to isolate a cDNA clone encoding a protein which has DNA recognition properties that overlap with those of H2TF1 and NF-κB (Singh et al, 1988). The technique is ideally suited for the isolation of clones encoding rare DNA-binding proteins and a number of other genes have subsequently been isolated (Table 4.1). The screening of λ expression libraries with recognition-site DNA probes has become one of the prefered techniques in the isolation of genes coding for transcription factors.

This technique has certain limitations, however, one of which is that the overexpression of β-galactosidase proteins in *E.coli* may lead to the formation of insoluble aggregates called inclusion bodies. A small proportion of DNA-binding proteins in an inclusion body may actually be in the correctly folded conformation suitable for binding to its ligand. As an improvement to the method, Vinson et al, (1988) showed that exposing protein immobilised on nitrocellulose filters to the denaturing agent, guanidine hydrochloride, enhances the presence of active molecules.

This modified method was used to identify cDNA clones encoding trans-acting factors which are unique in their binding to the human α2(1) procollagen promoter. Since the complex II/III proteins may be novel to regulation of the human α2(1) procollagen gene, efforts were concentrated on the identification of these proteins.

Table 4.1

CLONES ENCODING SEQUENCE-SPECIFIC DNA BINDING PROTEINS ISOLATED BY SCREENING OF EXPRESSION LIBRARIES

Clone	Binding site	Organism	Reference
MBP-1	GGGGATTCCCC	Human	Singh et al, 1988 Balwin et al, 1990
Oct-2	ATGCAAAT	Human	Clerc et al, 1988 Muller et al,1988 Staudt et al, 1988
Oct-1	ATGCAAAT	Human	Sturm et al, 1988
E12	GGCAGGTG	Human	Murre et al, 1989
XBP	GCTGGCAACTGTGTGACGTCATCACAAGA	Mouse	Hsiou-Chi et al,1988
RF-X	CCCCCTAGCAACAG	Human	Reith et al, 1989
IRF-1	AAGTGA	Mouse	Miyamoto et al, 1988
Pit-1	GATTACATGAATATTCATGA	Rat	Ingraham et al, 1988
CREB	TGACGTC	Human	Hoeffler et al, 1988
CNBP	GTGCGGTG	Human	Rajavashist et al, 1989
HBP-1	ACGTCA	Wheat	Tabata et al, 1989
ITF-1, 2	AACACCTGCAGCAGCAGCTGGCAGG	Human	Henthorn et al, 1990
Isl-1	TTAATAATCTAATTA	Rat	Karlsson et al, 1990
PU.1	GAGGAA	Mouse	Klemsz et al, 1990
ICSBP	AGTTTCACTTCT	Mouse	Driggers et al, 1990
Ets-1	AGCCACATCCTCTGGAA	Human	Ho et al, 1990
OBF1	GATCCAAGTGCCGTGCATAATGATGTGGG	Yeast	Biswas et al, 1990
FRG-Y1, Y2	CTGATTGGCCAA	*Xenopus*	Tafuri et al, 1990
TEF-1	GTGGAATGT	Human	Xiao et al, 1991
HTF4	CAGCTGG	Human	Zhang et al, 1991
NF-E1	CCTCCATC	Human	Park et al, 1991
AT-BP1	GATCCCAGCCAGTGGACTTAGCCC	Rat	Mitchelmore et al, 1991
AT-BP2	AACTGGGGTGACCTTGGTTAATATTCACCAG	Rat	Mitchelmore et al, 1991
E1A-F	ACAGGAAGTGACACGGATGTGGC	Human	Higashino et al, 1993

modified from Singh, 1993.

4.2 RESULTS

4.2.1 Selection of probes for screening a λgt11 expression library

As previously described, the complex I, II and III proteins bind distinct elements within the proximal α2(1) procollagen promoter (-107 to -60). The element involved in binding complex I proteins is the G/CBE and the element binding both the complex II and III proteins is the CME (see chapter 2). These two elements are distinct protein binding sequences as observed by competition experiments using unlabelled double-stranded DNA oligonucleotides corresponding to these sites as well as mutated oligonucleotides in EMSA's (Collins et al, 1997). The complex II and III proteins bind the CME with similar affinity (section 2.2.6) and appear unique in their binding to the proximal human α2(1) procollagen promoter.

Since the affinity of these protein complexes for the CME are similar, the use of the CME as a probe to screen an expression library, would allow the identification of genes encoding either of these factors. Oligonucleotides used in screening a λgt11 expression library were synthesised and tested for their ability to form DNA-protein complexes. Labelled G/CBE oligonucleotides specifically bound complex I proteins while CME specifically bound complex II and III proteins (fig 4.1). This confirmed that the G/CBE and CME are distinct DNA elements. An expression library could therefore be screened effectively with the CME oligonucleotide to identify complex II or III proteins without the possibility of isolating genes encoding the CCAAT-binding proteins.

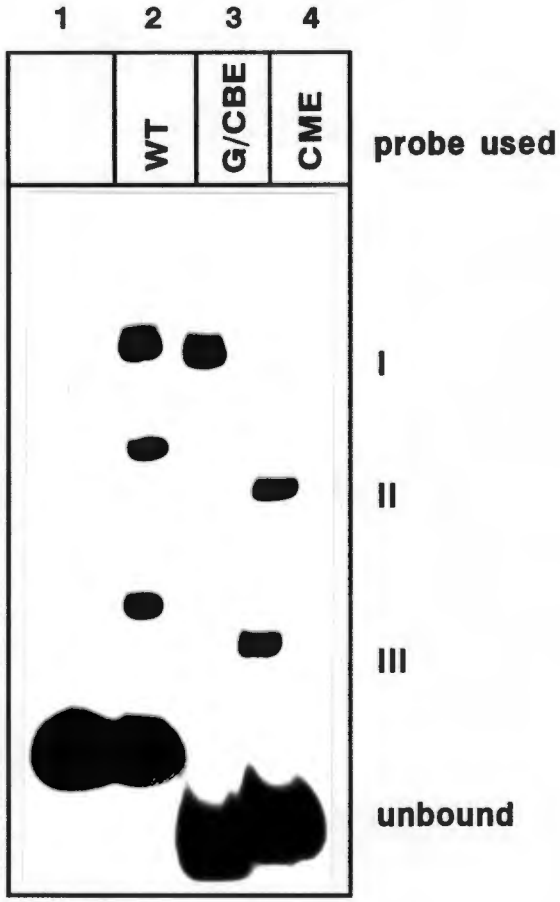


Fig 4.1 Checking of G/CBE and CME oligonucleotide probes in protein-binding assays. SVWI-38 nuclear extracts were incubated with labelled probes; lane 1, -107 to -60 bp wild type (WT) fragment of the COL1A2 promoter (no nuclear proteins added), lane 2, nuclear proteins incubated with WT fragment, lane 3, 29bp G/CBE containing fragment released from plasmid pNJ-29 incubated with nuclear proteins and lane 4, 32bp CME containing fragment released from plasmid pNJ-32 incubated with nuclear proteins. 4μg of nuclear proteins were incubated with the probes, electrophoresed on non-denaturing 5% polyacrylamide gels and exposed to X-ray film for 16 hours.

Since screening a phage library with a concatemerised probe greatly improves the detection of positive clones (Vinson et al, 1988), a CME oligonucleotide that allowed for concatemerisation of the DNA-binding site was generated (fig 4.2.A). Annealed oligonucleotides were concatemerised as described in section 6.10.1 and observed as ladders on 1% agarose gels (fig 4.2.B). The concatemerised oligonucleotides were radioactively labelled by nick translation for use in screening a λgt11 expression library. Initial results using nick translated probes were not successful and other labelling techniques were explored. The polymerase chain reaction (PCR) was subsequently used to prepare probes with high specific activities. Concatemerised CME was cloned into the vector pUC19 and used as a template in PCR. A Clontech human heart λgt11 library was screened with the radioactively labelled concatemerised probe. The choice of library is crucial in ensuring that the proteins of interest are expressed in the tissue/cell type used to make the library. The choice of screening a human heart library was based on the following;

(1) A human library had to be used since the proteins of interest are human α2(1) procollagen promoter-binding proteins.
(2) A previous study showed that although heart tissue contained type I collagen, it is expressed at intermediary levels. Heart tissue therefore contain cells that synthesise type I collagen and others that do not.

Screening a heart cDNA expression library would therefore enable the identification of genes involved the possible inactivation or activation of the collagen genes (ie. complex II or III expressing).

A

```
5'-AATTCGGAGGCCCTTTTGGAGG-3'
3'-GCCTCCGGGAAAACCTCCTTAA-5'
```

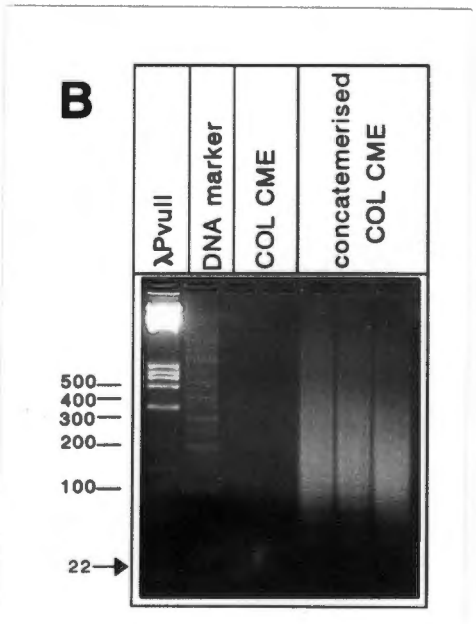


Fig 4.2 Concatemerisation of the CME oligonucleotides. **(A)** An oligonucleotide containing the complex II/III binding site (CME) was generated. It contained 5'-AATT overhangs that allowed sticky end ligations (section 6.10.1). **(B)** Detection of concatemerised CME oligonucletides by electrophoresis on 1% agarose gels. The lane marked COL CME contain the 22bp COL CME oligomer shown in (A).

4.2.2 Screening of expression libraries

4.2.2.1 Screening of a heart λgt11 expression library

The heart λgt11 library was screened according to the modified method of Vinson et al, (1988). Near-confluent plates (150 mm diameter) were used in an initial screen to maximise the chance of detecting potential positives. Each plate in the initial screening contained approximately 25,000-30,000 plaques. A total of 10 plates were screened at a time and the primary screening process was performed several times to allow for maximal positive detection. The plates containing plaques were overlayed with nitrocellulose membranes for 6-18 hours at 37°C. After marking, removing and drying the filters, the proteins bound to the filters were denatured/renatured as described in section 6.10.3. Filters were blocked with 5% fat-free milk powder, after which they were probed with radioactively labelled concatemerised CME probe. To identify putative positive phage plaques, autoradiographs were aligned with the original plates. Few putative positive plaques were detected in the first screen (fig 4.3.A). This was expected, since nuclear proteins represent less than 0.01% of total cellular proteins. Putative positive plaques from an initial screen were cored out, the phage eluted and subjected to another round of screening so as to; (i) exclude false positives and (ii) to enrich for potential positive clones by several rounds of purification (fig 4.3.B). This purification procedure was repeated four times. In the third screening, all plaques from a single plaque in the second round provided positive signals (fig 3.4.C). As confirmation, a fourth screen was done, showing 100% of the plaques expressing a protein that binds the CME (fig 4.3.D). The radioactive signals provided by these positives were very strong and autoradiographs could usually be developed within one hour of exposing the filters to X-ray film. A number of initially positive clones proved to be false on subsequent screening and were discarded. Further analyisis was performed on clones that remained positive after four screens and two of these proved to be informative. The significance of the others will be addressed in the discussion.

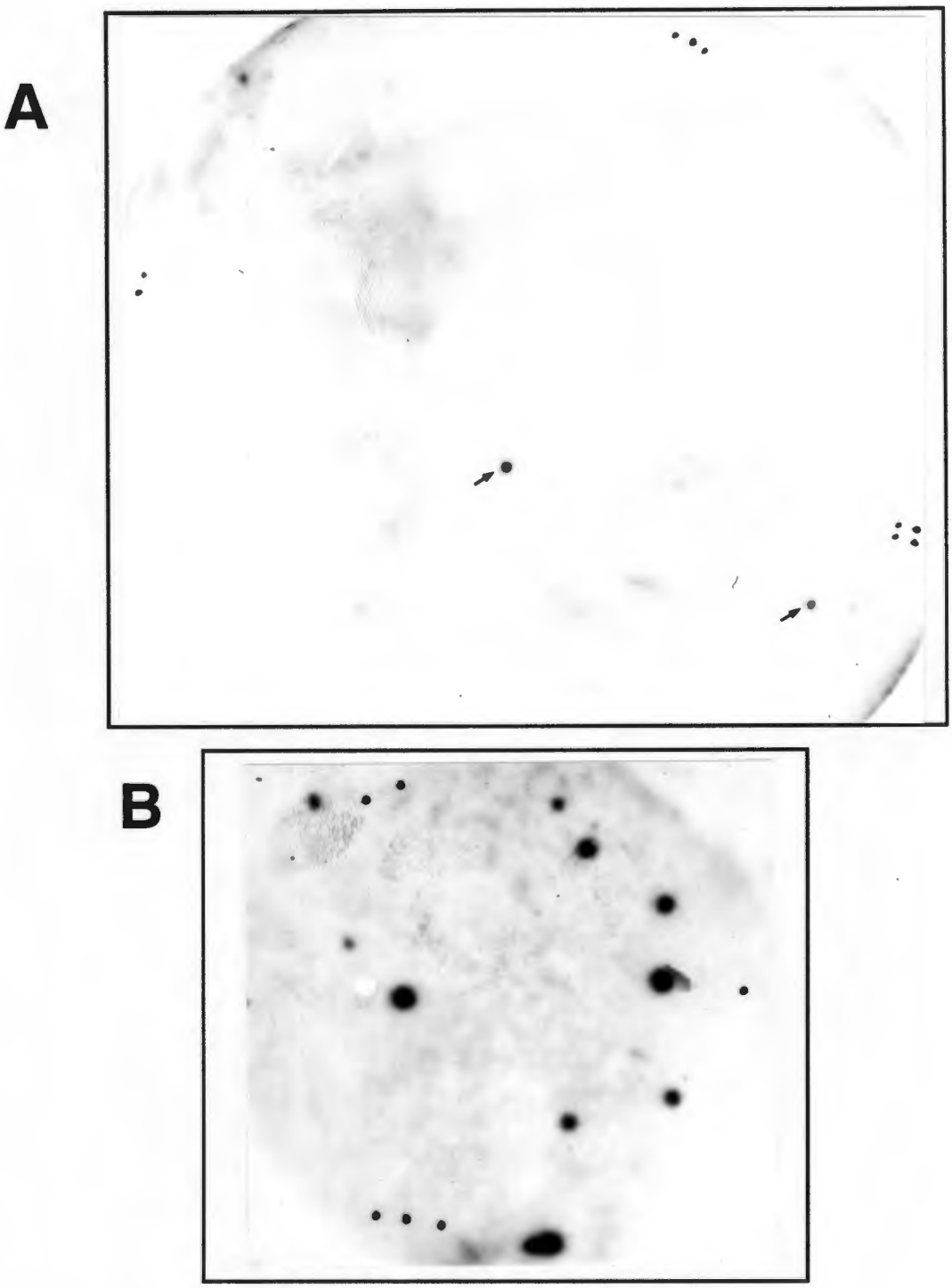


Fig 4.3 Screening of a λgt11 heart cDNA library. **(A)** Primary screen: 30,000 pfu on 150 mm plates were overlayed with IPTG-treated nitrocellulose membranes for 6 to 16 hours. Proteins immobilised on the membranes were screened with labelled concatemerised CME probes as described in section 6.10.3. Membranes were washed and exposed to X-ray film for 16 hours. Potential positives are indicated by the arrows. Plaques at and in the vicinity of a positive signal were cored out, the phage eluted and subjected to a second screening procedure. **(B)** Secondary screen: Phage eluted after a primary screen was replated and rescreened in order to isolate single clones for purification.

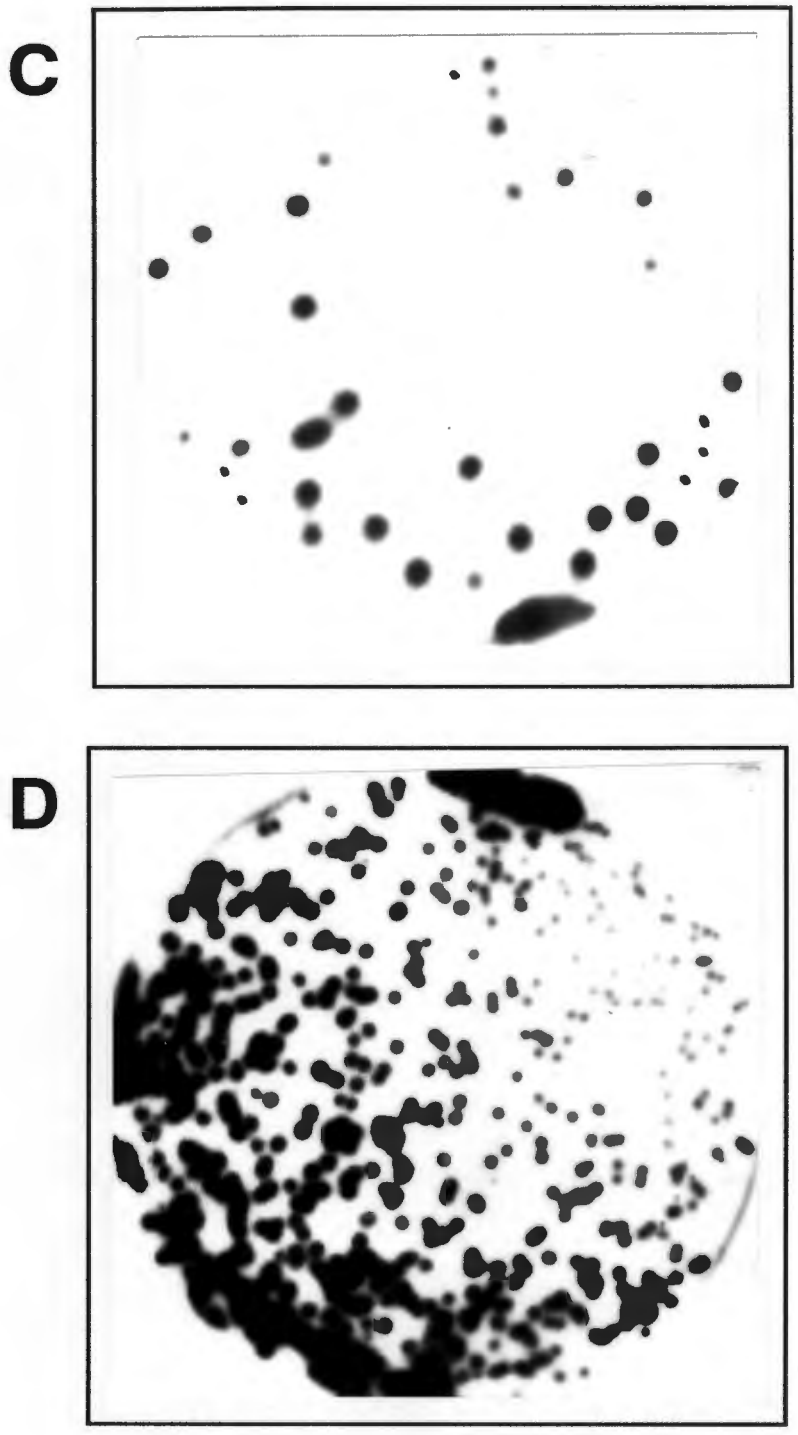


Fig 4.3 continued...
(C) Third screen: A single plaque from a secondary screen was cored out, the phage eluted and subjected to a third round of screening. At this stage all plaques were usually positive. **(D)** Fourth screen: A final round of plating and rescreening was included to confirm that a recombinant clone has been purified to homogeneity.

4.2.2.2 Controls for screening with the concatemerised CME

Due to the likelihood of detecting false postives when screening expression libraries, suitable controls are required. One such control was to screen potential positives with an oligonucleotide that does not bind the protein of interest. The -57 to -43 region (oligo A) of the α2(1) procollagen gene had been shown previously not to be involved in DNA-protein interactions (Smith, 1989). It was therefore used as a probe to detect non-specific binding. Putative positive phages were spotted in duplicate on a lawn of Y1090 cells immersed in top agarose on agar plates. After the appearance of plaques (about 3-4 hours at 42°C) the plates were overlayed with nitrocellulose membranes and incubated for a further 6 hours at 37°C. Filters were removed, cut in half and processed as described in section 6.10.5. No binding was detected with the non-specific DNA probe (oligo A), while the concatemerised CME probe bound all the potential positives spotted on the lawn of cells (fig 4.4.A).

To confirm that the labelled concatemerised CME probe bound specifically to the positive clones and not non-specifically to any protein immobilised on the nitrocellulose membranes, the ability of the probe to bind clones expressing known proteins was assessed. These proteins included chicken ovalbumin and β-galactosidase expressed by λovalbumin and λgt11 respectively. A phage which on initial screening proved to be negative for binding to the CME probe (λneg) was also included. Phages expressing these proteins were spotted on a lawn of Y1090 cells as previously described and filters probed with the concatemerised CME. While the probe bound putative positive clones, λ2, λ3 and λ4, no binding was detected on chicken ovalbumin, β-galactosidase and the λneg expressing phages as observed by the unreacted zones of clearing. The clone marked λ1 reacted positively with the labelled probe in the first screen but on subsequent screening proved to be negative for CME binding activity (fig 4.4.B). These screening results

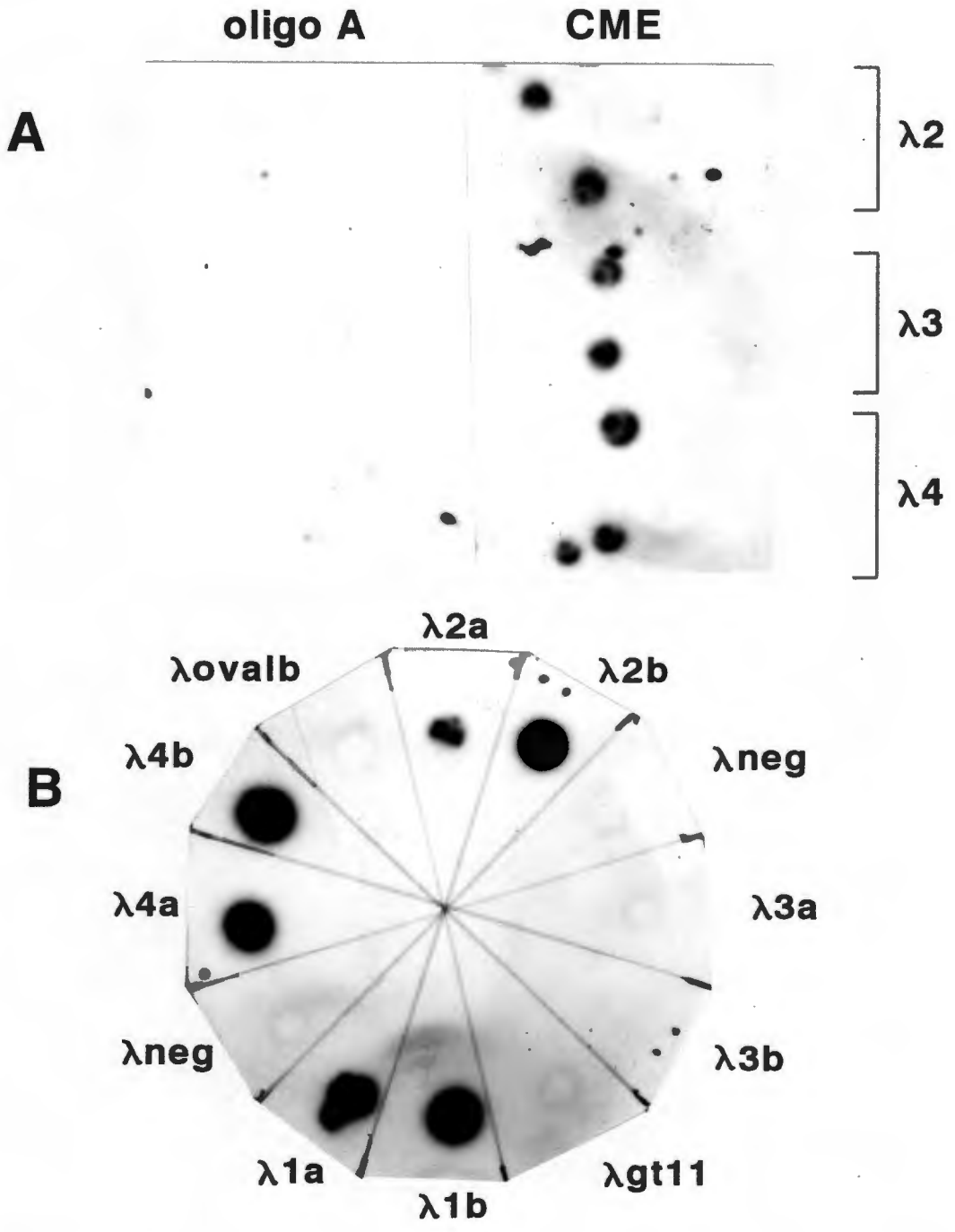


Fig 4.4 Controls for screening with the concatemerised CME. **(A)** Phage from putative positives were spotted in duplicate on a lawn of Y1090 cells on luria agar plates. After appearance of plaques, plates were overlayed with nitrocellulose membranes for 6 hours at 37°C. The filters were cut in half, one half was probed with a labelled concatemerised CME probe and the other half was probed with oligo A (-57 to -43 of α2(1) procollagen gene, (shown previously not to form any DNA-protein complexes). **(B)** Phages, λgt11 (expressing β-galactosidase), chicken λovalbumin, λnegative and putative positives, λ1, λ2, λ3 λ4 (a and b represent duplicates) were spotted on a lawn of Y1090 cells on luria agar plates as described in (A). Filters containing immobilised proteins were screened with labelled concatemerised CME as described in section 6.10.5. Membranes were exposed to X-ray film for 16 hours.

therefore indentified three potential clones expressing recombinant proteins which interact with the CME in the human α2(1) procollagen promoter. Subsequent characterisation of the λ3 proved to be difficult as lysogens expressing the fusion protein could not be isolated, hence the experiments described below will concentrate on the characterisation of λ2 and λ4.

4.2.3 Detection of lysogens harboring λgt11 recombinant phages.

To obtain direct evidence that the positive recombinant phages isolated express CME-binding β-galactosidase fusion proteins, lysogens harboring the recombinant phage were prepared. Y1090 cells containing the recombinant phages were streaked onto each of two plates. One plate was incubated at 32°C and the other at 42°C. This was done to test for temperature-sensitive growth as clones growing at 32°C but not at 42°C represent lysogens. Representitive recombinants incubated at these temperatures are depicted in fig 4.5. While all colonies grew at 32°C, eight failed to grow at 42°C. Crude proteins extracts were prepared from the lysogens as described in section 6.10.6. The protein extracts were stored at -70°C for subsequent use in characterisation of DNA-binding activity and SDS-polyacrylamide gel elecrophoresis.

4.2.3.1 South-Western Blotting

Additional evidence that the recombinant proteins encoded by the putative positive clones had the capacity to bind the CME in a sequence-specific manner was provided by South-western blotting using the labelled concatemerised CME as a probe. Lysogens of the recombinant phages were induced with IPTG for 1, 2 or 3 hours before fusion proteins were extracted. Proteins were fractionated on a 7% polyacrylamide SDS gel and transfered to nitrocellulose membranes. The immobilised proteins were denatured/renatured as described for the screening

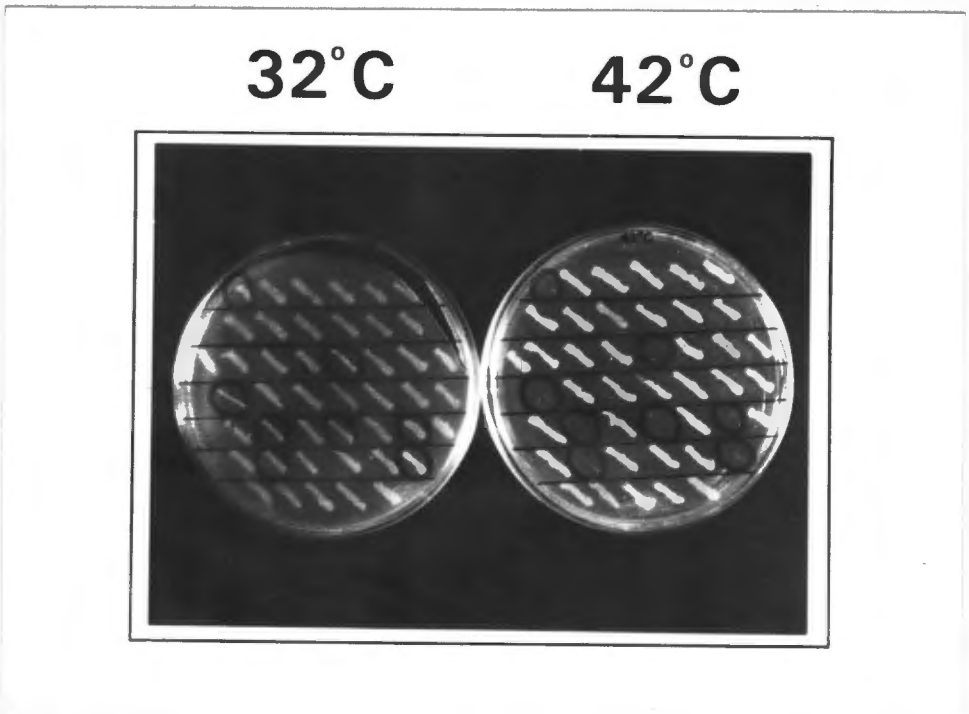


Fig 4.5 Detection of lysogens harboring λgt11 recombinant phages. Putative positives were plated on duplicate luria agar plates. One plate was incubated at 32°C and the other at 42°C. Clones growing at 32°C but not at 42°C represent lysogens. Lysogens are indicated by the circles.

protocol (section 6.10.3). The processed filters were blocked and probed with labelled concatemerised CME. Crude extracts from the two positive clones, λ2 and λ4, contained proteins which were IPTG inducible and able to bind to the probe (fig 4.6.A and B). Extracts from lysogens not induced with IPTG contained no detectable binding activity. An increase in binding of the probe to proteins was observed with increasing time of IPTG induction, indicating that these proteins were fusion products of the lac Z gene. The λ2 extract contained a fusion protein with an apparent molecular size 139 ±8 kD (fig 4.6.A), while the λ4 extract contained a fusion protein of approximately 153 ±10 kD (fig 4.6.B). Since the β-galactosidase segment of the fusion protein is approximately 116kD, the cDNAs encode proteins with molecular weights of approximately 23 ±8 kD and 37 ±10 kD respectively. Lysogens containing λgt11 and λnegative had no DNA binding activity even after IPTG induction. A SVWI-38 nuclear extract was used to check binding of the probe to the filter (fig 4.6, lane marked SV). The observed binding pattern in SVWI-38 nuclear extracts are similar to those presented by Collins, (1993).

4.2.3.2 Western Blotting

The filters used in South-western blotting were stripped of radioacitivity and reprobed with an anti-β-galactosidase antibody (section 6.10.3). The anti-β-gal antibody detected a number of bands (fig 4.7). In λ2 and λ4 lysogen extracts, the slowest migrating proteins corresponded exactly in size to that seen in South-western blots (fig 4.7A and B). This was observed by overlaying autoradiographs of South-western blots onto the Western blots. The appearance of several bands of lower molecular weight is probably due to degradation of the larger proteins. Similar degradation patterns of fusion proteins have been reported by Klemsz et al, (1990). The lanes containing proteins from the λgt11 and λneg lysogens contained predominantly the 116 kD β-galactosidase band and some degradation products.

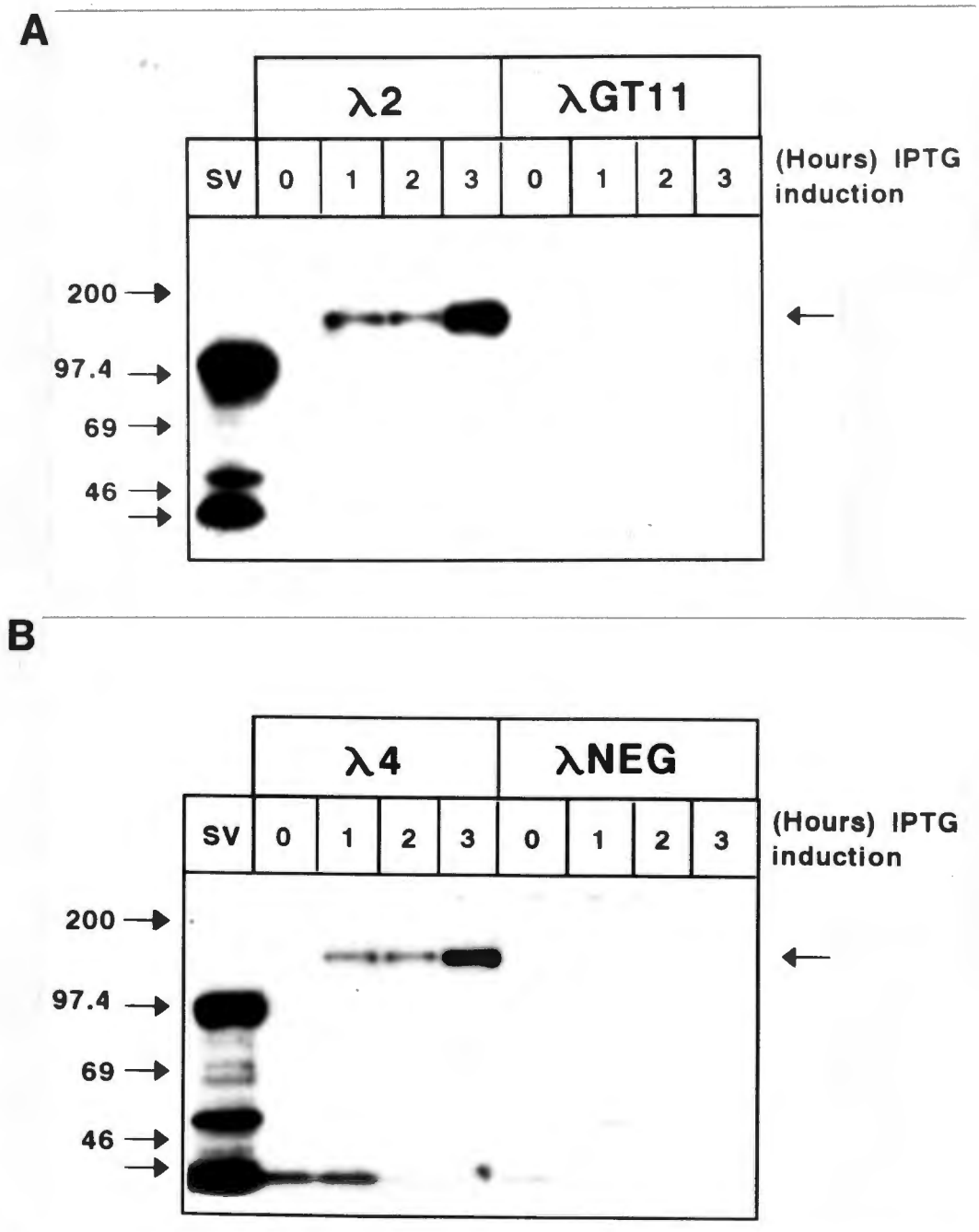


Fig 4.6 South Western blots of proteins isolated from λ2 and λ4 lysogens. (A) λ2 and λgt11 lysogens were induced with IPTG for the indicated times. After induction, crude proteins were seperated by SDS-PAGE and transfered to nitrocellulose membranes as described in section 6.5.2. Nitrocellulose filters were probed with labelled concatemerised CME probes. The lane marked SV represent crude nuclear proteins from SVWI-38 cells. **(B)** Proteins from λ4 and λnegative lysogens were separated on SDS-PAGE and probed as described in (A). The approximate molecular weight of the λ2 fusion protein is 139 ±8 kDand that of λ4 is 153 ±10 kD. Standard deviations were obtained by calculating the approximate molecular weights from three seperate South Western blots.

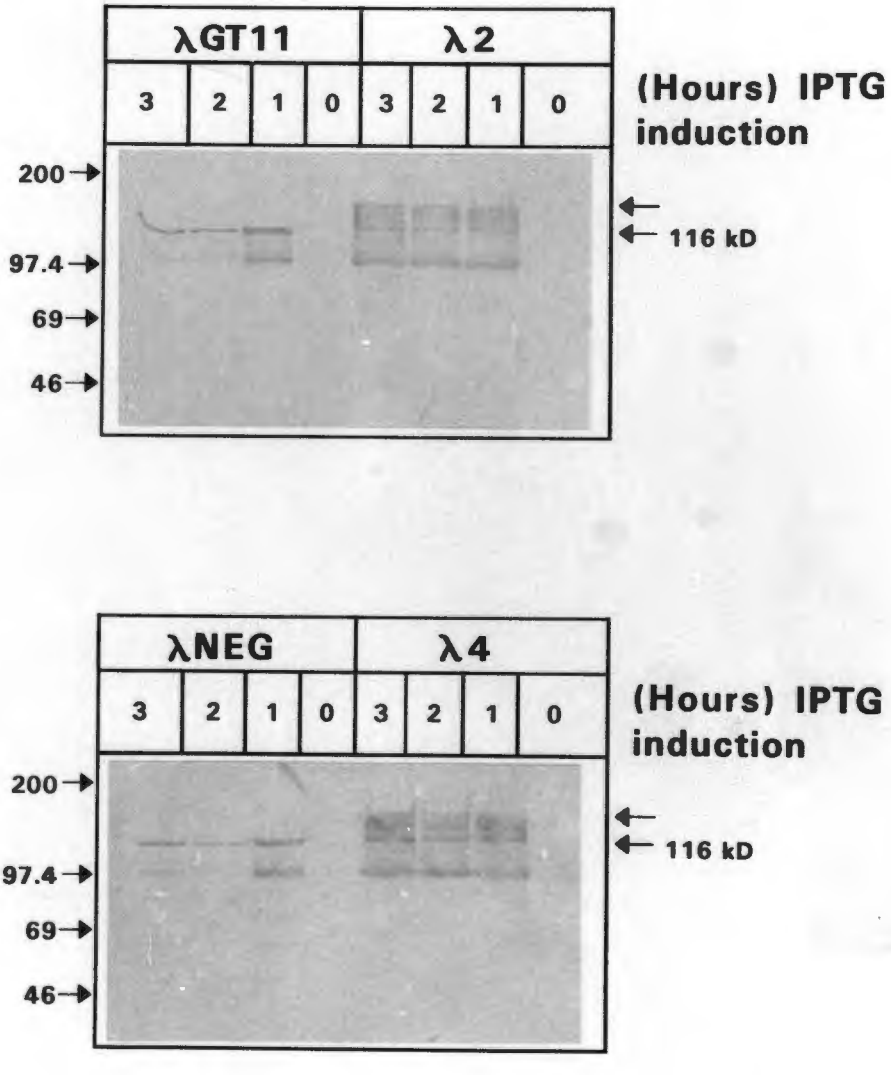


Fig 4.7 Western blot analysis using anti-β-galactosidase antibodies. The filters used in fig 4.6 were stripped of radioactivities and reprobed with an anti β-galactosidase antibody . A secondary antibody was used to visualise β-galactosidase fusion proteins as described in section 6.10.7. **(A)** Anti β-galactosidase antibodies detected the presence of several cross reacting bands in λ2 lysogen protein extracts. The slowest migrating band corresponding to that observed in South western blots is indicated by an arrow. Lysogens harboring λgt11, shows expression of the 116 kD β-galactosidase. **(B)** λ4 and λnegative lysogen protein extracts probed in the same way as described in (A). The band corresponding to the fusion protein in λ4 extracts is indicated by a arrow. λneg extracts also expressed β-galactosidase.

The results of the south-western and western blots provided additional evidence that at least two of the original putative positive phages, λ2 and λ4, express proteins which bind the CME of the human α2(1) procollagen promoter. To confirm this for one of the clones, λ2 lysogen protein extracts were used in electrophoretic mobility shift assays.

4.2.3.3 EMSA analysis of the λ2 fusion proteins

The COL1A2 proximal promoter fragment (-107 to -60) was used in EMSA's to confirm the binding activity of the β-galactosidase fusion proteins. This probe contains a single binding site for the CME-binding proteins. 30 μg of crude lysogen protein extract was used in each reaction. A DNA-protein complex with was detected in both λ2 protein extracts (fig 4.8.A). This migration pattern was slightly higher than that observed for complex I proteins and is as a result of the fusion of β-galactosidase and the unknown protein segment expressed by the recombinant cDNA. It should also be noted that these DNA-protein complexes where detected after exposure of the gels for three days to X-ray film. The DNA-binding activity detected in extracts from λ2 lysogens was not found in protein extracts from uninfected Y1090 cells.

Specificity for the DNA-binding activity was assayed by using an unlabelled CME oligonucleotide as a competitor for binding. Increasing amounts of the CME oligonucleotide specifically competed out binding of the DNA-binding protein complex found with λ2 lysogen extracts. At 200 molar excess of unlabelled CME, practically all DNA-protein interaction was competed out (fig 4.8.B). EMSA's with the unlabelled G/CBE oligonucleotide had no effect on binding of the fusion protein (fig 4.8.C). These results therefore confirmed that the clone obtained in screening a human heart λgt11 expression library, expressed a fusion protein which specifically bind the CME of the COL1A2 promoter.

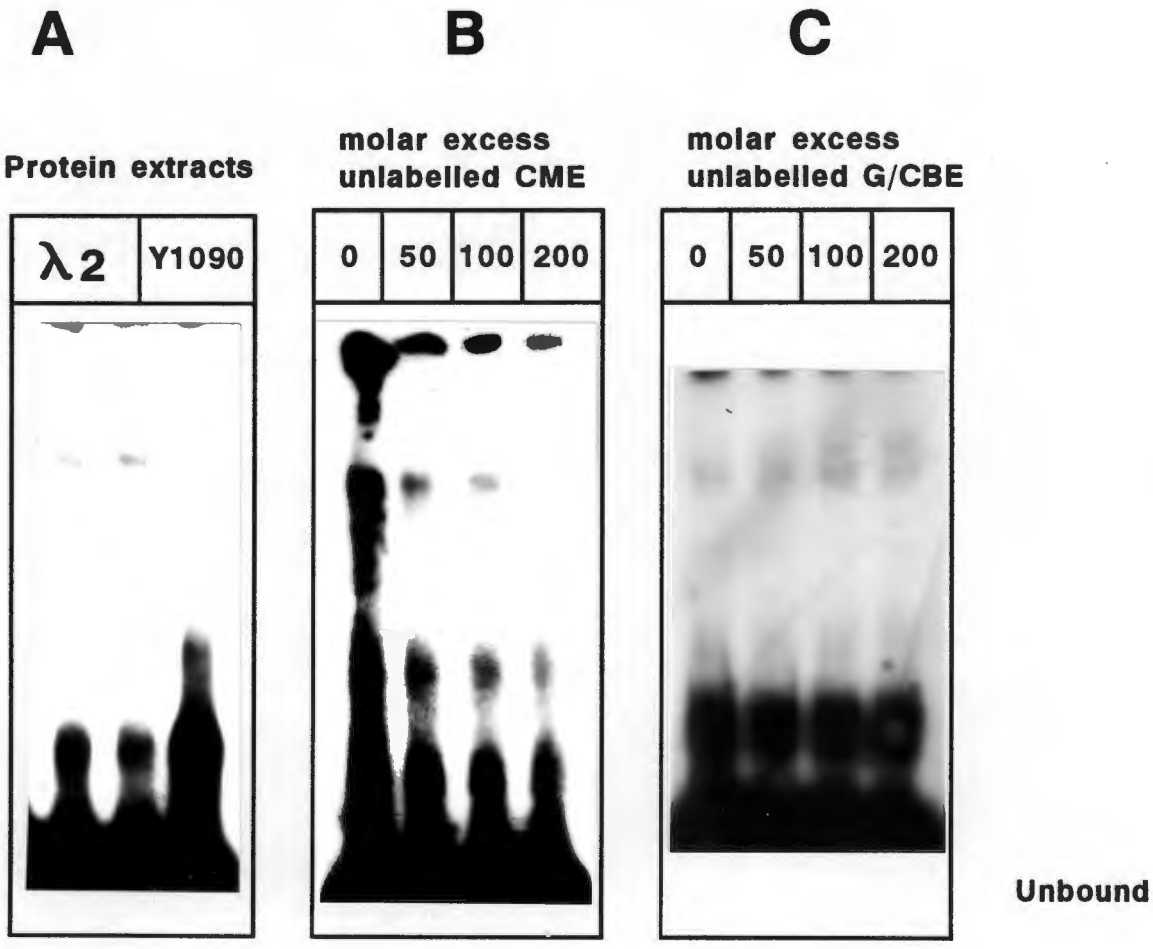


Fig 4.8 DNA-binding specificity of the λ2 fusion proteins. **(A)** 30μg of crude λ2 lysogen proteins (in duplicate) and Y1090 proteins were incubated with a labelled COL1A2 proximal promoter fragment (-107 to -60). 2μg poly dIdC.poly dIdC was included in the reaction prior to the addition of the probe. DNA-protein complexes were fractionated on non-denaturing 5% polyacrylamide gels. Dried gels were exposed to X-ray film for 3-5 days. **(B)** Increasing molar excess of unlabelled CME was used as a competitor for formation of the DNA-protein complex in λ2 lysogen extracts. EMSA's were performed as previously described, except that competitor DNA was incubated with protein extracts prior to the addition of probe. Dried gels were exposed to X-ray film for 3 days. **(C)** Increasing molar excess of unlabelled G/CBE used as a competitor as described in (B).

Having established that the clone, λ2 expressed a DNA-binding protein, further characterisation of the cDNA expressing the putative DNA-binding protein was performed. This included determining the size of the cDNA, followed by cloning into pUC 19 and subsequent sequence analysis.

4.2.4 Analysis of the λ2 cDNA inserts

The polymerase chain reaction was used to determine the approximate cDNA insert size of clone λ2. λgt11 primers flanking the cDNA inserts were used in the reactions. These primers were 24 mers and flanked the EcoR1 site into which the cDNA's had been cloned (fig 4.9.A). A cDNA insert of approximately 1kb was obtained, while PCR of λgt11 and λneg DNA as negative controls provided no PCR products (fig 4.9.B).

The λ2 PCR product was blunt ended using klenow polymerase and cloned into pUC19 as described in section 6.2.4. Inserts were checked by digestion of the cloned plasmid with EcoR1 or Kpn1/BamH1 double digestion, both providing the correct size of approximately 1kb (fig 4.10). The cloned insert was sequenced using the M13 universal primers and internal primers designed to complete the sequencing by DNA automated sequencing. The complete sequence is shown in fig 4.11. Computer-assisted analysis of the sequence was performed using the University of Wisconsin Genetics Computer Group Sequence Analysis Software Package (Devereux et al, 1984, 1989). The *FastA* program was used to search the European Molecular Biology Laboratory (EMBL) and GenBank nucleic acid databases for similarity between λ2 and sequences in the databases. No similarity was obtained with any of the sequences in the database, indicating a possibility that the λ2 encodes an as yet undescribed DNA-binding protein.

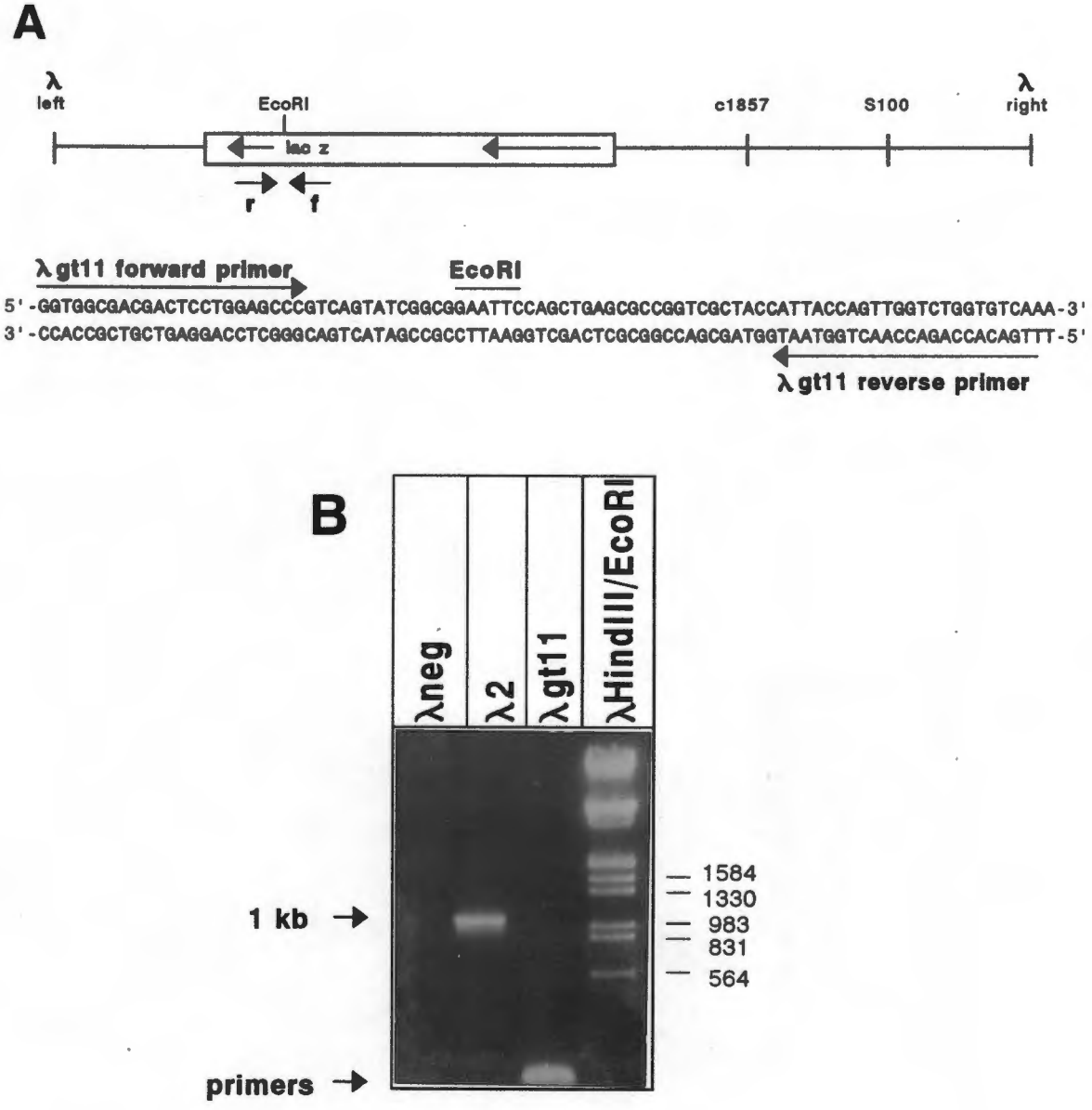


Fig 4.9 Size analysis of λ2 cDNA insert. λgt11 primers that flank cDNA inserts in λgt11 **(A)** were used in PCR reactions. PCR conditions for optimal product formation was set at 95°C,1min; 60°C,1min; 72°C,1min for 30 cycles. **(B)** 10μl of a 50μl PCR mix was loaded on 1% agarose gels and products were separated at 45 mA constant current. The insert size for λ2 is approximately 1 kb. λgt11 and λneg was used as negative controls.

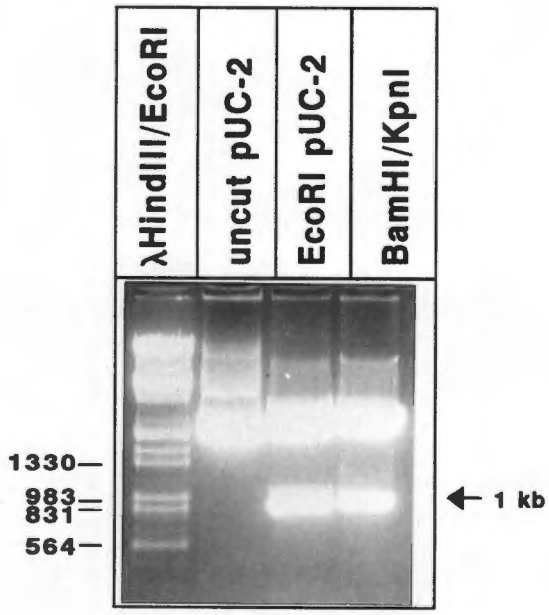


Fig 4.10 Cloning of the λ2 cDNA insert in pUC19. The PCR product of clone λ2 was blunt ended using Klenow polymerase. The blunt ended product was cloned in a SmaI digested pUC19 vector, the resultant plasmid (pUC-2) was cut with the restriction enzymes, EcoR1, and a BamH1/Kpn1 double digest and electrophoresed on 1% agarose gels. A product of approximately 1 kb was detected.

```
5'-
GAATTCGGGGCCGGGACCATTTTTAAATGTCAGTTGAAAATTATGGCTGTACTATTGCTT   60
  E  F  G  A  G  T  I  F  K  C  Q  L  K  I  M  A  V  L  L  L

AAACAAAACTGGAACTGTTGTTCGAAATCCATAGCCAATACTTTTCACGCCAACCTGTGT  120
  K  Q  N  W  N  C  C  S  K  S  I  A  N  T  F  H  A  N  L  C

ACTGAACATAGTAGATTGACATCTAATTCAAGATTACAACATCTGTTACATTCTAAGTGT  180
  T  E  H  S  R  L  T  S  N  S  R  L  Q  H  L  L  H  S  K  C

GTTCAGGCTTCTGAAGGTAAAGGGACACTGGATCCAGAAGCTATGGAACCAGCAGTTGAT  240
  V  Q  A  S  E  G  K  G  T  L  D  P  E  A  M  E  P  A  V  D

TTCTTGTATTTCCTGATTAACCTACTTGTAAACTTGAAAGCAAGACCTTGATTGCACCAA  300
  F  L  Y  F  L  I  N  L  T  C  K  L  E  S  K  T  L  I  A  P

CAGGTCCAGAGTATGAGTGCAAGCAAAGCAGAACTCTCATGCGTGACCTGAGCAGACAGG  360
  T  G  P  D  M  S  A  S  K  A  E  L  S  C  V  T  *

CTGGTATTTAACAGGTGCCTCGTGTTGACATTACGCTGCCTTAATGTAACACAGTCTGGC  420

AGTTGCTAAATTTGTGTTCCCATTTTAAATTGACCAATTTTGGGGTGTGACACTTTTGTG  480

CGGTTGAATTGGGAGAATGAAGATAAGTTAATTTACCTGTCCCAGGATCAAAAGAAGCCC  540

TAGAAAAGAAGCAGTAATCTACCTCCTGCCGATAACCTGTTTAAGATGACTCCAGCCAGA  600

AACACCCGCCGTTTCCACTCCTATTGGTCCAATTCCATGTGGCTGACTAAGTCCAATTTT  660

TTTTCTGAACAAAAGAAAGGCTAAACACTATGTAAATGTGAATGGAAACTTGGAAAATAC  720

TCGTTTTTATTTAACTACAAAAACTTTTGCTCTGTTTATCAGGAAATCCATACTTTATTT  780

TGTAATTAACAGACAAGCCTGTGTGGATGATTTTTTTGAACTTGGTAGTTCATAAAGGTT  840

TACAGTGAATAAAAGGACTATCATCTTGAGGATAGCAATATCAAAAGAATTCAGTAGTTA  900

CTGTCTGTTTAGGAATATAAGGGATTAAGATATCATATGGGTCAGGTTACCCGAATTC -3'  958
```

Fig 4.11 Nucleic acid and deduced amino acid sequence of clone λ2. The calculated size of the peptide encoded by the open reading frame is 12.5 kD. The peptide was also found to contain multiple consensus phosphorylation sites (see fig 4.12) for the cAMP-dependent protein kinase and protein kinase C (Kennely and Krebs, 1991).

Translation of the DNA sequence was performed and this showed an open reading frame of 116 amino acids with a calculated peptide molecular weight of 12.5kD. Discrepancies are often found between the calculated molecular weight of a protein and the apparent molecular weight observed by SDS-PAGE (23 ±8 kD, fig 4.6.A). A search of the protein databank by computation at the National Centre for Biotechnology Information (NCBI) using the blast network service, provided very little similarities between the 116 amino acid peptide with other proteins. Some similarity however, was obtained with the sex-determining region Y (SRY) protein (fig 4.12.A). Although the similarity is minimal, it is interesting to note that SRY is a transcriptional activator which regulates the genetic switch in male development responsible for initiating male sex determination. It also contains a HMG box which recoginises DNA by partial interaction in the minor groove (Payen and Cotinot, 1993). The murine (m)SRY and human (h)SRY have differences in DNA binding activity in that (m)SRY shows more extensive major-groove contacts with DNA and it has a higher sequence specificity than (h)SRY (Giese et al, 1994).

The peptide also contains multiple potential consensus phosphorylation sites for the cAMP-dependent protein kinase and protein kinase C (fig 4.12.B). The consensus sequences recognised by these kinases are listed in table 4.2.

Table 4.2 Consensus sequences most frequently recognised by cAMP-dependent protein kinase and protein kinase C.

Protein kinase	Consensus sequence
cAMP-PK	R-R/K-X-S*/T* > R-X_2-S*/T* = R-X-S*/T*
PKC	(R/K_{1-3},X_{2-0})-S*/T*-(X_{2-0},R/K_{1-3}) > S*/T*-(X_{2-0}, R/K_{1-3}) ≥ (R/K_{1-3}, X_{2-0})-S*/T*

Modified from Kennely and Krebs (1991)
Note: The phospho-acceptor group is denoted by an asterisk.
Where two amino acids functions interchangeably, both are listed with a slash (/) separating them.
Sequences judged to be neutral are denoted by an X.

Although these potential phosphorylation sites are found in the peptide, this does not imply that the protein is necessarily phosphorylated as secondary and tertiary structures can play a significant role in substrate recognition. These findings can only be confirmed when the protein has been purified and on secondary structure predictions.

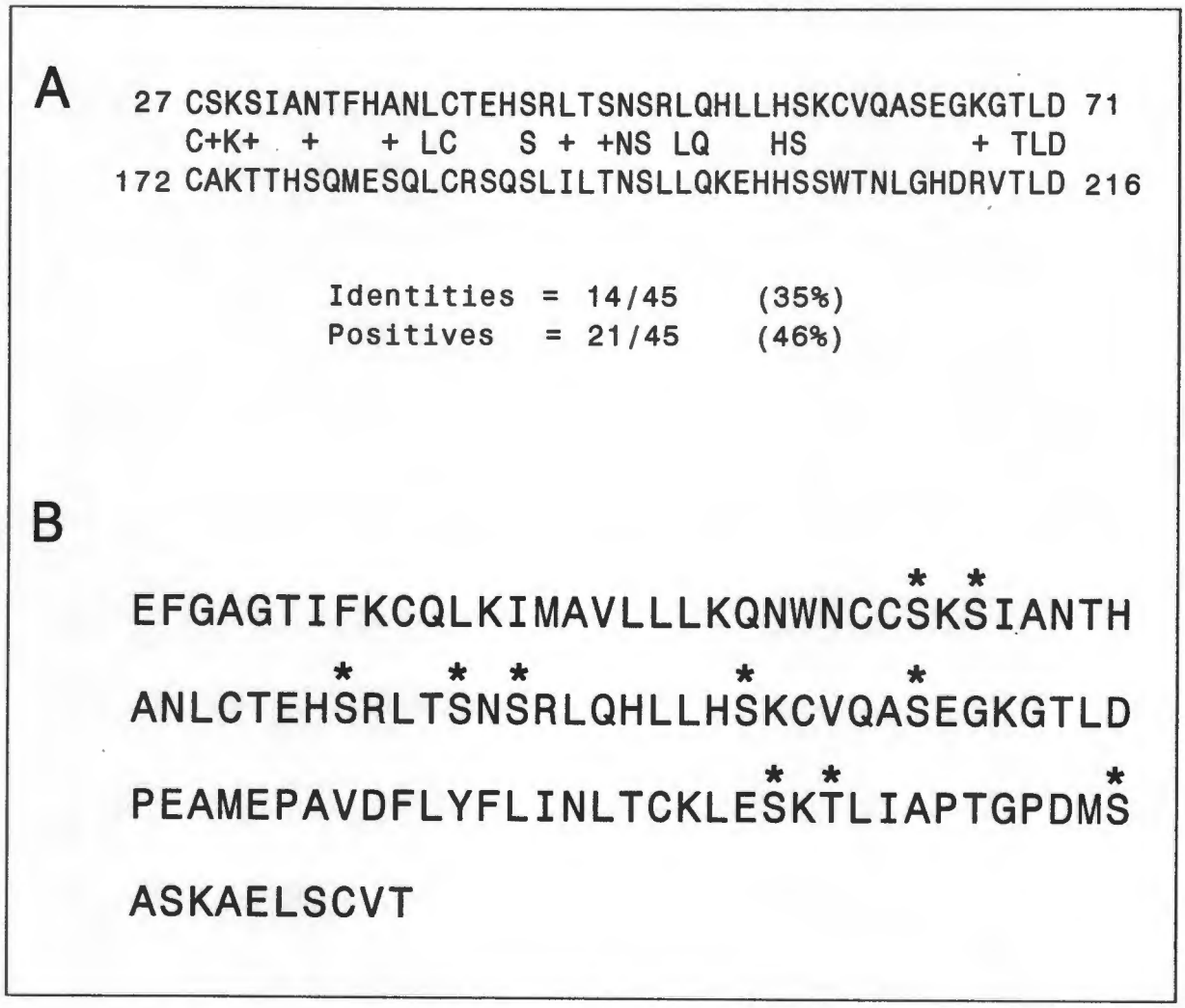


Fig 4.12 Sequence similarities and potential phosphorylation sites. **(A)** Computation performed using the NCBI blast network service provided similarities between the 116 amino acid peptide and the sex-determining region Y (SRY) protein. **(B)** Potential phosphorylation sites (indicated by the asterisks) using the phosphorylation consensus sequences listed in table 4.2.

4.2.5 Expression of λ2 mRNA in different cell lines

Northern blot analysis was performed to assess the expression of the λ2 encoding gene in eukaryotic cell lines. Cytoplasmic RNA was extracted from cell lines which showed varying levels of α2(1) collagen gene expression. The λ2 cDNA insert was released from pUC-2, radioactively labelled by random priming (section 6.6.2.2) and used as a probe to detect the presence of λ2 message. A β-actin probe was used as a control for RNA loading. A mRNA transcript of appoximately 4 kb was detected in varying amounts in the different cell lines (fig 4.13.A). The cell lines where α2(1) collagen gene expression occurs (FG_o and HT1080) (section 2.2.2.2), contained low amounts of the transcripts. In cell lines where α2(1) collagen gene expression was not observed (MDA, MCF, ZR and LB), the amount of λ2 expression was enhanced by comparison to that in FG_o and HT1080 cells. Bar graphs of the expression of λ2 relative to that of β-actin in the different cell lines are shown in fig 4.13.B.

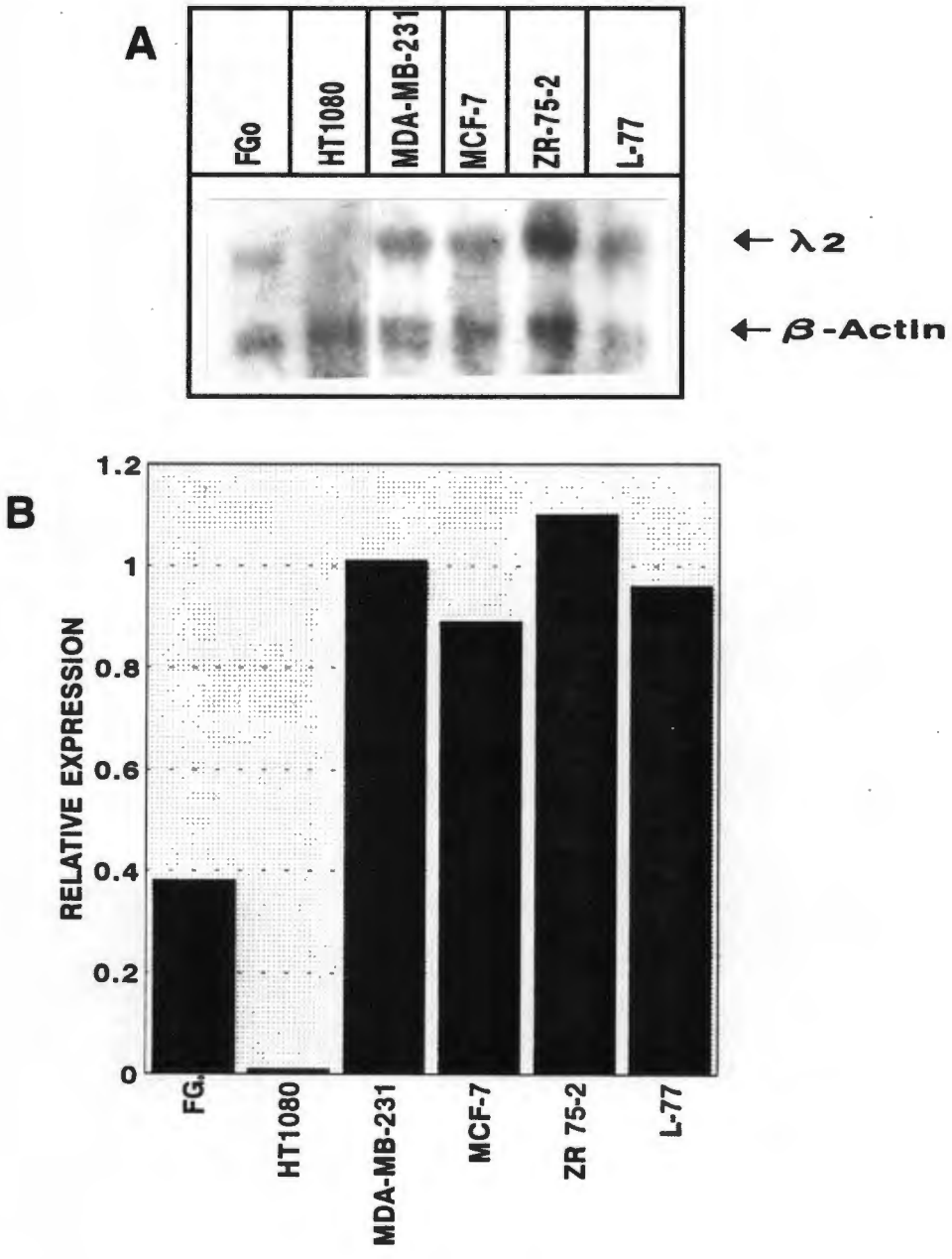


Fig 4.13 Expression of λ2 transcript in different cell lines. **(A)** Total RNA was extracted from the cell lines by the method of Chomczinsky and Sacchi, (1987). 5μg of total RNA was separated on 1% agarose formaldehyde gels, transferred to nylon membranes and probed with labelled λ2 and β-actin probes as described in section 6.6.2. After washing, the filters were exposed to X-ray film for 16 hours. **(B)** Densitometric scans of the autoradiograph in (A) indicating the ratio of λ2 expression relative to that of β-actin.

4.3 DISCUSSION

This study describes the isolation of cDNA clones coding for proteins which may in part be responsible for the cell- and species-specific expression of the human α2(1) procollagen gene. Screening of a cDNA expression library with the CME provided two clones, λ2 and λ4. The proteins encoded by these clones showed weak binding to the CME in EMSA's using crude protein extracts from lysogens. This weak binding can be attributed to the following: (i) A limitation of the screening strategy used is that it can result in the isolation of fusion proteins in lysogen extracts that bind the recognition site with low affinity or no detectable affinity when assayed in soluble fractions (Singh, 1993). (ii) The technique also relies on the expression of a functional DNA-binding domain in *E.coli* where post-translational modification of the fusion proteins does not occur as it would in eukaryotic cells. Chapter 3 reports that complex II and III proteins require phosphorylation for high affinity binding to their recognition site. This may explain the low affinity binding observed in EMSA's using *E.coli* lysogen protein extracts. The high affinity binding to the concatemerised CME probe detected on nitrocellulose membranes is as a result of denaturing and renaturing the proteins on the filters, and could allow the proteins to take up the correct conformation for DNA binding.

Although low affinity binding of λ2 fusion proteins was observed in EMSA's, competition experiments with unlabelled CME confirmed the presence of a sequence- specific DNA-binding protein in extracts from lysogens. The molecular size of the protein expressed by clone λ2 was approximately 23 ±8 kD and the cDNA size of 1kb represented a segment of the coding region of the gene based on the findings by Northern blot analysis indicating a transcript size of approximately 4 kb.

Since λ2 transcripts are found predominantly in cells where the α2(1) procollagen gene is not expressed, the expression of these transcripts may be controlled in a cell-specific manner. This cell-specific expression is consistent with work presented in chapter 2, showing cell-specific detection of the complex II proteins which bind the CME in the human α2(1) procollagen promoter. The implication of this is that λ2 expresses a peptide that could possibly be a component of the complex II protein(s). Although the DNA-sequence used to screen the library was able to bind both complex II and III proteins, it is more likely that the clone obtained encodes a component of complex II since λ2 transcript detection in the different cell lines correlates with the presence of complex II proteins and not that of complex III.

Translation of the DNA sequence provided a 116 amino acid peptide which shares 35% homology over 45 amino acids with the sex-determining region Y (SRY) protein, an activator of transcription (Payen and Cotinot, 1993). This is however not a significant similarity. Interestingly, recent work has shown that other members of the SRY-related family of genes, the SOX genes, play a role in the regulation of the type II collagen gene. SOX9, acts as an activator by interacting with an enhancer sequence located in the first intron of the α1(II) gene (Lefebvre et al, 1997). Members of the SOX family of proteins may therefore have regulatory effects on the other collagen genes. It is possible that the clone isolated in this study represent an as yet undescribed DNA binding protein. Using the consensus sequences for phosphorylation sites published by Kennely and Krebs (1991), a number of potential phosphorylation sites in the 116 amino acid peptide were noted. Identification of these phosphorylation sites are in agreement with a model depicting phosphorylation as a requirement for high affinity complex I, II and III DNA-protein interactions (section 3.2.5.2).

In initial attempts to isolate clones expressing COL1A2 promoter binding proteins, the -107 and -60 region of the α2(1) procollagen gene was used as a probe to screen a HepG2 λgt11 expression library. This fragment includes the inverted CCAAT box and the CME. Sequence analysis of a potential positive and a database search indicated 85% homology between the clone and the bacteriophage lambda CI protein. It is interesting to note that the one of the DNA binding sites

for the lambda CI protein is highly homologous to that of the DNA probe used to screen the library, especially around the CCAAT box in the COL1A2 promoter (fig 4.14).

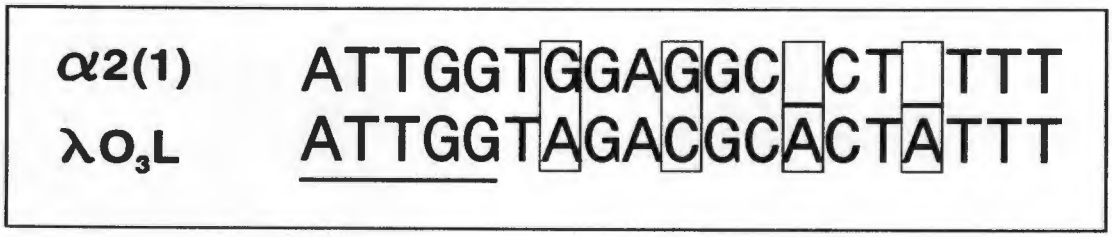


Fig 4.14 Sequence homology between COL1A2 and λO_3L. Sequence homology between the human α2(1) procollagen promoter and lambda operator 3 binding site. The bases not homologous or gaps are boxed. The CCAAT box is underlined.

The λCI protein binds three regions in the λ left and right operator regions. The homology between the COL1A2 promoter and the CI protein DNA binding sites is found in the third lambda operator box, λO_3L. A comparison of the two sequences showed a 4 bp mismatch over a 19 bp region. It is therefore possible that the λCI proteins have the ability to bind this region in α2(1) promoter, hence the detection of this clone. The manner in which λCI DNA came about in the cloning region of the λgt11 expression vector remains unexplained. It is possible that this may have arisen due to a rearrangement of the λgt11 DNA. A rearrangement of this nature could have resulted in the placement of the λCI gene in the vicinity of the lac Z promoter allowing the expression and subsequent detection of the protein. To substantiate the speculation of rearrangement events, several restriction enzymes were used to compare the digestion patterns of the clone isolated and wild type λgt11 cut with the same restriction enzymes. Digestion patterns between the clone (assuming it was λgt11 containg an insert) and wild type λgt11 were not consistent with the what was expected. These patterns implied differences in the cutting sites

between the clone and WT λgt11. It should be noted that the HepG2 library used initially was a Clontech library. The purity of certain Clontech libraries had been questioned at the time and it is also possible that the initial choice of library may have been problematic. The subsequent screening of a heart λgt11 library was therefore performed only after thorough checking of the library. Screening of this library revealed the clones discussed here.

It should be noted that the results presented in this chapter are preliminary and a precise conclusion as to the exact role of the clone identified can only be made after extensive analysis with a recombinant protein using both wild type and mutant sequences. The detection of these clones will enable the to identification of full length DNAs encoding the proteins that bind the proximal α2(1) procollagen promoter. This will allow a detailed analysis of their precise role in α2(1) procollagen gene expression.

5. CONCLUSION

CHAPTER 5:

CONCLUSION

The regulation of α2(1) procollagen gene expression has been associated with different trans-acting factor interactions within the proximal promoter region (-107 to -60). A α2(1) collagen expressing cell line (CT-1), formed two major DNA-protein complexes, complex I and III, while a cell line in which the gene is inhibited (SVWI-38), contained an additional DNA-protein complex, II (Parker et al, 1989, 1992). Preliminary functional data also showed that the DNA binding sequences of these proteins were essential for the expression of a proximal COL1A2 promoter construct (Collins et al, 1997). The differences in trans-acting factor binding could explain the differences in α2(1) procollagen promoter activities in the two cell lines and formed the basis of the investigation presented here.

5.1 Objectives

In an attempt to identify mechanisms in the transcriptional regulation of the human α2(1) procollagen gene, the objectives of this investigation were as follows:

1. To determine if the findings reported by Parker et al, (1992) were specific for the cell lines used in that particular study or whether cell differentiation in terms of procollagen gene expression could be associated with trans-acting factor interactions in a number of different cell lines.

2. Having established the above; to determine if the model proposed for cell-specific α2(1) procollagen gene expression, holds true in different species.

3. Since the human complex I proteins and the mouse CBF bound similar elements in their respective promoters; to determine their relationship with each other.

4. To study the effect of dephosphorylation and kinase inhibitors on trans-acting factor interations at the proximal promoter and subsequent expression of the α2(1) procollagen gene.

5. To identify the gene(s) coding for α2(1) procollagen promoter-binding proteins and to correlate these findings with the cell and species-specific regulation of the α2(1) procollagen gene.

5.2 Significant findings

The present study provides evidence that the association between α2(1) procollagen gene expression with trans-acting factor switching is a fairly general phenomenon. A range of cell lines was examined to determine α2(1) collagen chain synthesis, steady state mRNA levels and new transcript formation. The different levels of α2(1) collagen transcripts in the various cell lines was correlated with transcription factor interaction with the proximal α2(1) procollagen promoter. This clear correlation led to the proposal of a model for the cell-specific regulation of the α2(1) procollagen promoter. The model proposes that two DNA-protein complexes, I and III are associated with normal expression of the gene. Complex I proteins involves binding to the G/CBE (which includes the CCAAT box) and complex III proteins bind the CME. These two factors are possibly co-activators of α2(1) procollagen gene expression. In cells where the gene is inhibited, two scenarios may be responsible for the down regulation of the gene, (1) the absence of complex III proteins and (2) the presence of a different DNA-protein complex, complex II, is proposed as a repressor of α2(1) collagen synthesis. Both these events may be required in the inhibition of the gene. The fact that the complex III

proteins were not detected in cell extracts where the α2(1) procollagen gene is permanently inactive, could be due to the silencing of the gene(s) expressing complex III protein(s) or modification of the protein(s) in such a manner that it is prevented from interacting with it's recognition sequence. In cells where the α2(1) procollagen gene is transiently silenced, the presence of all three DNA-protein complexes, I, II and III were detected. It is possible that the relative amounts or the ratio of complex II to III proteins may be important in regulating expression of the gene. Where the amounts of complex II proteins exceed that of complex III, α2(1) collagen synthesis is inhibited (eg SVWI-38) and vice versa (eg CT-1). Surprisingly, a cell line (HT1080) where no α2(1) collagen chains or steady state mRNA were detected, contained complex I and III binding activity. Identification of newly formed RNA transcripts in this cell line by nuclear run-on transcription assays, showed that the gene was transcriptionally active and is therefore in line with the proposed model. This finding implies that post-trancriptional events are responsible for the absence of α2(1) collagen mRNA and polypeptide synthesis in HT1080 cells. A similar result was reported in HeLa cells, where the gene is active, but post-transcriptional modifications lead to absence of steady-state message and polypeptide synthesis (Furth et al, 1991).

The use of a G/CBE oligonucleotide which exclude the downstream GGAGG box, (GCCCTCCC**ATTGG**) in competiton experiments, indicated that the two binding elements, G/CBE and CME, are distinct but adjacent binding sites. This sequence competed out binding of complex I as effectively as the proximal α2(1) procollagen promoter sequence (-107 to -60). That these elements are distinct is supported by competition assays using the CME oligonucleotide (**GGAGG**CCCTTTT) which competed out complexes II and III binding and not that of complex I. These factors also bind the proximal promoter with similar affinities.

The validity of the above model was tested in two species, human and mouse. The results indicated that the proposed model does not account for differences in α2(1) procollagen gene expression in rodent cell lines. This may be due to a 3 bp mismatch within the CME motif between the human and rodent promoters. While the human promoter clearly forms two distinct DNA-protein complexes, the rodent promoter forms only one DNA-protein complex, the CCAAT-binding factor (CBF) in the -107 to -60 region. Interestingly, mouse nuclear extracts contained complex II proteins, but these however, are not able to interact with the mouse proximal α2(1) procollagen promoter. It is possible that complex II protein may have DNA binding activity at distal sites or that it has a regulatory function in the expression of other genes. Since the CME binding proteins are proposed to have a regulatory function, this raises the possibility that the absence of this element in rodent promoters could lead to species-specific differences in the proximal α2(1) procollagen promoter acitivity. Transfection assays of both the human and mouse proximal promoter-CAT constructs into CT-1 (human) and 3T3 (mouse) cells showed differences in transcriptional activities between two species. In CT-1 cells the human promoter-CAT construct was more active than the mouse promoter. Since the only difference in terms of trans-acting factor binding sites between the two promoters is the 3 bp mismatch in the CME, the observed transcriptional differences can be ascribed to differences in trans-acting factor binding. The inability of complex III to bind the mouse promoter transfected into CT-1 cells may be responsible for the decreased transcriptional activity observed. Expression of the human and mouse promoter-CAT constructs in 3T3 cells showed no significant difference between the two, while a clear difference was observed when comparing the activity of the human promoter in the mouse and human cell lines. This difference is similar to that reported by Parker et al, (1992) showing a marked difference in α2(1) procollagen promoter activity in CT-1 and SVWI-38 cells.

The DNA-protein complex which is common to both the human (COL1A2) and mouse (Col1A2) promoters, is the G/CBE (complex I) binding proteins in human and CBF in mouse. This protein complex is ubiquitous and is found in all the cell lines used in this study. Similar findings reported by Vuorio et al, (1990) showed that the mouse CBF is a ubiquitous factor. Cross-species EMSA's using human nuclear extracts with the mouse proximal α2(1) procollagen promoter produced a single DNA-protein complex, having the same migration pattern as that observed in the interaction of CBF with its consensus sequence. Further characterisation of the complex I proteins using consensus binding sites of other known CCAAT binding proteins as competitors showed that double stranded oligonucleotides of the mouse CCAAT binding element and the NF-Y binding sequence competed out complex I formation whereas a consensus oligonucleotide of the CCAAT/enhancer binding protein (C/EBP) did not. The use of a mouse anti-CBF-B antibody in EMSA's using human nuclear proteins resulted in the supershift of the complex I proteins, confirming that complex I and CBF are related or similar proteins. The same antibody also resulted in the supershifting of the mouse CBF. These results and those presented by Collins et al, (1997) confirm that complex I is a member of the heterologous CCAAT-binding proteins.

Further analysis of the complex I, II an III proteins using phosphatase-treated nuclear proteins in EMSA's showed that all three proteins require phosphorylation for DNA binding activity. Dephosphorylation of nuclear proteins using calf intestinal phosphatase resulted in the loss of complex I formation on the COL1A2 promoter and CBF binding to the Col1A2 promoter. The serine/threonine phosphatase, PP2A resulted in the loss of complexes II and III formation. To confirm these findings, cells were incubated with the kinase inhibitor, staurosporin. This resulted in a decrease in all DNA-protein interactions with the proximal α2(1) procollagen promoter. It is possible that phosphorylation leads to an altered protein conformation or to altered interaction with other proteins that

would allow for DNA binding activity (reviewed by Hunter and Karin, 1992). Well documented examples of transcription factors where DNA binding activity is stimulated by phosphorylation are rather limited. An extensively studied example is the serum response factor (SRF), which binds and activates the c-fos promoter through the serum response element (SRE). The mechanism whereby phosphorylation affects DNA binding is unknown, but phosphorylation induces a conformational change in SRF (Treisman, 1986 and Janknecht et al, 1992). It is likely that phosphorylation of the complex I, II and III proteins occur at ser/thr residues since a tyrosine kinase inhibitor had no effect on their DNA binding ability. The experiments performed in this study do not shed light on exactly how and where phosphorylation of complexes I, II, III and CBF occur and whether it induces a conformational change in the proteins. These would be interesting questions to pursue once these proteins have been purified.

The inhibition observed in α2(1) procollagen promoter activity and gene expression (CT-1 cells) in the presence of the kinase inhibitor (staurosporin), correlates well with the loss of trans-acting factor interactions in the proximal promoter region and provides support for the proposal that phosphorylation events are important in the regulation of α2(1) procollagen gene expression. Sobel et al, (1983); Rabin et al, (1986) and Greenwel et al, (1997) reported that an activator of protein kinase C activity, PMA, resulted in inhibition of α2(1) procollagen gene expression. This is contradictory to findings by Stuiver et al, (1991) showing that PMA has a stimulatory effect on α2(1) procollagen expression in 3T3-L1 cells. It is possible that cell-specific signalling pathways are responsible for the differences observed and that a multitude of protein kinases play a role in expression of the collagen genes.

In an attempt to identify the complex II and III proteins, a concatemerised CME probe was used to screen a human heart λgt11 expression library. Two clones, λ2

and λ4, were identified. These clones expressed fusion proteins with approximate sizes of 139 kD and 153 kD as determined by south western blotting. They were also IPTG inducible, an indication that their expression was under the control of the lac Z promoter. When taking the size of β-galactosidase into account, the size of λ2 corresponds to approximately 23 ±8 kD and λ4 to 37 ±10 kD. The fusion protein expressed by one of the clones, λ2, was shown to be sequence-specific since unlabelled wild type CME could compete out formation of the DNA-protein complex detected in EMSA's. It contained a cDNA insert of approximately 1 kb and a databank search for sequence similarities between λ2 and published sequences did not provide any significant homologies. Also, a search in the protein databank for similarities between the λ2 peptide and other proteins did not provide any significant homologies. It is therefore possible that λ2 is a product of an as yet unidentified gene that may have a role in the regulation of α2(1) collagen synthesis. Northern blot analysis using λ2 as a probe showed that a 4 kb mRNA transcript is associated with cells where α2(1) collagen is not detected. This is similar to the association of complex II proteins with cells where α2(1) collagen expression is inhibited. The identification of genes which express proteins that bind the α2(1) procollagen promoter will allow investigations that will lead to an improved understanding of the mechanisms by which the type I collagen genes are regulated.

5.3 Novel aspects of this investigation

Tissue and cell-specific expression of the mouse type I collagen genes has been shown by a number of groups (Rossi and de Crombrugghe, 1987; Slack et al, 1991; Simkevich et al, 1992; Sokolov et al, 1995). These studies have focused on the identification of *cis*-regulatory elements and the trans-acting factors which bind to them. In this study, a clear correlation between trans-acting factor interaction and expression of the human α2(1) procollagen gene was identified. A

study by Walsh and Schimmel (1987) showed a similar association of trans-acting factor binding in the muscle α-actin promoter during myocyte differentiation, where binding of a specific trans-acting factor is greatly diminished in non-myocyte cells. The present study implicates cis-acting elements and their respective trans-acting factors with the expression of the α2(1) procollagen gene.

Evidence that a 3 bp mismatch between the human and rodent α2(1) procollagen promoters within the CME confers species-specific differences in promoter activities is provided. This finding is novel in that the mismatch prevents proposed activators and repressors (complexes II and III respectively) from binding to the rodent α2(1) procollagen promoter. The Col1A2 and COL1A2 promoters also respond to TGF-β via binding of different trans-acting factors to the individual promoters (Rossi et al, 1988; Inagaki et al, 1995; Chung et al, 1996; Greenwel et al, 1995, 1997). These studies showed that there are species-specific differences in the response of α2(1) procollagen promoter to cytokines like TGF-β. Our investigation contributes to towards studies implying species-specific differences in α2(1) procollagen gene expression, albeit in a different region of the α2(1) procollagen promoter. The promoter region investigated in this study is located downstream of the TGF-β response element.

Evidence for the requirement of phosphorylation for the DNA-binding activities of complex I, II, III and CBF is provided. This is of interest as no CCAAT-binding factor to date has been shown to require phosphorylation for DNA-binding activity. Although phosphorylation events are suggested to play a role in collagen gene expression (review by Slack et al, 1993), little direct evidence of its effect on trans-acting factor binding to the α2(1) procollagen promoter has been reported. A study by Greenwell et al, (1995) however show that tyrosine dephosphorylation of nuclear proteins result in enhanced DNA-protein interaction with the human α2(1) procollagen TGF-β response element. There is much speculation as to exactly

which trans-acting factors bind to this region (section 1.). Chung et al, (1996) implicated AP-1 binding and the components of AP-1, Fos and Jun are both influenced by phosphorylation. c-Jun DNA-binding activity is reversibly inhibited by phosphorylation of Ser-243, (Boyle et al, 1991). These findings may have a bearing on the study by Greenwel et al, (1995) which showed that dephosphorylation of nuclear proteins mimics TGF-β stimulation of α2(1) procollagen gene expression. The response observed may be due to dephosphorylation of proteins such as Sp1 (Greenwel et al, 1995; 1997) or Jun which can result in binding of Sp1 or AP-1 to the the TGF-β response element. It is possible that TGF-β stimulation results in the activation of a phosphatase(s) involved in dephosporylation of these proteins or other factors which affect trans-acting factor binding to the α2(1) procollagen promoter. These ideas are speculative but shows that phosphorylation and dephosphorylation events have key roles in the regulation of α2(1) procollagen gene expression.

The isolation of a cDNA whose expression is associated with the inhibition of α2(1) collagen synthesis will allow us to perform investigations that will lead to a greater understanding of α2(1) collagen gene expression.

5.4 Extended model of trans-acting factor interactions with the -107 to +54 α2(1) promoter.

The results of this study can be used to formulate a model which encomposes the major findings (fig 5.1).

1. An external ligand leads to the activation of the protein kinase C pathway and other signalling pathways.

2. Protein kinase C phosphorylates a serine/threonine kinase amongst other proteins. It may also phosphorylate nuclear factors directly.

3. The ser/thr kinase (and possibly other kinases) phosphorylate proteins including the nuclear proteins forming complexes I, II, III in human and CBF in mouse.

4. The phosphorylated proteins are stimulated to bind their recognition sequences leading to activation/inhibition of collagen gene expression in a cell and species-specific manner.

5. The interaction of complex I and III proteins with the proximal α2(1) procollagen promoter leads to expression of the gene, while the interaction of complex II causes downregulation of the gene. The loss of complex III binding may also account for decreased expression.

6. Complex I and CBF are similar in their binding activities and are possibly members of the same family of proteins.

This study contributes towards our understanding of the mechanisms involved in the transcriptional regulation of the human α2(1) procollagen gene. The proximal promoter binding proteins discussed in this thesis are likely to interact with each other and the transcription initiation machinery to either activate or down regulate transcription of the human α2(1) procollagen gene in a cell-specific manner.

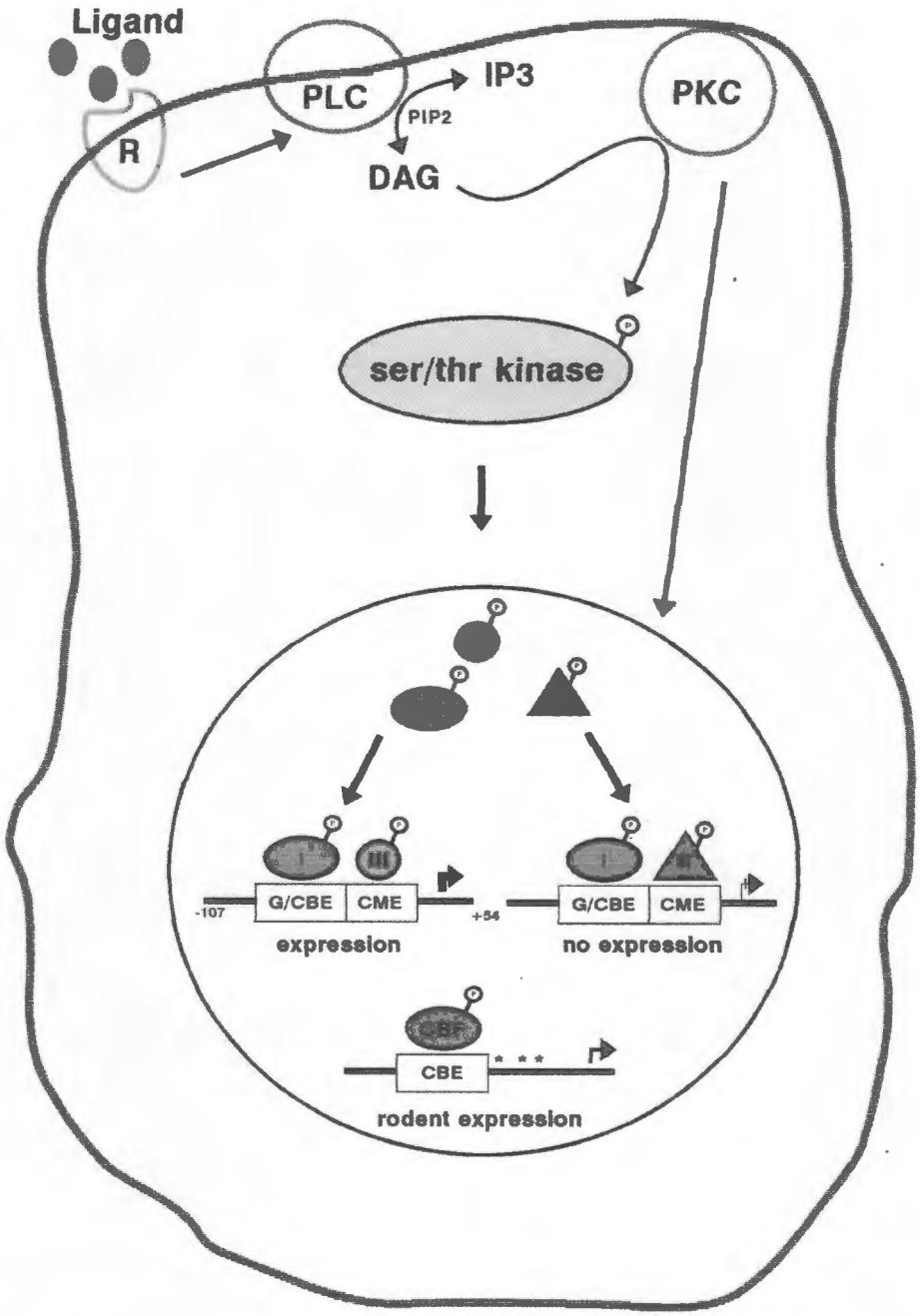


Fig 5.1 Proposed model for the cell- and species-specific regulation of the α2(1) procollagen promoters. An extracellular ligand initiates the activation of a mebrane bound receptor (R). This activated receptor results in the activation of phospholipase C, which cleaves PI 4,5-bisphosphate (PIP_2) to generate diacylglycerol (DAG) and inositol 1,4,5-triphosphate (IP_3). DAG stimulates the activation of protein kinase C (PKC) which in turn can activate a number of other protein kinases. Protein kinase C or other kinases (possibly ser/thr kinases) in turn phosphorylate the trans-acting factors which interact with the proximal α2(1) procollagen promoter resulting in the binding of these factors to consensus sequences in the promoter and ultimate regulation of the gene. The thickness of the arrows indicate the degree of promoter activity.

6. MATERIALS AND METHODS

CHAPTER 6:

MATERIALS AND METHODS

6.1 Cell Culture

6.1.1 Maintenance of cells in culture

Cells were grown in DMEM containing 10% fetal calf serum, 100μg/ml streptomycin and 250u/ml penicillin. The cells were incubated at 37°C in a humidified 95% air, 5% CO_2 incubator. Upon reaching confluence, cells were trypsinised with 0.05% trypsin in PBS containing 10mM EDTA, and split at a ratio of 1:4 for transformed cells and 1:2 for non-transformed cells. The mouse macrophage cell line, $P388D_1$ and the human monocyte cell line, U937 were grown in suspension in RPMI medium containing 100μg/ml streptomycin and 250u/ml penicillin.

The cell lines used in this study were:

1. human skin fibroblasts, (FG_o) (van der Westhuysen et al, 1984)
2. human embryonic lung fibroblasts, (WI-38) (ATCC CCL-75)
3. γ-radiation-transformed WI-38 fibroblasts, (CT-1) (Namba et al, 1980)
4. simian virus 40 transformed WI-38 fibroblast, (SVWI-38) (de Haan et al, 1986)
5. human fibrosarcoma cells, HT1080 (ATCC CCL-121)
6. normal human lymphoblastoid cells, L-77
7. human hepatocellular carcinoma cells, HepG2 (ATCC HB 8065)

8. human mammary epithelial tumour cell lines, ZR-75-2 (ATCC CRL-1500) and MDA-MB-231 (ATCC HTB-26).
9. human monocytic cell line (U937) (ATCC CRL 1593)
10. Swiss mouse 3T3 fibroblasts (ATCC CCL 92)
11. Swiss mouse 3T6 fibroblasts (ATCC CCL 96)
12. Rat-1 lung fibroblast (Rat-1) (Slack et al, 1992)
13. Ras-transformed Rat-1 fibroblasts, (Rasrat-1) (Slack et al, 1992)
14. mouse macrophages, ($p388D_1$) (Koren et al, 1975)

6.1.2 Treatment of cells with the kinase inhibitors, staurosporin or genistein and the phosphatase inhibitor, okadaic acid

Cells in suspension were pelleted and resupended in RPMI medium. Approximately 5 X 10^6 cells in 10 ml of medium was treated with the required concentration of inhibitor (see section 3.2.5.2). Incubation was continued for between 3 and 17 hours and cell death observed by trypan blue exclusion. After incubation with inhibitors the cells were washed with PBS and nuclear proteins extracted as described below (section 6.4.1).

6.1.3 Collagen synthesis

For the determination of collagen synthesis , cells were plated at a density of $2x10^5$ cells per 30 mm diameter dish. The cells were grown at 37°C in a CO_2 incubator until approximately 80% confluent, and labelled for 18 hours in serum-free medium containing 10 μCi/ml of 2,3-[^{3}H]-proline, 50 μg/ml ascorbate and 50 μg/ml of β-aminopropionitrile (BAPN).

The medium (2ml) from cells labelled with 2,3-[^{3}H]-proline was added to 200 μl of a solution containing 1 mM PMSF and 0.8 mg/ml n-ethylmaleimide. To this

mix was added 50 µl of 1 mg/ml rat tail collagen in 0.5 M acetic acid. The cell layer was rinsed in 200 µl PBS which was added to the above and 600 µl of ice cold 96% ethanol was added to precipitate the collagen at -20°C for 16 hours. The collagen was pelleted at 10 000 rpm at 4°C for 30 min in a Beckman JA20 rotor. The ethanol was decanted and the pellet washed twice in 70% ethanol. All traces of ethanol was removed under vacuum and the collagen dissolved in 100 µl of 0.5 M acetic acid. Non-collagenous proteins were digested by the addition of 5 µl of 1 mg/ml pepsin and incubated overnight at 4°C.

Five µl of the protein sample was incubated with 5 µl of 20% TCA on ice for 1 hour. This was followed by incubation at 90°C for 30 min, on ice for 30 min, the sample pippeted onto GF/C filters, washed three times in 5% TCA and a final wash with 70% ethanol. The filters were allowed to air dry and radioactivity counted in 10 ml scintillation fluid.

6.2 Preparation and transformation of competent *E. coli* cells

6.2.1 Preparation of competent cells

Glycerol stocks of *E.coli* DH5α or DK-1 cells were used to prepare competent cells. Fifty µl of a glycerol stock was inoculated into 10 ml of luria broth and grown overnight at 37°C with shaking. Three ml of the overnight culture was inoculated into 300 ml of luria broth in a 2 litre conical flask and incubated at 37°C with vigorous shaking until the cultures reached an OD_{650nm} of 0.2 to 0.4. The cells were pelleted by centrifugation at 5 000 rpm at 4°C in a Beckman JA10 rotor for 10 min. The pellet was resuspended in 40 ml of ice cold 60 mM $CaCl_2$, 10 mM Pipes pH 7.2. This suspension was left on ice for 30 min, transferred to Corex glass centrifuge tubes and centrifuged at 5 000 rpm at 4°C for 5 min in a Beckman JA20 rotor. Four ml of a solution of 60 mM $CaCl_2$, 10 mM Pipes and

15% glycerol was added to the pellet and gently mixed. The cell suspension was stored in aliquots of 200 µl at -70°C.

6.2.2 Transformation of competent cells

The transformation efficiency of the cells were checked using the vector pUC19. An aliquot of competent cells were thawed on ice and 100 µl transferred to a 15 ml sterile plastic tube. To the cells was added 1 ng of pUC19 DNA, mixed gently and incubated on ice for 30 min, followed by heat shock at 42°C for 3 min before the addition of 1 ml of prewarmed (37°C) L-broth and incubation at 37°C for 1 hour. 50, 100 and 200 µl aliquots were plated out on luria agar plates containing 50µg/ml and incubated overnight at 37°C. A successful competent cell batch usually formed between 10^7 and 10^8 colonies per µg of pUC19 DNA.

6.2.3 Preparation of plasmid DNA using the rapid plasmid extraction method

An alkaline lysis method (Birnboim and Doly, 1979; Birnboim, 1983) was used for small scale crude plasmid extracts. Ten ml of L-broth containing 50 µg/ml ampicillin was inoculated with 50 µl of a glycerol stock and incubated overnight at 37°C with shaking. Before harvesting of the overnight culture, glycerol stocks were prepared. The cells were pelleted in a Beckman TJ-6 centrifuge at 2000 rpm for 10 min at 4°C. The cell pellet was resuspended in 200 µl of solution 1 (25 mM Tris-Cl pH 8.0, 10 mM EDTA and 50 mM glucose), and incubated at room temperature for 5 min. Cells were lysed by the addition of 400 µl of solution 2 (1M NaOH, 1% SDS) and left on ice for 5 min. The genomic DNA and cell membranes were precipitated by the addition of 300 µl of 3M potassium acetate, pH 4.8 and incubation on ice for 10 min. The cell debris and chromosomal DNA was removed by centrifugation in a microfuge for 5 min. Supernatants were

transferred to a fresh tube and plasmid DNA was precipitated by the addition of an equal volume of isopropanol and left for 30 min at -20°C. The plasmid DNA was pelleted by centrifugation in a microfuge, washed twice with 70% ethanol and dried under vacuum. The pellet was resuspended in 50-80 µl of TE.

6.2.4 Blunt end ligation of DNA fragments into plasmid vectors

DNA fragments with 5'-overhanging ends were incubated with the klenow fragment of *E. coli* DNA polymerase in order to fill in recessed 3'-ends of double stranded DNA (Sambrook et al, 1989). One unit of klenow polymerase in klenow buffer containing 0.5 mM each of dGTP, dATP, dTTP and dCTP was incubated with 100 -1000 ng of double stranded DNA in a final volume of 20 µl. The reaction was incubated at room temperature for 30 min, extracted with phenol:chloroform:isoamyl alcohol (25:24:1) and the DNA precipitated with 96% ethanol. DNA was washed, dried and resuspended in a small volume of sterile water.

For blunt end ligation, plasmids were digested with the restriction enzyme SmaI. A DNA insert : vector ratio of 5:1 was the prefered ratio used in ligations. Ligation reactions were performed in a final volume of 10 µl as recommended by the manufacturers of the ligase or by in gel ligations (section 6.2.5).

6.2.5 In gel ligation

In gel ligations were performed as previously described (Struhl, 1985). DNA samples were digested with the appropriate restriction enzymes and electrophoresed on 0.8% low melting agarose gels in TAE buffer (section 6.11). The desired bands were cut out in a small volume using a razor blade and transferred to a microfuge tube. The gel slices were melted at 70°C for 10 min and

added to a ligation reaction containing insert and vector DNA in a volume of 9 µl. The tube was placed at 37°C and 11µl of a cocktail containing 2x T4 DNA ligase buffer (section 6.11), 1 mM ATP and 1unit T4 DNA ligase was added and incubated overnight at room temperature. After the ligation was completed, the gelatenous mixture was remelted at 70°C for 5 min and 5 to 10 µl was added to 100 to 200 µl of competent cells.

Ligation of vector and inserts were detected by blue/white selection of DH5α competent cells on L-agar ampicillin (50 µg/ml) plates onto which 50 µl of 100 mM IPTG and 20 µl of 2% X-gal had been spread. Plasmids with inserts were isolated as described in section 6.2.3.

6.3 Preparation of DNA fragments for EMSA's

6.3.1 Preparation of plasmid DNA

A 500 ml overnight culture of E.coli containing the desired plasmid was harvested at 5000 rpm for 10 min in a Beckman JA10 rotor at 4°C. The Qiagen plasmid DNA extraction kit (Diagen GmbH, Dusseldorf, Germany) was used to prepare plasmid DNA as described by the manufacturers. After precipitation, the DNA was washed with 70% ethanol and resuspended at a concentration of 1 µg/µl in sterile TE (section 6.11) or sterile water depending on the final use of the DNA.

6.3.2 Isolation of DNA fragments

Plasmid DNA was digested with the required restriction enzymes for 2-16 hours and the digested material resolved on 7% non-denaturing polyacrylamide gels or 1% agarose gels. The band/s of interest was eluted from the gels using one of the following methods.

6.3.2.1 Crush soak method

After fractionating the released DNA fragments DNA on non-denaturing polyacrylamide gels, the gels were stained with ethidium bromide. The band of interest was excised from the gel using a sharp blade, sliced into smaller pieces and transferred to a microfuge tube containing 500 µl of DNA extraction buffer (section 6.11). This was incubated overnight at 37°C with gentle agitation. The buffer was removed and transferred to a fresh microfuge tube. Remaining gel slices were washed with 100 µl of DNA extraction buffer which was added to the above buffer. The DNA was then precipitated by the addition of 1/10 volume of 3M sodium acetate and 2.5 volumes of ice cold 96% ethanol. The DNA was pelleted by centrifugation, washed with 70% ethanol, dried under vacuum and resuspended in 1 x TE. Recovered DNA was stored for later use in ligation reactions or DNA-protein binding experiments.

6.3.2.2 Qiaex extraction method

DNA fragments were fractionated on 0.8% low melting agarose gels in 1 x TAE buffer (section 6.11). The required DNA fragment was cut out of the gel and the DNA was extracted using the Qiaex gel extraction kit. The method of extraction is as described by the manufacturers, (Diagen GmbH, Dusseldorf, Germany). DNA prepared using this method was used in ligation experiments only.

6.4 DNA-protein interactions

6.4.1 Preparation of nuclear proteins

Nuclear proteins were extracted from cells using the method of Dignam et al, (1983). Briefly, cells grown in 150 mm diameter tissue culture dishes were harvested in ice cold PBS using a rubber policeman and pelleted in a Beckman TJ-

6 centrifuge at 2000 rpm at 4°C for 10 min. Cells grown in suspension were centrifuged in a Beckman JA10 rotor. The cells were washed in ice cold PBS, pelleted and resuspended in 5 packed cell volumes of buffer A (section 6.11). The cells were allowed to swell on ice for 10 min, pelleted again and resuspended in 2 original packed cell volumes of buffer A. The resuspended cells were lysed using 10-15 strokes of a tight pestle in a Dounce homogeniser. Cell lysis was checked by microscopy. The nuclei were pelleted by centrifugation at 13 000 rpm at 4°C in a Beckman JA20 rotor for 30 min. Pelleted nuclei were resuspended in 1-3 ml of buffer C (section 6.11) containing 0.5 mM PMSF, 1 μg/ml pepstatin and 1 μg/ml leupeptin and homogenised with 10-15 strokes with the tight pestle of a Dounce homogeniser. The homogenised nuclei were transferred to a sterile 50 ml tube and gently stirred on ice for 30 min. The salt soluble proteins were separated from nuclear debris by centrifugation in a Beckman JA20 rotor for 30 min and 13 000 rpm at 4°C. The supernatant was dialysed against 50 volumes of buffer D (section 6.11) for 5 hours at 4°C. The dialysate was centrifuged at 13 000 rpm for 30 min in a Beckman JA20 rotor at 4°C and the supernatant stored at -70°C in aliquots of 20-100 μl.

6.4.2 End-labelling of DNA

DNA fragments were end-labelled using the klenow fragment of *E. coli* DNA polymerase. 3'-recessed ends of double stranded DNA were labelled as described by Sambrook et al, (1989). One unit of Klenow DNA polymerase was incubated with approximately 200-500 ng of DNA in 1X klenow buffer containing 3.3 mM each of dGTP, dATP and dTTP and 20 μCi of [α-^{32}P]-dCTP. The reactions were incubated at room temperature for 30 min and stopped by the addition of an equal volume of chloroform:isoamylalcohol (24:1 (v/v)). Unincorporated nucleotides were separated from the labelled DNA using sephadex G-50 chromatography. Fractions were collected in 150 μl of sterile water and the tubes containing the

initial peak of radioactivity were pooled. Ten µl of the pooled fractions were counted to calculate the specific activity of the labelled DNA.

6.4.3 Electrophoretic mobility shift assays (EMSA)

Four µg of crude nuclear extract was incubated with 2 µg of poly dIdC/poly dIdC and 4 µl of incubation buffer (section 6.11) in a volume of 20µl for 10 min on ice prior to the addition of labelled probe. Labelled probe was added at 10^4cpm (approximately 1 ng DNA) and the incubation continued for a further 30 min on ice (Fried and Crothers, 1981). At the end of the incubation period, 2 µl of 0.25% bromophenol blue in 35% glycerol was added. The DNA-protein complexes were separated on non-denaturing 5% polyacylamide gels at 150 V for approximately 2 hours at 4°C in 0.5X TBE. The gels were dried under a vacuum and exposed to X-ray film for 16 hours.

6.4.3.1 EMSA Competition experiments

Competition experiments were performed by incubating the nuclear extracts with 50 to 200 molar excess of double stranded oligonucleotides for 10 min before the addition of the labelled probe. EMSA's were then completed as described above.

Sense and antisense oligonucleotides were synthesised using a Beckman system 1000A DNA synthesiser. Double stranded oligonucleotides were produced by incubating an equal concentration of complementary oligomers at 90°C for 5 min, 65°C for 30 min and annealing at 37°C for 1 hour followed by slow cooling to room temperature. Annealed oligonucleotides were checked on 12% non-denaturing polyacrylamide gels.

Oligonucleotides used were:

NF-Y	(a)	$^{5'}$AACATTTTTCCTGATTGGTTAAAAGTTG$^{3'}$
	(b)	$^{5'}$CAACTTTTAACCAATCAGAAAAATGTT$^{3'}$
C/EBP	(a)	$^{5'}$AATTCAATTGGGCAATCAGG$^{3'}$
	(b)	$^{5'}$AATTCCTGATTGCCCAATTG$^{3'}$
CME (EcoR1 5'overhangs)	(a)	$^{5'}$AATTCGGAGGCCCTTTTGGAGG$^{3'}$
	(b)	$^{5'}$AATTCCTCCAAAAGGGCCTCCG$^{3'}$
G/CBE (BamH1 5'overhangs)	(a)	$^{5'}$GATCCGCCCTCCCATTGGTG$^{3'}$
	(b)	$^{5'}$GATCCACCAATGGGAGGGCG$^{3'}$

6.4.3.2 EMSA Supershift experiments

One, 5 and 10 µl of a 1/100 dilution of antibodies against the CCAAT-binding factor (CBF) (a kind gift from Drs Heide Eberspaecher and Benoit de Crombrugghe, University of Texas) was incubated with nuclear extracts for 30 min at 37°C in incubation buffer prior to the addition of probe. 2 µg of poly dIdC/poly dIdC was incubated with the antibody/nuclear extract mix for a further 10 min on ice, followed by the addition of labelled probe as described above. DNA-protein complexes were analysed as described in section 3.2.4.2.

6.5 SDS-Polyacrylamide gel electrophoresis

6.5.1 Preparation and electrophoresis of gels

7% Separating gel (for mini gels, final volume 10 ml) were made up as follows:

distilled water	5 ml
1.5 M Tris-Cl (pH 8.8)	2.5 ml
10% SDS	100 µl
30% acrylamide:bisacrylamide stock (29:1)	2.33 ml
10% ammonium persulphate	50 µl
TEMED	5 µl

The gel was cast in a Biorad minigel system, covered with a thin layer of water and allowed to set before addition of the stacking gel.

4% Stacking gel (final volume 5 ml):

distilled water	3 ml
0.5 M Tris-Cl (pH 6.8)	1.25 ml
10% SDS	50 µl
30% acrylamide:bisacrylamide stock (29:1)	0.65 ml
10% ammonium persulphate	25 µl
TEMED	5 µl

The stacking gel was poured on top of the separating gel, a comb inserted and the gel left to set at room temperature.

Protein samples were lyophilised, resuspended in 10-15 µl of 2x treatment buffer (section 6.11) and 2 µl of stop buffer (section 6.11). The samples were heated at 95°C for 5 min and immediately loaded onto the gel. Gels were electrophoresed in approximately 300 ml 1x SDS PAGE buffer (section 6.11) at 100 V for 1 hour.

For collagen gels, protein samples corresponding to radioactivity from 10^6 cells were dried, resuspended in 10 µl 2X treatment buffer, heated and electrophoresed as described above. The gels were soaked in EN^{3}HANCE autoradiography enhancer (Du Pont) for 1 hour , rinsed in water for 1 hour and dried. Dried gels were exposed to X-ray film for between 16 and 18 hours.

6.5.2 Transfer of proteins to nitrocellulose membranes

Transfer of proteins to Hybond C nitrocellulose membranes was done using the Biorad minigel transfer system. Transfer buffer (section 6.11) was prepared and cooled to 4°C. Nitrocellulose membrane and Whatman 3M paper was cut to the

size of the gel, and soaked in transfer buffer for 15 min. A sandwich of filter paper, nitrocellulose membrane, gel and filter paper was prepared and placed in the transfer apparatus for transfer at 100 V for 1,5 hours at room temperature.

6.6 RNA extraction and Northern blot analysis

6.6.1 Preparation of RNA

RNA was prepared using the method of Chomczynski and Sacchi, (1987). Cells were grown to near confluence as described in section 6.1.1 and fresh meduim was added 24 hours before harvesting. The medium was removed and the cell surface rinsed with ice cold PBS. Cells growing in suspension were centrifuged at 5 000 rpm in a Beckman JA20 rotor for 5 min at 4°C. To adherent cells in petri dishes or the suspension cell pellet was added 1 ml of guanidinium isothiocyanate (GITC) solution D (section 6.11) per 100 mm dish and shaken by gentle oscillation. The mixture was transferred to a sterile 15 ml tube followed by the addition of 100 µl of 2 M sodium acetate (pH4) and gentle mixing by inversion. An equal volume of water saturated phenol was added and the mixture vortexed to disrupt any clumps. 200 µl of chloroform:isoamylalcohol (49:1) was added to the mixture, vortexed and incubated on ice for 15 min. The final suspension was centrifuged at 10 000 rpm in a Beckman JA20 rotor for 20 min. The aqueous phase was removed, mixed with an equal volume of isopropanol and the RNA precipitated overnight at -20°C. The RNA was pelleted by centrifugation and washed with 70% ethanol, dried under vacuum and dissolved in 20-50 µl of sterile distilled water. The concentration of RNA was determined spectrophotometrically and 2 µg of RNA was checked on 1% agarose formaldehyde gels. Gels were stained with ethidium bromide and the integrity of the RNA was ascertained by the intactness of the 18S and 28S ribosomal RNA.

6.6.2 Northern Blot analysis

6.6.2.1 Electrophoresis of RNA on formaldehyde agarose gels

Agarose formaldehyde minigels were perpared by boiling 0.4g high-gelling agarose in 23.3 ml distilled water. After the solution cooled down to approximately 65°C, 8.53 ml of formaldehyde and 4 ml of 10X RNA running buffer (section 6.11) was added.

To 5 μg of cytoplasmic RNA in 3.3μl sterile distilled water was added 1.5μl 10X RNA running buffer, 7.5μl deionised formamide, 2.7μl formaldehyde and 1.5μl RNA loading buffer (section 6.11)

The samples were heated for 10 min at 55°C, before loading on 1% agarose formaldehyde gels. Gels were run at 30 mA in 1X RNA running buffer (section 6.11), with stiring of the buffer every 30 min. RNA was transferred overnight to nylon membranes (Amersham Hybond-N) using the sandwich method (Sambrook et al, 1989). The success of transfer was ascertained by staining the gel with ethidium bromide after transfer. The membranes were rinsed in 6X SSC and the RNA cross-linked using an ultraviolet cross-linker (Spectrolinker, XL-1000 UV crosslinker).

6.6.2.2 Random prime labelling of DNA probes

The labelling reaction was performed as described by the manufacturers (Boeringher Mannheim). 25 ng of DNA was heated at 95°C for 5 min and labelled in a reaction volume of 20 μl containing, 2 μl reaction buffer, 3.3 mM of dGTP, dATP, dTTP, 5 μl of [α-^{32}P]-dCTP and 1 unit of Klenow DNA polymerase for 1

hour at 37°C. Labelled DNA was seperated from unincorporated nucleotides by chromotography on sephadex G-50 columns.

6.6.2.3 Hybridisation of probes to Northern blots

UV Cross-linked membranes were prehybridised for 4 hours at 42°C in a hybridisation oven in 10 ml of hybridisation buffer (section 6.11).

DNA was labelled with [α-^{32}P]-dCTP using the random primed labelling system, heated at 99°C for 5 min, cooled on ice and added to the hybridisation mix at a concentration of 1-2 X 10^6 cpm/ml. Hybridisation was for 18-48 hours at 42°C in a Hybaid hybridisation oven. The membranes were washed twice at room temp for 15 min in 2X SSC and 0.1% SDS, followed by two stringent washes at 65°C in 0.1X SSC and 0.1% SDS. Membranes were sealed in plastic bags and exposed to X-ray film for 16 hours.

6.7 Nuclear Run-on Transcription Assay

One to five 150 mm culture dishes or 175 cm^2 tissue culture flasks containing 5 to 50 X 10^6 cells were used for the isolation of nuclei as described in Current Protocols, (1990). Extracted nuclei were stored in 200 µl glycerol storage buffer at -70°C.

6.7.1 Nuclear run-on transcription

Nuclear run-on transcription assays were performed using the method of Greenberg and Ziff, (1984). 200 µl of nuclei was added to 200 µl of 2x reaction buffer (section 6.11) containing 1 mM of each, GTP, ATP, CTP and 5 µl [α-^{32}P]UTP (10 mCi/ml) for 30 min at 30°C. Six hundred µl of a solution containing

40μg RNase-free DNase in HSB buffer (section 6.11) was added to the labelled nuclei and incubated for a further 5 min at 30°C, followed by the addition of 200 μl of SDS/Tris buffer (section 6.11), 10 μl of 20 mg/ml proteinase K and further incubation at 42°C for 30 min. The sample was mixed vigorously with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) and centrifuged at 2000 rpm in a Beckman TJ-6 centrifuge at room termperature. The aqueous layer was removed and 2 ml of sterile water, 3 ml 10% TCA/60 mM sodium pyrophosphate and 10 μl of 10 mg/ml *E. coli* tRNA was added to precipitate the RNA for 30 min on ice. The precipitate was filtered onto Millipore type HA filters (0.45-μm) and washed 3 times in 10 ml of 5% TCA/30 mM sodium pyrophosphate. The filters were transferred to a scintillation vial and incubated for 30 min at 37°C in 1.5 ml of DNase I buffer (section 6.11) containing 37.5 μl of 1 mg/ml RNase-free DNase I. The reaction was quenched by the addition of 45μl of 0.5 M EDTA and 68μl of 20% SDS, heated at 65°C for 10 min to elute the labelled RNA. The supernatant was removed and 1.5 ml elution buffer (section 6.11) was added to the filter, incubated for 10 min at 65°C and the two supernatants combined. 4.5 μl of 20 mg/ml proteinase K was added to the combined supernatants and incubated at 37°C for 30 min followed by adding an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1), centrifugation and removal of the aqueous phase. 0.75 ml of 1M NaOH was added to the solution and left on ice for 10 min. The reaction was neutralised by the addition of 1.5 ml of 1M HEPES (free acid). Labelled RNA was precipitated overnight at -20°C by the addition of 0.53 ml (1/10 the volume) of 3M sodium acetate and 14.5 ml (2.5 volumes) absolute ethanol. The precipitant was pelleted by centrifugation at 10 000 rpm at 4°C in a Beckman JA20 rotor for 30 min and resuspended in 0.5 ml TES solution (section 6.11). 5 μl RNA samples were counted in liquid scintillation fluid to check the incorporation of labelled UTP into RNA. To the balance of the RNA solution an

equal volume of TES/NaCl solution (section 6.11) was added for use in hybridisation reactions.

6.7.2 Hybridisation of RNA transcripts to cDNA

6.7.2.1 Preparation of nitrocellulose filters for nuclear run-on transcription assays

200μg of plasmid was linearised by digestion with an appropriate restriction endonuclease. 49 μl of 1M NaOH was added to the linearised DNA (in a volume of 440μl) and incubated for 30 min at room temperature. After incubation, 4.9 ml of 6X SSC was added to neutralise the sample and the mixture placed on ice. The slot blot apparatus was set up with 0.45-μm nitrocellulose membrane and 125 μl of sample (~5μg cDNA plasmid) was added per slot under a low vacuum provided by a venturi apparatus. The slots were rinsed with 500 μl 6X SSC, the nitrocellulose filters dried overnight and baked for 2 hours in a vacuum oven at 80°C.

6.7.2.2 Hybridisation reactions

cDNA plasmids immobilised on nitrocellulose membranes as described in section 6.7.2.1 were hybridised with 1 x 10^6 cpm of RNA in 1ml of TES/NaCl solution (section 6.11) for 48 hours at 65°C in sealed vials. After hybridisation, the filters were washed twice in 2X SSC for 30 min at 65°C. The filters were removed from the vials and transferred to 8 ml of 2X SSC containing 80 μg of RNase A and incubated for 15 min at 37°C. A final wash with 2X SSC at 37°C for 10 min was performed, the filters were blotted on Whatman 3MM paper and exposed to X-ray film for 24-48 hours.

6.8 DNA sequencing

Sequencing was performed by the method of Sanger et al, (1977), using the dideoxy chain termination reaction. The protocol was performed as described by the manufacturer of Sequenase™ (United States Biochemical Corporation, Cleveland, Ohio). 2-3 μg of double stranded plasmid DNA was denatured in 0.2M NaOH in 0.2 mM EDTA for 5 min at room temperature and neutralised by the addition of 1/10 the volume of 2M ammonium acetate (pH 4.6) and 2.5 to 3 volumes of 96% ethanol to precipitate the DNA at -70°C for 15 min. The DNA was pelleted, washed in 70% ethanol, dried and redissolved in 7 μl of sterile distilled water. To this was added 1 μl of primer (50 ng/μl), 2 μl of sequenase buffer (section 6.11) and incubated at 65°C for 10 min for the primers to anneal. The reaction was slow cooled by incubation at 37°C for 30 min followed by cooling at room temperature for 1 hour. Labelling and termination reactions were carried out as described by the manufacturers and 3 μl of the final sample was analysed on 6% denaturing-polyacrylamide sequencing gels.

Automated sequencing were performed by a core facility using an Applied Biosynthesis Model 373A Version 2.10.1S automated sequencer.

6.9 Transient Transfection Experiments

6.9.1 Preparation of cells

Cells were grown on 150 mm diameter tissue culture dishes as described in section 6.1. Transient transfections were performed in triplicate using the calcium phosphate-DNA precipitation method (Graham and van der Eb, 1973). Confluent cells on a 150 mm tissue culture dish was trypsinised and seeded into ten 100 mm tissue culture dishes. These were grown to 30-40% confluency in the required

medium prior to the addition of the DNA-$CaPO_4$ precipitate. Four hours before addition of the DNA-$CaPO_4$ precipitate the medium was replaced with fresh growth medium containing 10% heat inactivated foetal calf serum.

6.9.2 Preparation of the DNA-$CaPO_4$ precipitate

$CaCl_2$, TTE and DNA cocktail was prepared in a clear 10 ml plastic tube as follows (100 mm dish):

1X TTE	345 µl
2M $CaCl_2$	60 µl
promoter-CAT construct	10 µl (10µg)
internal control plasmid, CMV-βgal	5 µl (5µg)

This cocktail was added dropwise to 420 µl of 2X HBS with constant agitation. A fine precipitate formed and was left at room temperature for 30 min. The suspension was distributed dropwise onto the cells and was incubated for 24 hours in 5% CO_2 at 37°C. The medium was aspirated, cells rinsed with PBS and fresh medium was added for an additional 24-48 hours.

6.9.3 Extraction of protein from transfected cells

The medium was removed, replaced with 5 ml of trypsin and incubated at 37°C for 2 min. 5 ml of fresh medium was added and the cells dislodged from the substratum by a gentle pipetting motion. The mixture was transferred to a 10 ml tube on ice and centrifuged at 2000 rpm at 4°C in a Beckman JA20.1 rotor for 5 min. The pellet was washed with ice cold PBS, centrifuged and resuspended in 100 µl of 0.25 M Tris-HCl (pH 8) The cell extract was freeze/thawed for three cycles, by freezing in liquid nitrogen and immediate thawing in a 37°C waterbath. Released cellular proteins were separated from cell debri by centrifugation in a

microfuge for 5 min. The supernatant was removed, frozen and stored at -20°C for β-galactosidase and CAT assays.

Protein concentrations were determined using the Biorad reagent as described by the supplier.

6.9.4 β-galactosidase assay

50 μg of cell extract was mixed with 50 μl of a 4 mg/ml O-nitrophenyl-β-D-galactopyranosidase (ONPG), 3 μl of 100X Mg^{2+} buffer (section 6.11), adjusted to 300 μl with 0.1 M $NaPO_4$ and incubated at 37°C for 1 hour by which time a yellow colour was clearly visible. The reaction was stopped by the addition of 500 μl of 1M Na_2CO_3 and the optical density measured at 420 nm. Protein extracts from cells where no DNA was included in the transfection was used as a blank.

6.9.5 CAT Assays

CAT activity was measured using the protocol of Seed and Sheen, (1988). Briefly, 50 μg of cell extract was incubated with 2 μl of ^{14}C-chloramphenicol, 5 μl of n-butyryl CoA (5mg/ml) and the volume adjusted to 125μl with 0.25 M Tris/HCl pH 8.0. After incubation at 37°C for 16 hours, 300 μl of xylene (this extracts the butyrylated chloramphenicol into the xylene phase while unbutyrylated chloramphenicol remains in the aqueous phase) was added, mixed by vortexing, followed by centrifugation and the xylene phase (upper) removed to a fresh microfuge tube. 100 μl of 0.25 M Tris pH 8 was added to the xylene phase , re-extracted as above, and 200 μl of the xylene phase transferred to a scintillation vial. Label incorporation was determined by counting in a Packard Tri-carb 4640 liquid scintillation counter.

The relative CAT activity was calculated as follows:

$$\frac{\text{(CAT activity in experimental sample) - (background CAT activity)}}{\text{(β-galactosidase activity of experimental sample) - (background β-gal activity)}}$$

6.10 Screening of λgt11 expression libraries

6.10.1 Concatemerisation of CME oligomer

Equal concentrations of CME oligonucleotides (a) and (b) (section 6.4.3.1) were incubated with 20 µl of 10 x polynucleotide kinase buffer in a total volume of 130 µl and incubated for 2 min at 90°C, 10 min at 65°C, 10 min at 37°C and 5 min at room temp. 15 µl of a 20 mM ATP stock and 10 units of T4 polynucleotide kinase was added and the reaction incubated for 2 hours at 37°C. The kinase reaction was stopped by the addition of 50 µl of 10 M ammonium acetate, 100 µl of water and heated for 15 min at 65°C. This was followed by cooling to room temperature, the phosphorylated oligonucleotides were precipitated with 2.5-3 volumes of 96% ethanol and ligated to form concatemers.

6.10.2 PCR-generated probes for screening λgt11 expression libraries

Harline et al, (1992) demonstrated that probes labelled by PCR are superior to probes prepared by conventional labelling methods in the successful screening of λgt11 cDNA expression libraries. Concatemerised CME oligonucleotides were cloned into pUC19 and the M13 universal primers were used to prepare labelled probes by PCR. 5 ng of plasmid DNA (containing concatemerised CME), 50 ng of each of forward and reverse primers, 1 mM of dGTP, dATP and dTTP , 10 µl of [α-^{32}P]-dCTP, 5 µl of 10x Taq polymerase reaction buffer and 2 units of Boeringher Mannheim Taq polymerase were prepared in a reaction volume of 50 µl. The conditions for the PCR amplification were 30 cycles of 95°C for 1 min, 58°C for 1 min and 72°C for 1 min for 30 cycles.

After PCR, the reaction volume was increased to 150 µl with sterile water and extracted with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1). The labelled PCR products was separated by chromatography on a sephadex G-50 column.

6.10.3 Screening of a λgt11 expression library

50 µl of *E. coli* Y1090 from a glycerol stock was inoculated in 10 ml of L-broth containing 50 µg/ml ampicillin and 0.4% maltose. The culture was incubated overnight at 37°C with shaking. 50 µl of the overnight culture was inoculated into fresh L-broth and incubated at 37°C until an OD_{550} of between 0.6 and 0.9 was reached. The cells were pelleted by centrifugation at 2000 rpm in a Beckman TJ-6 centrifuge for 10 min and resuspended in 10 ml of 10 mM $MgSO_4$. The cell suspension was stored at 4°C and could be used for up to two weeks.

λgt11 cDNA libraries, titre 1 x 10^9 pfu/ml, were serially diluted in SM buffer to 10^4, 10^3 and 10^2 pfu/ml. 600 µl of Y1090 cells were incubated with 400 µl of diluted phage in SM buffer, incubated at 37°C for 20 min followed by the addition of 8 ml of top agar (section 6.11) (at 45°C) and mixed by gentle inversion of the tube. The top agar mixture was poured onto 150 mm luria agar plates. Plates were left to set at room temperature and transferred to a 42°C incubator until pin-prick sized plaques were visible (usually 3-4 hours). They were then overlayed with nitrocellulose membranes (Hybond-C, Amersham) saturated in 10 mM IPTG and incubated overnight at 37°C.

After overnight incubation, the position of the filters were marked with needle holes and the filters removed from the plates. The plates were stored at 4°C and the filters dried with the protein side up on Whatman 3MM paper for 20 min. The

filters were processed by denaturation/renaturation using 6.6 M guanidine HCl in 1x binding buffer (Vinson et al, 1988). Denaturation/renaturation was performed by sequential dilution of the 6.6 M guanidine HCl as described below. The filters were washed in a solution containing 450 ml of 6.6 M guanidine HCl, 50 ml 10x binding buffer (section 6.11) and 0.5 ml 1M DTT for 5 min. This was decanted, replaced with fresh 6 M guanidine HCl and the wash repeated. The used 6 M guanidine was diluted to 3M with 1x binding buffer and the filters washed for 5 min. Washes with diluted guanidine HCl were continued down to 1.5, 0.75, 0.375 and 0.188 M guanidine HCl. The filters were rinsed twice in 1x binding buffer and then blocked with 5% instant non-fat dry milk in 1x binding buffer containing 0.05% Tween-20 with gentle shaking for 2 hours.

Binding of PCR labelled concatemerised CME probe (10^6 cpm/ml) to proteins on filters was performed in 1x binding buffer, 0.25% milk powder, 5 μg/ml herring sperm DNA, 0.5% Tween-20 for 1-2 hours at 4°C. The filters were washed for 15 min each in three changes of binding buffer containing 0.25% milk powder. Excess moisture was removed by blotting the filters on Whatman 3MM paper. The filters were sealed in plastic bags and exposed to X-ray film for 2-16 hours.

6.10.4 Purification of positive clones

Positive clones producing proteins that bound the radioactive probe were detected by autoradiography. The position of the filters were marked on the autoradiographs and aligned with the plates. The positive plaques were cored out using a sterile pasteur pipette and transferred to 500 μl of SM buffer in a microfuge tube. 5 μl of chloroform was added and the tube incubated overnight at 4°C. The titre of the eluted phage was determined and the phage plated at a low density for rescreening. The screening process (section 6.10.3) was repeated until all the plaques on the plates were positive.

6.10.5 Controls for screening

In order to determine whether the CME probe may cross react with phages other than the potential positives, 600 µl of *E. coli* Y1090 cells in 10 mM $MgSO_4$ was added to 4 ml of top agar (45°C), inverted and poured on L-agar plates. The plates were allowed to set for 20 min and then incubated at 37°C for 2 hours. Purified phage stocks (including the negative controls; λgt11 chicken ovalbumin, λgt11 only and negative clones isolated from primary screens) were spotted onto the lawn of cells in clearly marked positions and incubated at 42°C for 3 hours. At the appearance of tiny plaques, IPTG saturated filters were placed on the lawn and the plates incubated for 6 hours at 37°C. The filters were removed and processed as described in section 6.10.3.

Positive clones were also spotted in duplicate and the filters prepared as described above. The filters were then divided into two for hybridisition with either the CME or an oligomer spanning -57 to -43 of the α2(1) procollagen promoter previously shown not to bind any trans-acting factors.

6.10.6 Rapid analysis of λgt11 fusion proteins

Liquid cultures of *E. coli* Y1090 infected with recombinant λgt11 were induced directly to produce fusion proteins as described by Runge, (1992). 400µl of Y1090 in 10 mM $MgSO_4$ was incubated for 2 hours at 37°C with 100µl phage dilution (1/100) in 8ml of L-broth containing 50µg/ml ampicillin. After 2 hours a 1ml aliquot was removed and processed as described below. To the rest of the culture (7ml), 70µl of 1M IPTG was added and the incubation continued. One ml aliquots were collected on a hourly basis up 3 hour post-induction at 37°C. Processing of aliquots: The infected cells were pelleted by centrifugation in a microfuge for 1 min at room temperature. The supernatant was removed and the pellet resuspended

in 100µl of 1x SDS loading buffer containing 1µl of 100mM PMSF, 1µg/ml each of leupeptin and pepstatin. The samples were boiled for 3 min and stored at -70°C for later analysis. 25µl of sample were analysed on 7% polyacrylamide SDS gels and proteins transferred to nitrocellulose membranes as described in section 6.5.2.

South western blotting was performed using the same method described in screening of λgt11 expression libraries with labelled concatemerised probes (section 6.10.3) Membranes were exposed to X-ray film for 16 hours at -70°C.

6.10.7 Western Blotting using an anti-β-galactosidase antibody

β-galactosidase fusion proteins immobilised on nitrocellulose membranes were prepared as described in section 6.10.6. The membranes were blocked with Tris-buffered saline (TBS) (section 6.11) containing 5% milk powder and 0.2% Tween 20 for 1 hour at room temperature. The filters were incubated with 500µg/ml mouse anti-β-galactosidase antibody in incubation solution (TBS containing 2% milk powder and 0.2% Tween 20) for 45 min at room temperature, followed by repeated washing in incubation solution without antibody. A 1/750 dilution of sheep anti-mouse IgG-peroxidase in incubation solution was reacted with the membranes for 45 min at room temperature. The membranes were washed in incubation solution followed by a final wash in TBS. The membranes were immersed in 60mg of 4-chloro-naphthol dissolved in a solution containing 20ml ice cold methanol, 100ml TBS and 60µl 30% hydrogen peroxide until bands were visualised.

6.10.8. Recombinant lysogen detection

E. coli Y1090 lysogens were isolated by picking colonies surrounding the plaques of positive phages. The colonies were steaked on L-agar plates and incubated

overnight at 32°C. A number of overnight colonies were selected and plated on duplicate L-agar plates. One plate was incubated at 32°C and its duplicate at 42°C. Colonies growing at 32°C but not at 42°C represent lysogens. Lysogens were inoculated in L-broth and grown overnight at 32°C for preparation of glycerol stocks.

For protein assays, crude proteins were extracted as described by Singh et al, (1989). 2 ml of L-broth containing 50μg/ml ampicillin was inoculated with 20μl of an overnight culture of lysogen and incubated at 32°C with good aeration until it reached an OD_{600} of 0.5. The cultures were moved to a 44°C waterbath and incubated for 20 min. IPTG was added to a final concentration of 10 mM to induce the β-galactosidase promoter and incubated for a further 1 hour at 37°C. The induced cells were pelleted by centrifugation in a microfuge for 1 min and the pellet resuspended in 200μl of extraction buffer (section 6.11). The resuspended cells were frozen in liquid nitrogen, thawed and incubated with 0.5mg/ml lysozyme for 15 min on ice. NaCl was added to a final concentration of 1M and the suspension mixed on a rotator for 15 min at 4°C. The lysates were centrifuged for 30 min at 4°C and the supernatants dialysed on Millipore filters (type VS, 0.025μm) against 100ml extraction buffer for 1 hour at 4°C. The dialysed extract was frozen and stored at -70°C for later use.

6.10.9 Analysis of DNA from positive λgt11 clones

5μl of positive phage in SM buffer was heated at 95°C for 5 min followed by cooling on ice for 5 min. A PCR reaction mix was prepared consisting of 5μl phage, 50ng each of λgt11 forward and reverse primers, 1mM dNTP mix, Taq polymerase buffer and 2 units of Taq polymerase in a final volume of 50μl. The

PCR reaction procedure was as follows, 95°C for 1 min, 66°C for 1 min and 72°C for 1 min for 30 cycles.

PCR products were phosphorylated using T4 polynucleotide kinase (Ausubel et al, 1987). One unit of T4 polynucleotide kinase was added to 10μl of PCR product in polynucleotide kinase buffer containing 1mM ATP in a final volume of 20μl for 1 hour at 37°C.

Phosphorylated PCR products were blunt-ended using Klenow DNA polymerase. One unit of Klenow, in 1x klenow reaction buffer containing 0.5mM each of dGTP, dATP, dTTP and dCTP was incubated with 10μl of PCR product for 30 min at room temp. The blunt-ended DNA was ligated to pUC19 digested with Sma1 and transformed into competent DH5α as described in section 6.2.2. Recombinant plasmids were isolated as described in section 6.2.3.

6.11 Buffers and Solutions

Acrylamide Stock Solution (30%) (29:1)
29% Acrylamide
1% Bisacrylamide

Acrylamide Stock Solution (40%) (38:2)
38% Acrylamide
2% Bisacrylamide

Dignam buffer A
10 mM Hepes (pH 7.9)
1.5 mM $MgCl_2$
10 mM KCl
0.5 mM DTT

Dignam buffer C
20 mM Hepes (pH 7.9)
25% glycerol
0.45 M NaCl
1.5 mM $MgCl_2$
0.2 mM EDTA
0.5 mM PMSF
1 μg/ml leupeptin
1 μg/ml pepstatin

Dignam buffer D
20 mM Hepes (pH 7.9)
25% glycerol
0.1 M KCl
0.2 mM EDTA
0.5 mM DTT, 0.5 mM PMSF, 1 μg/ml each of leupeptin and pepstatin

DNA Elution Buffer

0.5 mM Ammonium Acetate

10 mM Magnesium Acetate

1 mM EDTA

Elution Buffer (nuclear run-on)

1% SDS

10 mM Tris-Cl (pH 7.5)

5 mM EDTA

GITC Solution D

4 M GITC

25 mM Sodium Citrate

0.5% Sarkosyl

Add 36µl 2-Mercaptoethanol to 5 ml before use

Glycerol Storage Buffer

50 mM Tris-Cl (pH 8.3)

40% Glycerol

5 mM $MgCl_2$

0.1 mM EDTA

HBS (10 x) (transfections) (per 100 ml of distilled water)

5.0 g Hepes

8.0 g NaCl

0.37 g KCl

1.0 g Dextrose

0.103 g Na_2HPO_4 (anhydrous - 1.45 mM)

HSB Buffer (nuclei extraction)

0.5 M NaCl

50 mM $MgCl_2$

2 mM $CaCl_2$

10 mM Tris-Cl (pH 7.4)

Hybridisation Buffer (Northern blot)

Dextran sulphate	1 g
distilled water	0.8 ml
20x SSC	2.5 ml
500 mM sodium pyrophosphate	1 ml
deionised formamide	5 ml
herring sperm DNA (10 mg/ml)	0.1 ml
10% SDS	0.1 ml

Incubation Buffer (EMSA) (5 x)

100 mM Hepes (pH 7.9)

250 mM KCl

2.5 mM DTT

1.0 mM EDTA

5 mM $MgCl_2$

20% Ficoll 400

L-agar (per litre)

10 g Tryptone

5 g Yeast extract

5 g NaCl

15 g Bacto-agar

L-Broth (per litre)

10 g Tryptone

5 g Yeast Extract

5 g NaCl

Lysogen Extraction Buffer

50 mM Tris-Cl (pH 7.5)

1 mM EDTA

1 mM DTT

1 mM PMSF

PBS (pH 7.3)

137 mM NaCl

2.7 mM KCl

4.3 mM $Na_2HPO_4.2H_2O$

1.4 mM KH_2PO_4

Protein Transfer Buffer

25 mM Tris

192 mM Glycine

20% Methanol

Reaction Buffer without nucleotides (2 x)

10 mM Tris-Cl (pH 8)

5 mM $MgCl_2$

0.3 M KCl

RNA Loading Buffer

1 mM EDTA

0.25% Bromophenol Blue

0.25% Xylene Cyanol

50% Glycerol

RNA Running Buffer (10 x)

200 mM Morpholinopropanesulfonic acid (MOPS) (pH 7.0)

50 mM Sodium Acetate

10 mM EDTA

Screening Binding Buffer (10 x)

200 mM Hepes

30 mM $MgCl_2$

400 mM KCl

10 mM DTT

SDS-PAGE Running Buffer (5 x)

0.25 M Tris-Cl (pH 8.3)

1.92 M Glycine

0.5% SDS

SDS/Tris Buffer

5% SDS

0.5 M Tris-Cl (pH 7.4)

0.125 M EDTA

SM λ Diluent Buffer

10 mM Tris-Cl (pH 7.5)

10 mM $MgCl_2$

Solution 1 (Rapid plasmid prep)

25 mM Tris-Cl (pH 8.0)

10 mM EDTA

50 mM Glucose

Solution 2 (Rapid plasmid prep)

200 mM NaOH

1% SDS

Solution 3 (Rapid plasmid prep)

3 M Potassium Acetate (pH 4.8)

SSC (20 x) (per litre)

3M NaCl

0.3M sodium citrate.$2H_2O$ (pH 7.0)

Stop Buffer

0.05% Bromophenol Blue

0.5% SDS

1 mM EDTA

50% Glycerol

T4 Polynucleotide kinase buffer (10 x)

500 mM Tris-Cl (pH 7.6)

100 mM $MgCl_2$

50 mM DTT

1 mM Spermidine

1 mM EDTA (pH 8.0)

TAE Electrophoresis Buffer (1 x)

40 mM Tris

5 mM Sodium Acetate

1 mM EDTA

pH 7.8 with glacial acetic acid

TBE Electrophoresis Buffer (1 x) (pH 8.3)

90 mM Tris

90 mM Boric Acid

2.5 mM EDTA

TE buffer (10 x) (pH 8.0)

100 mM Tris-Cl

10 mM EDTA

TES Solution

10 mM N-tris(hydroxymethyl)methyl-2aminoethanesulfonic acid (TES) pH 7.4

10 mM EDTA

0.2 % SDS

TES/NaCl Solution

10 mM TES (pH 7.4)

10 mM EDTA

0.2% SDS

0.6 M NaCl

Top agar (per litre)

10 g Tryptone

5 g Yeast extract

5 g NaCl

7 g Bacto-agar

Treatment Buffer (2 x)

0.125 M Tris-Cl (pH 6.8)

4% SDS

20% Glycerol

5 mM EDTA

10% 2-Mercaptoethanol

Tris Buffered Saline (TBS)

25 mM Tris (pH 7.5)

0.9% w/v NaCl

TTE (transfections)

1 mM Tris-Cl (pH 8.0)

0.025 mM EDTA

REFERENCES

Adams, S.L. (1989) *American J. Resp. Cell and Mol. Biol.* **1,** 161-168.

Agarwal, A.R., Goldstein, R.H., Lucey, E., Ngo, H.Q. and Smith B.D. (1996) *J. Cell. Biochem.* **63,** 135-148.

Ala-Kokko, L., Yuan, C.M., Le Guellec, D., Franc, S., Fertala, A., Khillian, J.S., Sokolov, B.P. and Prockop, D.J. (1996) *Ann. N. Y. Acad. Sci.* **785,** 202-203.

Alevizopoulos, A. and Mermod, N. (1996) *J. Biol. Chem.* **271,** 29672-29681.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. eds (1987) *Current Protocols in Molecular Biology,* Greene Publishing Associates, New York.

Avivi, A., Skorecki, K., Yayon, A. and Givol, D. (1992) *Oncogene* **7,** 1957-1962.

Balwin, A.S., LeClair, K.P., Singh, H. and Sharp, P.A. (1990) *Mol. Cell.Biol.* **10,** 1406-1414.

Barberis, A., Superti-Furga, G. and Busslinger, M. (1987) *Cell* **50,** 347-359.

Barnhart, K.M., Kim, C.G., Banerji, S.S. and Sheffery, M. (1988) *Mol. Cell. Biol.* **8,** 3215-3226.

Baxevanis, A.D., Arents, G., Moudrianakis, E.N., and Landsman, D. (1995) *J. Biol. Chem.* **11,** 1475-1489.

Bedalov, A., Breault, D.T., Sokolov, B.P., Lichtler, B.P., Bedalov, I., Clark, S.H., Mack, K., Khillan, J.S., Woody, C.O., Kream, B.E. and Rowe, D.W. (1994) *J. Biol. Chem.* **269,** 4903-4909.

Bedolov, A. (1995) *Biochem. Mol. Biol. Int.* **37,** 159-166.

Beier, F., Vornehm, S., Poschl, E., von der Mark, K. and Lammi, M.J. (1997) *J.Cell Biochem.* **66,** 210-218.

Birk, D.E., Fitch, J.M., Babiarz, J.P., Doanne, K.J. and Linsenmayer, T.F. (1990) *J. Cell Science* **95,** 649-657.

Binetruy, B., Smeal, T. and Karin, M. (1991) *Nature* **351,** 122-127.

Birnboim, H.C. (1983) *Methods in Enzymology* **100,** 243-255.

Birnboim, H.C. and Doly, J. (1979) *Nucleic Acids Res.* **7,** 1513-1523.

Biswas, E.E., Stefanec, M.J. and Biswas, S.B. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 6689-6692.

Boast, S., Su, M-W., Ramirez, F., Sanchez, M. and Avvedimento, E.V. (1990) *J. Biol. Chem.* **265,** 13351-13356.

Bogdanovic, Z, Bedalov, A., Krebsbach, P.H., Pavlin, D., Woody, C.O., Clark, S.H., Thomas, H.F., Rowe, D.W., Dream, B.E. and Lichtler, A.C. (1994) *J. Bone Miner. Res.* **9,** 285-292.

Bornstein, P. (1996) *Matrix Biology* **15,** 3-10.

Bornstein, P. and McKay, J. (1988) *J. Biol. Chem.* **263,** 1603-1606.

Bornstein, P. and Sage, H. (1980) *Annu. Rev. Biochem.* **49,** 957-1003.

Bornstein, P. and Sage, H. (1989) *Prog. Nucleic Acid Res. Mol. Biol.* **37,** 67-106.

Bornstein, P., McKay, J., Liska, D.J., Apone, s. and Devarayalu, S. (1988) *Mol. Cell. Biol.* **8,** 4851-4857.

Bornstein, P., McKay, J., Morishima, J.K., Devarayalu, S. and Gelinas, R.E. (1987) *Proc. Nat. Acad. Sci. USA* **84,** 8869-8873.

Bossone, S.A., Asselin, C., Patel, A.J. and Marcu, K.B. (1992) *Proc. natl. Acad. Sci USA* **89,** 7452-7456.

Bou-Gharios, G., Garrett, L.A., Rossert, J., Niederreither, K. Eberspaecher, H., Smith, C., Black, C. and de Crombrugghe, B. (1996) *J. Cell Biol.* **134,** 1333-1344.

Boulikas, T. (1995) *Crit. Rev. Eukaryot. Gene Expr.* **5,** 1-77.

Boyes, J. and Bird, A. (1991) *Cell* **64,** 1123-1134.

Bradford, M.M. (1967) *J. Mol. Biol.* **26,** 365-369.

Bradham, D.M., in der Wiesche, B., Precht, P., Balakir, R. and Horton, W. (1994) *J. Cell Physiol.* **158,** 61-68.

Brenner, D.A., Rippe, R.A. and Veloz, L. (1989) *Nucleic Acids Res.* **17,** 6055-6064.

Brenner, D.A., Rippe, R.A., Rhodes, K., Trotter, J.F. and Breindl, M. (1994) *J. Lab. Clin. Med.* **124,** 755-760.

Broome, S. and Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 2746-2749.

Burgeson, R.E. (1988) *Ann. Rev. Cell Biol.* **4,** 551-577.

Chan, E. and Goldberg, H. (1995) *J. Biol. Chem.* **270,** 4473-4477.

Chan, H., Hartung, S. and Breindl, M. (1991) *Mol. Cell. Biol.* **11,** 47-54.

Chen, H., Campisi, J. and Radmanabhan, R. (1996) *J. Biol.Chem.* **271,** 13959-13967.

Chen, S.S., Ruteshouser, E.C., Maity, S.N. and de Crombrugghe, B. (1997) *Nuc. Acids Res.* **25,** 3261-3268.

Chodosh, L.A., Baldwin, A.S., Carthew, R.W. and Sharp, P.A. (1988a) *Cell* **53,** 11-24.

Chomczynski, P. and Sacchi, N. (1987) *Anal. Biochem.* **162,** 156-159.

Chung, K-Y., Agarwal, A., Uitto, J., and Mauviel, A. (1996) *J. Biol. Chem.* **271,** 3272-3278.

Clark, A.R. and Docherty, K. (1993) *Biochem. J.* **296,** 521-541.

Clerc, R.G., Corcoran, L.M., LeBowitz, J.H., Baltimore, D. and Sharp, P.A. (1988) Gene Dev. **2,** 1570-1581.

Cohen, R.B., Sheffery, M. and Kim, C.C. (1986) *Mol. Cell. Biol.* **6,** 821-832.

Collins, M. (1993) *Ph.D. Thesis,* University of Cape Town.

Collins, M., Leaner, V.D. and Parker, M.I. (1997) *Biochem. J.* **322,** 199-206.

Courey, A.J., Holtzman, D.A., Jackson, S.P. and Tijian, R. (1989) *Cell* **59,** 827-836.

Coustry, F., Maity, S.N. and de Crombrugghe, B. (1995) *J. Biol. Chem.* **270,** 468-475.

Coustry, F., Maity, S.N., Sinha, S. and de Crombrugghe, B. (1996) *J. Biol. Chem.* **271,** 14485-14491.

Cowell, I.G. and Hurst, H.C. (1994) *Nucleic Acids Res.* **22,** 59-65.

Daniel, S., Zhang, S., DePaoli-Roach, A.A. and Kim, K-H. (1996) *J. Biol. Chem.* **271,** 14692-14697.

Davis, B.H., Chen, A. and Beno, D. (1996) *J. Biol. Chem.* **271,** 11039-11042.

de Crombrugghe, B., Karsenty, G., Maity, S., Vuorio, T., Rossi, P., Ruiteshouser, E.C., McKinney, H. and Lozano, G. (1990) *Ann. N.Y. Acad. Sci.* **580,** 88-96.

de Groot, R.P. and Sassone-Corsi, P. (1992) *Oncogene* **7,** 2281-2286.

de Haan, J.B., Gevers, W. and Parker, M.I. (1986) *Cancer Res.* **46,** 713-716.

de Wet, W., Bernard, M., Benson-Chanda, V., Chu, M-L., Dickson, L., Wel, D. and Ramirez, F. (1987) *J. Biol. Chem.* **262,** 16032-16036.

De Wet, W.J., Chu, M-L. And Prockop, D.J. (1983) *J. Biol. Chem.* **258,** 14385-14389.

Derijard, B., Hibi, M., Wu, I.H., Barrett, T., Su. B., Deng, T., Karin, M. and Davis, R.J. (1994) *Cell* **76,** 1025-1037.

Devereux, J. (1989) *The GCG Sequence Analysis Software Package,* **Version 6.0,** Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wisconsin, USA, 53705.

Devereux, J.P., Haeberli, P. and Smithies, O. (1984) *Nucleic Acids Res.* **12,** 387-395.

Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) *Nucleic Acids Res.* **11,** 1475-1489.

Dorn, A., Bollekens, J., Staub, A., Benoist, C. and Mathis, D. (1987) *Cell* **50,** 863-872.

Driggers, P.H., Ennist, D.L., Gleason, S.L., Mak, W., Marks, M.S., Levi, B., Flanagan, E., Appella, E. and Ozato, K. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 3743-3747.

Eckes, B., Mauch, C., Huppe, G. and Krieg, T. (1993) *FEBS Lett.* **318,** 129-133.

Eizenberg, O. and Oren, M. (1991) *Biochim. Biophys. Acta* **1129,** 34-42.

Fine, A. and Goldstein, R.H. (1987) *J. Biol. Chem.* **262,** 3897-3902.

Flemington, E., Bradshaw, H.D., Traina-Dorge, V., Slagel, V. And Deininger, P.L. (1987) *Gene* **52,** 267-277.

Flink, I.L. and Morkin, E. (1990) *J. Biol. Chem.* **265,** 11233-11237.

Focht, R. and Adams, S.L. (1984) *Mol. Cell. Biol.* **4,** 1843-1852.

Forsburg, S.L. and Guarente, L. (1989) *Genes and Dev.* **3,** 1166-1178.

Foulkes, N.S. and Sassone-Corsi, P. (1992) *Cell* **68,** 411-414.

Fouser, L., Sage, H.E., Clark, J. and Bornstein, P. (1991) *Proc. Nat. Acad. Sci. USA* **88,** 10158-10162.

Fried, M. and Crothers, D.M. (1981) *Nucleic Acids Res.* **9,** 6505-6525.

Furth, J.J., Wroth, T.H. and Acherman, S. (1991) *Experimental Cell Res.* **192,** 118-121.

Galera, P., Musso, M., Ducy, P. And Karsenty, G. (1994) *Proc. Natl. Acad. Sci. USA* **9,** 9372-9376.

Galera, P., Park, R.W., Ducy, P., Mattei, M.G. and Karsenty, G. (1996*) J. Biol. Chem.* **271,** 21331-21339.

Gallinari, P., La Bella, F. and Heintz, N. (1989) *Mol. Cell. Biol.* **9,** 1566-1575.

Giese, K., Pagel, J. and Grosschedl, R. (1994) *Proc. Natl. Acad. Sci. (USA)* **91,** 3368-3372.

Ghosh, S. and Baltimore, D. (1990) *Nature* **344,** 678-682.

Goldberg, H. Helaakosi, T., Garrett, L.A., Karsenty, G. Pellegrino, A., Lozano, G., Maity, S., de Crombrugghe, B. (1992) *J. Biol. Chem.* **267,** 19622-19630.

Graham, F.L. and Van der Eb. A.J. (1973) *Virology* **52,** 456-457.

Green, H., Goldberg, B. and Torado, G.J. (1966) *Nature* **212,** 631-633.

Greenberg, M.E. and Ziff, E.B. (1984) *Nature* **311,** 433-438.

Greenwel, P., Inagaki, Y., Hu, W., Walsh M. and Ramirez, F. (1997) *J. Biol. Chem.* **272,** 197738-19745.

Greenwel, P., Hu, W., Kohanski, R.A. and Ramirez, F. (1995) *Mol. Cell. Biol.* **15,** 6813-6819.

Gregori, C., Kahn, A. and Pichard, A.L. (1993) *Nucleic Acids Res.* **21,** 897-905.

Guenette, D.K., Ritzenthaler, J.D., Foley, J., Jackson, J.D. and Smith, B.D. (1992) *Biochem. J.* **283,** 699-703.

Gunning, P., Ponte, P., Okayama, H., Engel, J., Blau, H. and Kedes, L. (1983) *Mol. Cell. Biol.* **3,** 787-795.

Haas, T. and Plow, E.F. (1994) *Curr. Opin. Cell Biol.* **6,** 656-662.

Handy, D.E. and Gavras, H. (1996) *Hypertension* **27,** 1018-1024.

Harbers, K., Kuehn, M., Delius, H. and Jaenisch, R. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 1504-1508.

Harline, M.C., Kandala, J.C., Sage, R.D., Guntaka. R. and DeAngelo, A. (1992) *Biotechniques* **13,** 388-391.

Harniel, a., Welge-Lussen, U., Kuhn, K. and Poschl, E. (1995) *J. Biol. Chem.* **270,** 11204-11215.

Harrison, J.R., Vargas, S.J., Petersen, D.N., Lorenzo, J.A. and Kream, B.E. (1990) *Mol Endocrinol.* **4,** 184-190.

Hasegawa, T., Takeuchi, A., Miyaishi, O., Isobe, K-I. and de Crombrugghe, B. (1997) *J. Biol. Chem.* **272,** 4915-4923.

Hasegawa, T., Zhou, X., Garrett, L.A., Ruteshouser, E.C., Maity, S.N. and de Crombrugghe, B. (1996) *Nucleic Acid Res.* **24,** 3253-3260.

Hata, R. (1995) *Cell Biol. Int.* **19,** 735-741.

Hatamochi, A., de Crombrugghe, B. and Krieg, T. (1993) *FEBS Lett.* **327,** 325-331.

Hatamochi, A., Golumbek, P.T., Van Schaftingen, E. and de Crombrugghe, B. (1988) *J. Biol. Chem.* **263,** 5940-5947.

Hatamochi, A., Paterson, B. and de Crombrugghe, B. (1986) *J. Biol. Chem.* **261,** 11310-11314.

Heinrichs, A.A.J., Banerjee, C., Bortell, R., Owem, T.A., Stein, J.L., Stein, G.S. and Lian, J.B. (1993) *J. Cell. Biochem.* **53,** 240-250.

Heinrichs, A.A.J., Bortell, R., Bourke, M., Lian, J.B., Stein, G.S. and Stein, J.L. (1995) *J. Cell. Biochem.* **57,** 90-100.

Henthorn, P., Kiledjian, M. and Kadesch, T. (1990) *Science* **247,** 467-470.

Hernandex-Munain, C. and Krangel, M.S. (1994) *Mol. Cell. Biol.* **14,** 473-483.

Higashino, F., Yoshida, K., Fujinaga, Y., Kamio, K. and Fujiaga, K. (1993) *Nucleic Acids Res.* **21,** 547-553.

Ho, I., Bhat, N.K., Gottschalk, L.R., Lindsten, T., Thompson, C.B., Papas, T.S. and Leiden, J.M. (1990) *Science* **250,** 814-818.

Hoeffler, J.P., Meyer, T.E., Yun, Y., Jameson, J.L. and Habener, J.F. (1988) *Science* **242,** 1430-1433.

Hoemann, C.D. and Zarbl, H. (1990) *Cell Growth Differ.* **1,** 581-590.

Hooft van Huijsduijnen, R., Li, X-Y., Blank, D., Matthes, H., Benoist, C. And Mathis, D. (1990) *EMBO J.* **9,** 3119-3127.

Houglum, K., Buck, M., Alcorn, J., Contreras, S., Bornstein, P. and Chojkier, M. (1995) *J. Clin. Invest.* **96,** 2269-2276.

Huerre, C., Junien, Claudine., Weil, D., Chu, M-L., Morabito, M., Van Cong, N., Myers J.C., Foubert, C., Gross, M-S., Prockop, D.J., Boue, A., Kaplan, J-C., De La Chapelle, A. and Ramirez, F. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 6627-6630.

Hunter, T. (1997) *Cell* **88,** 333-346.

Hunter, T. and Karin, M. (1992) *Cell* **70,** 375-387.

Ichiki, Y., Smith, E.A., Leroy, E.C. and Trojanowska, M. (1997) *J. Rheumatology* **24,** 90-95.

Ignotz, R.A. and Massague, J. (1986) *J. Biol. Chem.* **261,** 4337-4345.

Ignotz, R.A., Endo, T. And Massague, J. (1987) *J. Biol. Chem.* **262,** 6443-6446.

Ihn, H., Ohnishi, K., Tamaki, T., LeRoy, E.C. and Trojanowska, M. (1996) *J. Biol. Chem.* **271,** 26717-26723.

Imagawa, M., Chiu, R. and Karin, M. (1987) *Cell* **51,** 251-260.

Inagaki, Y., Truter, S, Tanaka, S., Di Liberto, M. and Ramirez, F. (1995) *J. Biol. Chem.* **270,** 3353-3358.

Inagaki, Y., Truter, S. and Ramirez, F. (1994) *J. Biol. Chem.* **269,** 14828-14838.

Inagaki, Y., Truter, S., Greenwel, P, Rojkind, M., Unoura, M., Kobayashi, K. and Ramirez, F. (1995) *Hepatology* **22,** 573-579.

Inagaki, Y., Truter, S., Tanaka, S., Di Liberto, M., and Ramirez, F. (1995) *J. Biol. Chem.* **270,** 3353-3358.

Inagaki, Y., Tsunokawa, Y., Sakamoto, H., Setsuo, H., Kenichi, K., Hattori, N., Ramirez, Terada, M. and Sugimura, T. (1987) *Biochem. Biophys. Res. Comm.* **148,** 869-875.

Ingraham, H.A., Chen, R., Mangalam, H.J., Eisholtz, H.P., Flynn, I., Lin, C.R., Simmons, D.M., Swanson, L. and Rosenfeld, M.G. (1988) *Cell* **55,** 519-529.

Jahner, D. and Jaenisch, R. (1985) *Nature* **315,** 594-597.

Janknecht, R., Hipskind, R.A., Houthaeve, T., Nordheim, A. and Stunnenberg, HG. (1992) *EMBO J.* **11,** 1045-1054.

Johnson, A.C., Ishii, S., Jinno, Y., Pastan, I. and Merlino, G.T. (1988) *J. Biol. Chem.* **263,** 5693-5699.

Johnson, A.D. (1995) *Cell* **81,** 655-658.

Jost, J.-P. and Hofsteenge, J. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 9499-9503.

Kadonaga, J.T. and Tijian, R. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 5889-5893.

Kageyame, R. and Pastan, I. (1989) *Cell* **59,** 815-825.

Kahari, V-M., Chen, Y.Q., Su, M.W., Ramirez, F. and Uitto, J. (1990) *J. Clin. Invest.* **86,** 1489-1495.

Karin, M. and Hunter, T. (1995) *Current Biology* **5,** 747-757.

Karlsson, O., Thor, S., Norberg, T., Ohlsson, H. and Edlund, T. (1990) *Nature* (London) **344,** 879-882.

Karsenty, G. and de Crombrugghe, B. (1990) *J. Biol. Chem.* **265,** 9934-9942.

Karsenty, G. and de Crombrugghe, B. (1991) *Biochem. Biophys. Res. Comm.* **177,** 538-544.

Karsenty, G. and de Crombrugghe, B. (1991) *J. Biol. Chem.* **265,** 9934-9942.

Karsenty, G. and Park, R-W. (1995) *Intern. Rev. Immunol.* **12,** 177-185.

Karsenty, G., Golumbek, P. and de Crombrugghe, B. (1988) *J. Biol. Chem.* **263,** 13909-13915.

Karsenty, G., Ravazzolo, R. and de Crombrugghe, B. (1990) *J. Biol. Chem.* **266,** 24842-24842.

Kawamura, H., Nagata, K., Masamune, Y. and Nakanishi, Y. (1993) *Biochem. Biophys. Res. Comm.* **192,** 1424-1431.

Kennely, P.J. and Krebs, E.G. (1991) *J. Biol.Chem.* **266,** 15555-15558.

Keshet, I., Lieman-Hurwitz, J. and Cedar, H. (1986) *Cell* **44,** 535-543.

Khilan, J.S., Schmidt, A., Overbeek, P.A., de Crombrugghe, B. And Westphal, H. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 725-729.

Kikuchi, K., Hartl, C.W., Smith, E.A., LeRoy, E.C. and Trojanowska, M. (1992) *Biochem. Biophys. Res. Comm.* **187,** 45-50.

Kim, C.G. and Sheffery, M. (1990) *J. Biol. Chem.* **265,** 13362-13369.

Kim, C.G., Swendeman, S.L., Barnhart, K.M. and Sheffery, M. (1990) *Mol. Cell. Biol.* **10,** 5958-5966.

Kim, I-S., Sinha, S., de Crombrugghe, B. and Maity, S.N. (1996) *Mol. Cell. Biol.* **16,** 4003-4013.

Klemsz, M.J., McKercher, S.R., Celada, A., Van Beveren, C. and Maki, R.A. (1990) *Cell* **61,** 113-124.

Kobayashi, M., Shimomura, A., Hagiwara, M. and Kawakami, K. (1997) *Nucleic Acids Res.* **25,** 877-882.

Kofford, M.W., Schwartz, L.B., Schechter, N.M., Yager, D.R., Diegelman, R.F. and Graham, M.F., (1997) *J. Biol. Chem.* **272,** 7127-7131.

Koren, H.S., Handwerger, B.S. and Wunderlich, J.R. (1975) *J. Immun.* **114,** 894-897.

Kovacs, A., Kandala, J.C., Weber, K.T. and Guntaka, R.V. (1996) *J. Biol. Chem.* **271,** 1805-1812.

Koyama, H., Raines, E.W., Bornfeldt, K.E., Roberts, J.M. and Ross, R. (1996) *Cell* **87,** 1069-1078.

Kratochwil, K., von der Mark, K., Kollar, E.J., Jaenisch, R., Mooslehner, K., Schwarz, M., Haase, K., Gmach, I. And Harbers, K. (1989) *Cell* **57,** 807-816.

Krebsbach, P.H., Nakata, K., Bernier, S.M., Hatano, O., Miyashita, T., Rhodes, C.S. and Yamada, Y. (1996) *J. Biol. Chem.* **271,** 4298-4303.

Kuroki, Y., Shiozawa, S., Sugimoto, T. and Fujita, T. (1992) *Biochem. Biophys. Res. Comm.* **182,** 1389-1394.

Langholz, O., Rockel, D., Mauch, C., Kozlowska, E., Bank, I., Krieg, T. and Eckes, B. (1995) *J. Cell Biol.* **131,** 1903-1915.

Lefebvre, V., Huang, W., Harley, V.R., Goodfellow, P.N. and de Crombrugghe, B. (1997) *Mol. Cell. Biol.* **17,** 2336-2346.

Levine, M. and Manley, J.L. (1989) *Cell* **59,** 405-408.

Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F. and Bird, A. (1992) *Cell* **69,** 905-914.

Li, X-Y., Hooft van Huijsduijnen, R., Mantovani, R., Benoist, C. And Mathis, D. (1992) *J. Biol.Chem.* **267,** 8984-8990.

Lievens, P.M.J., Donady, J.J., Tufarelli, C. and Neufeld, E.J. (1995) *J. Biol. Chem.* **270,** 12745-12750.

Lim, A., Greenspan, D.S. and Smith, B.D. (1994) *Matrix Biol.* **14,** 21-30.

Liou, H-C., Boothby, M.R. and Glimcher, L.H. (1988) *Science* **242,** 69-71.

Liska, D.J., Reed, M.J., Sage, E.H. and Bornstein, P. (1994) *J. Cell Biol.* **123,** 695-704.

Liska, D.J., Robinson, V.R. and Bornstein, P. (1992) *Gene Expression* **2,** 379-389.

Liska, D.J., Slack, J.L. and Bornstein, P. (1990) *Cell Regulation* **1,** 487-498.

Liu, X., Wu, H., Byrne, M., Krane, S. and Jaenish, R. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 1852-1856.

Lochter, A. and Bissel, M.J. (1995) *Cancer Biology* **6,** 165-173.

Lum, L.S., Sultzman, L.A., Kaufman, R.J., Linzer, D.I. and Wu, B.J. (1990) *Mol. Cell. Biol.* **10,** 6709-6717.

Maatta, A., Glumoff, V., Paakkonen, P., Liska, D.J., Penttinen, P.K. and Elima, K. (1993) *Biochem. J.* **194,** 365-371.

Mahoney, C.W., Shuman, J., McKnight, S.L., Chen, H-C. and Huang, D-P. (1992) *J. Biol. Chem.* **267,** 19396-19403.

Maity, S.N. and de Crombrugghe, B. (1992) *J. Biol. Chem.* **267,** 8286-8292.

Maity, S.N., Golumbek, P.T., Karsenty, G. and de Crombrugghe, B. (1988) *Science* **241,** 582-585.

Maity, S.N., Sinha, S., Ruteshouser, E.C. and de Crombrugghe, B. (1992) *J. Biol. Chem.* **267,** 16574-16580.

Marchant, J.K., Hahn, R.A., Linsenmayer, T.F. and Birk, D.E. (1996) *J. Cell Biol.* **135,** 1415-1426.

Marziali, G., Perrotti, E., Ilari, R., Testa, U., Coccia, E.M. and Battistini, A. (1997) *Mol. Cell. Biol.* **17,** 1387-1395.

Massague, J. (1990) *Ann. Rev. Cell Biol.* **6,** 597-641.

Massague, J. and Weis-Garcia, F. (1996) *Cancer Surveys* **27,** (Cell Signalling) 41-64.

Mc Nabb, D.S., Xing, Y. and Guarente, L. (1995) *Genes Dev.* **9,** 47-58.

McVey, J.H., Noruma, S., Kelly, P., Mason, I.J. and Hogan, B.L. (1988) *J. Biol. Chem.* **263,** 11111-11116.

Mitchelmore, C., Troboni, C. and Cortese, R. (1991) *Nucleic Acids Res.* **19,** 141-147.

Miyamoto, M., Fujita, T., Kimura, Y., Maruyama, M., Harada, Y., Sudo, Y., Miyata, T. and Taniguchi, T. (1988) *Cell* **54,** 903-913.

Mufson, R.A. (1997) *FASEB J.* **11,** 37-44.

Muller, M.M., Ruppert, S., Schaffner, W. and Matthias, P. (1988) *Nature* (London) **336,** 544-551.

Murre, C., McCaw, P.S. and Baltimore, D. (1989) *Cell* **56,** 777-783.

Murre, C., Schonleber-McCaw, P. and Baltimore, D. (1989) *Cell* **56,** 777-783.

Myers, J.C., Chu, M-L., Faro, S.H., Clark, W.J., Prockop, D.J. and Ramirez, F. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 3516-3520.

Nah, H-D., Niu, Z. and Adams, S.L. (1994) *J. Biol. Chem.* **269,** 16443-16448.

Namba, M., Nishitani, K. and Kimoto, T. (1980) *Gann* **71,** 300-307.

Nehls, M.C., Grapilon, M.L. and Brenner, D.A.. (1992) *DNA Cell Biol.* **11,** 443-452.

Nehls, M.C., Rippe, R.A., Veloz, L. And Brenner, D.A. (1991) *Mol. Cell. Biol.* **11,** 4065-4073.

Niederreither, K., D'Souza, R., Metsaranta, M., Eberspaecher, H., Toman, P.D., Vuorio, E. and de Crombrugghe, B. (1995) *Matrix Biology* **14,** 705-713.

Niederreither, K., D'Souza, R.N. and de Crombrugghe, B. (1992) *J. Cell Biol.* **119,** 1361-1370.

Oikarinen, J., Hatamochi, A. and de Crombrugghe, B. (1987) *J. Biol. Chem.* **262,** 11064-11070.

Olesen, J.T. and Guarente, L. (1990) *Genes Dev.* **4,** 1714-1729.

Olsen, A.S. and Prockop, D.J. (1989) *Matrix Biol.* **9,** 73-81.

Olsen, A.S., Geddis, A.E. and Prockop, D.J. (1991) *J. Biol. Chem.* **266,** 1117-1121.

Pang, J.H. and Chen, K.Y. (1993) *J. Biol. Chem.* **268,** 2909-2916.

Park, K. and Atchison, M.L. (1991) *Proc. Natl.Acad. Sci. USA* **88,** 9804-9808.

Parker, M.I., de Haan, J.B. and Gevers, W. (1986) *J. Biol. Chem.* **261,** 2786-2790.

Parker, M.I. and Fitschen, W. (1980) *Nuc. Acids. Res.* **8,** 2823-2833.

Parker, M.I., Judge, K. and Gevers, W. (1982) *Nucleic Acids Res.* **10,** 5879-5891.

Parker, M.I., Smith, A.A. and Gevers, W. (1989) *J. Biol. Chem.* **264,** 7147-7152.

Parker, M.I., Smith, A.A., Mundell, K., Collins, M., Boast, S. and Ramirez, F. (1992) *Nucleic Acids Res.* **20,** 5825-5830.

Pavlin, D., Bedalov, A., Kronenberg, M.S., Kream, B.E., Rowe, D.W., Smith, C.L., Pike, J.W. and Lichtler, A.C. (1994) *J. Cell Biochem.* **56,** 490-501.

Pavlin, D., Lichtler, A.C., Bedalov, A., Kream, B.E., Harrison, J.R., Thomas, H.F., Gronowicz, G.A., Clark, S.H., Woody, C.O. and Rowe, D.W. (1992) *J. Cell Biol.* **116,** 227-236.

Payen, E.J. and Cotinot, C.Y. (1993) *Nucleic Acids Res.* **21,** 2772.

Penttinen, R.P., Kobayashi, S. And Bornstein, P. (1988) *Pro. Natl. Acad. Sci. USA* **85,** 1105-1108.

Phillips, N., Bashey, R.I. and Jimenez, S.A. (1995) *J. Biol. Chem.* **270,** 9313-9321.

Piccolo, S., Bonaldo, P., Vitale, P., Volpin, D. and Bressan, G.M. (1995) *J. Biol. Chem.* **270,** 19583-19590.

Pise-Masison, C.A., Dittmer, J., Clemens, K.E. and Brady, J.N. (1997) *Mol. Cell. Biol.* **17,** 1236-1243.

Pogulis, R.J. and Freytag, S. (1993) *J. Biol. Chem.* **268,** 2493-2499.

Poppleton, H.M. and Raghow, R. (1997) *Biochem. J.* **323,** 225-231.

Prockop, D.J. and Kivirikko, K.I. (1995) *Annu. Rev. Biochem.* **64,** 403-434.

Prockop, D.J., Colige, A., Helminen, H., Khillan, J.S, Periera, R. and Vandenberg, P. (1993) *J. Bone Miner. Res.* **8,** S489-S492.

Rabin, M.S., Doherty, P.J. and Gottesman, M.M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 357-360.

Raghow, R., Postlethwaite, A.E., Keshi-Oja, J., Moses, H.L. and Kang, A.H. (1987) *J. Clin. Invest.* **79,** 1285-1288.

Raibaud, O., Gutierrez, C. and Schwartz, M. (1985) *J. Bacteriol.* **161,** 1201-1208.

Rajavashisth, T.B., Taylor, A.K., Andalibi, A., Svenson, K.L. and Lusis, A.J. (1989) *Science* **245,** 640-643.

Ramirez, F. and Di Liberto, M. (1990) *FASEB. J.* **4,** 1616-1623.

Ramirez, F., Boast, S., D'Alessio, M., Lee, B., Prince, J., Su, M-W, Vissing, H. and Yoshioka, H. (1990) *Ann.N.Y.Acad.Sci.* **580,** 74-80.

Ramirez, F., Boast, S., D'Alessio, M., Prince, J., Su, M.W. and Vissing, H. (1989) *Connect. Tissue Res.* **21,** 79-88.

Ray, R. and Miller, D.M. (1991) *Mol. Cell. Biol.* **11,** 2154-2161.

Reith, W., Barrass, E., Satola, S., Kobr, M., Reinhart, D., Sanchez, C.H. and Mach, B. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 4200-4204.

Renkawitz, R. (1990) *Trends Genetics* **6,** 192-197.

Rhodes, D., Rippe, R.A, Umezawa, A., Nehls, M., Brenner, D.A. and Breindl, M. (1994) *Mol. Cell. Biol.* **14,** 5950-5960.

Rhodes, K., Hall, K., Lee, K.E., Razzaghi, H.and Breindl, M. (1996) *Gene Expr.* **6,** 35-44.

Rippe, R.A., Lorenzo, S.I., Brenner, D.A. and Breindl, M. (1989) *Mol. Cell. Biol.* **9,** 2224-2227.

Rippe, R.A., Umeza, A., Kimball, J.P., Breindl, M. and Brenner, D.A. (1997) *J. Biol. Chem.* **272,** 1753-1760.

Ristiniemi, J. and Oikarinen, J. (1989) *J. Biol. Chem.* **264,** 2164-2174.

Ritzenhaler, J.D., Goldstein, R.H., Fine, A. and Smith, D. (1993) *J. Biol. Chem.* **268,** 13625-13631.

Rosenshine, I., Ruschkowski, S. and Finlay, B.B. (1994) *Methods Enzymol.***236,** 467-476.

Rossert, J., Eberspaecher, H. and de Crombrugghe, B. (1995) *J. Cell Biol.* **129,** 1421-1432.

Rossert, J.A., Chen, S.S., Ebersphaecher, H., Smith, C.N. and de Crombrugghe, B. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 1027-1031.

Rossi, P. and de Crombrugghe, B. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 5590-5594.

Rossi, P., Karsenty, G., Roberts, A.B., Roche, N.S., Sporn, M.B. and de Crombrugghe, B. (1988) *Cell* **52,** 405-414.

Rossouw, C.M., Vergeer, W.P., du Plooy, S.J., Bernard, M.P., Ramirez, R. and de Wet, W.J. (1987) *J. Biol. Chem.* **262,** 15151-15157.

Roy, B. and Lee, A.S. (1995) *Mol.. Cell. Biol.* **15,** 2263-2274.

Roy, B. Li, W.W., Lee, A.S. (1996) *J. Biol. Chem.* **271,** 28995-29002.

Runge, S.W. (1992) *Biotechniques* **12,** 630-631.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning, A Laboratory Manual,* **2nd edition,** Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Sandmeyer, S., Gallis, B. and Bornstein, P. (1981) *J. Biol. Chem.* **256,** 5022-5028.

Sanger, F., Nicklen, S and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463-5467.

Sassone-Corsi, P. (1994) *EMBO J.* **13,** 4717-4728.

Schmidt, A., Rossi, P. and de Crombrugghe, B. (1986) *Mol. Cell. Biol.* **6,** 347-354.

Schmidt, A., Setoyama, C. and de Crombrugghe, B. (1985) *Nature* **314,** 286-289.

Schnieke, A., Harber, K. And Jaenisch, R. (1983) *Nature* **304,** 315-320.

Schreier, T., Friis, R.R., Winterhalter, K.H. and Trueb, B. (1988) *Biochem. J.* **253,** 381-386.

Seed, B. and Sheen, J.-Y. (1988) *Gene* **67,** 271-277.

Sgouras, D., Athanasiou, M.A., Beal, Jr., G.J., Fisher, R.J., Blair, D.G. and Mavrothalassitis, J. (1995) *EMBO J.* **14,** 4781-4793.

Sharf, R., Merazo, D., Azriel, A., Thornton, A.M., Ozato, K., Petricon, E.F., Larner, A.C., Schaper, F., Hauser, H. and Levi, B.Z. (1997) *J. Biol. Chem.* **272,** 9785-9792.

Sherwood, A.L. and Bornstein, P. (1990) *Biochem. J.* **265,** 895-897.

Sherwood, A.L., Bottenus, R.L., Martzen, M.R. and Bornstein, P. (1990) *Gene* **89,** 239-244.

Simkevich, C.P., Thompson, J.P., Poppleton, H. and Raghow, R. (1992) *Biochem. J.* **286,** 179-185.

Singh, H. (1993) *Methods in Enzymology* **218,** 551-567.

Singh, H., Clerc, R.G. and LeBowitz, J.H. (1989) *Biotechniques* **7,** 252-261.

Singh, H., LeBowitz, J.H., Baldwin, A.S. and Sharp, P.A. (1988) *Cell* **52,** 415-423.

Sinha, S., Kim, I-S., Sohn, K-Y., de Crombrugghe, B. and Maity, S.N. (1996) *Mol. Cell. Biol.* **16,** 238-337.

Sinha, S., Maity, S.N., Lu, J. and de Crombrugghe, B. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 1624-1628.

Sinha, S., Maity, S.N., Seldin, M.F. and de Crombrugghe, B. (1996) *Genomics* **37,** 260-263.

Skalnik, D.G., Strauss, E.C. and Orkin, S.H. (1991) *J. Biol. Chem.* **266,** 16736-16744.

Slack, J.L., Liska, D.J. and Bornstein, P. (1991) *Mol. Cell. Biol.* **11,** 2066-2074.

Slack, J.L., Liska, D.J. and Bornstein, P. (1993) *American J. Med. Genetics* **45,** 140-151.

Slack, J.L., Parker, M.I. and Bornstein, P. (1995) *J. Cell. Biochem.* **58,** 380-392.

Slack, J.L., Parker, M.I., Robinson, V.R. and Bornstein, P. (1992) *Mol. Cell. Biol.* **12,** 4714-4723.

Smith, A.A. (1989) *Ph. D. Thesis,* University of Cape Town.

Smith, B.D. and Marsilio, E. (1988) *Biochem. J.* **253,** 269-273.

Sobel, M.E., Dion, L.D., Vuust, J. and Colburn, N.H. (1983) *Mol. Cell. Biol.* **3,** 1527-2532.

Sokolov, B.P., Ala-Kokko, L., Dhulipala, R., Arita, M., Khillan, J.S. and Prockop, D.J. (1995) *J. Biol. Chem.* **270,** 9622-9629.

Sokolov, B.P., Mays, P.K., Khillan, J. and Prockop, D.J. (1993) *Biochemistry* **32,** 9242-9249.

Sonnenberg, A. (1993) *Curr. Top. Microbiol. Immunol.* **186,** 7-26.

Sorci-Thomas, M. and Kearns, M.W. (1991) *J. Biol. Chem.* **266,** 18045-18050.

Sorci-Thomas, M. and Kearns, M.W. (1995) *Biochim. Biophys. Acta* **1256,** 387-395.

Staudt, L.M., Clerc, R.G., Singh, H., LeBowitz, J.H., Sharp, P.A. and Baltimore, D. (1988) *Science* **241,** 577-580.

Struhl, K. (1985) *Biotechniques* **3,** 452-453.

Stuiver, I., Shimizu, Y. and Shimizu, N. (1991) *Biochem. J.* **278,** 369-373.

Sturm, R.A., Das, G. and Herr, W. (1988) *Genes Dev.* **2,** 1582-1599.

Sudbeck, B.D., Parks, W.C., Welgus, H.G. and Pentland, A.P. (1994) *J. Biol. Chem.* **47,** 30022-30029.

Suen, T-C. and Hung, M-C. (1991) *Mol. Cell. Biol.* **11,** 354-362.

Tabata, T., Takase, H., Takayama, S., Mikami, K., Nakatsuka, A., Kawata, T., Nakayama, T. and Iwabuchi, M. (1989) *Science* **245,** 965-967.

Tafuri, S.R. and Wolffe, A.P. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 9028-9032.

Tamaki, T., Ohnishi, K., Hartl, C., LeRoy, E.C. and Trojanowska, M. (1995) *J. Biol. Chem.* **270,** 4299-4304.

Thompson, J.P., Simkevich, C.P., Holness, M.A., Kang, A.H. and Raghow, T. (1991) *J. Biol. Chem.* **266,** 2549-2556.

Travers, H., French, N.S. and Norton, J.D. (1996) *Cell Growth Diff.* **7,** 1353-1360.

Treisman, R. (1986) *Cell* **46,** 567-574.

Tronche, F., Rollier, A., Bach, I., Weiss, M.C. and Yaniv, M. (1989) *Mol. Cell. Biol.* **9,** 4759-4766.

Tronche, F., Rollier, A., Sourdive, D., Cereghini, S. and Yaniv, M. (1991) *J. Mol. Biol.* **221,** 31-43.

Tsumaki, N. Kimura, T., Matsui, Y., Nakata, K. and Ochi, T. (1996) *J. Cell Biol.* **134,** 1573-1582.

Unemori, E.N., Amento, E.P., Bauer, E.A. and Horuk, R. (1993) *J. Biol. Chem.* **268,** 1138-1342.

van der Rest, M. and Garrone, R. (1991) *FASEB.J.* **5,** 2814-2823.

van der Westhuysen, D.R., Coetzee, G.A., Demasius, I.P.C., Harley , E.H., Gevers, W., Baker, S.G. and Seftel, H.C. (1984) *Arteriosclerosis* **4,** 238-247.

Vilen, B.J., Cogswell, J.P. and Ting, J.P.-Y. (1990) *Mol. Cell. Biol.* **11,** 2406-2415.

Vilen, B.J., Penta, J.F. and Ting, J.P.-Y. (1992) *J. Biol. Chem.* **267,** 23728-23734.

Vinson, C.R., LaMarco, K.L, Johnson, P.F., Landschulz, W.H. and McKnight, S.L. (1988) *Genes and Development* **2,** 801-806.

Vuorio, E. and de Crombrugghe, B. (1990) *Annu. Rev. Biochem.* **59,** 837-872.

Vuorio, T., Maity, S.N. and de Crombrugghe, B. (1990) *J. Biol. Chem.* **265,** 22480-22486.

Vuust, J., Sobel, M.E. and Martin, G.R. (1985) *Eur.J.Biochem.* **151,** 449-453.

Walsh, K. and Schimmel, P. (1987) *J. Biol. Chem.* **262,** 9429-9432.

Wang, Q. and Raghow, R. (1996) *Mol. Cell. Biochem.* **158,** 33-42.

Waurin, J., Mueller, C. and Schibler, U. (1990) **214,** 865-874.

Weiner, F.R., Giambrone, M-A., Czaja, M.J., Shah, A., Annoni, G., Takahashi, S., Eghbali, M. and Zern, M. (1990) *Hepatology* **11,** 111-117.

Westermarch, J., Ilvonen, E. and Kahari, V-M. (1995) *Biochem. J.* **308,** 995-999.

Wright, K.L., Vilen, B.J., Itoh-Lindstrom, Y., Moore, T.L., Li, G., Criscitiello, M., Cogswell, P., Clarke, J.B. and Ting, J.P.-Y. (1994) *EMBO J.* **13,** 4042-4053.

Wu, B.J., Williams, G.T., Morimoto, R.I. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 2203-2207.

Xiao, J.H., Davidson, I., Matthes, H., Garnier, J. and Chambon, P. (1991) *Cell* **65,** 551-568.

Xing, Y., Zang, S., Olesen, J.T., Rich, A. and Guarente, L. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 3009-3013.

Yang, B-S., Geddes, T.J., Pogulis, R.J., de Crombrugghe, B. and Freytag, S.O. (1991) *Mol. Cell. Biol.* **11,** 2291-2295.

Yoshioka, H., Greenwel, P., Inoguchi, K., Truter, S., Inagaki, Y., Ninomiya, Y. and Ramirez, F. (1995) *J. Biol. Chem.* **270,** 418-424.

Zhang, Y., Babin, J., Feldhaus, A.L., Singh, H., Sharp, P.A. and Bina, M. (1991) *Nucleic Acids Res.* **19,** 4555.