CHARACTERISATION OF *CIS*- AND *TRANS*-ACTING FACTORS THAT REGULATE THE HUMAN ALPHA 2(I) PROCOLLAGEN GENE

THESIS PRESENTED BY

AGATHA MARIPANYANE MASEMOLA

IN FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

MEDICAL BIOCHEMISTRY

IN THE

FACULTY OF HEALTH SCIENCES

UNIVERSITY OF CAPE TOWN

JULY 1999

University of Cape Town

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

CERTIFICATE OF SUPERVISOR

In terms of paragraph eight of the "General regulations for the degree of Ph.D.", I, as supervisor of the candidate **A.M. Masemola**, 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

I would like to express my sincere gratitude to Prof Iqbal Parker for his excellent supervision and guidance throughout the duration of this project;

Prof. Lutz Thilo and Dr Cynthia Sikakana for constructive criticism and encouragement;

Dr Virna Leaner and Mr Howard Donninger for continued assistance, advice and making my transition to Molecular Biology very pleasant; 'Matladi Ndlovu, Yolanda Giyose and Emmerentia Russel for support and friendship; Hajira Karjiker for all the oligonucleotides, laughs and cakes; Fundisile Mbangi, Fred Wamunyokoli, Erin Dietzsch, Robea Ballo, Bijou Nguza for all the discussions; past and present postgraduate students and staff in the Department of Medical Biochemistry for a wonderful working environment;

Mrs Sylva Schwager for all the help with the protein work, especially MALDITOF-ms analyses; Mr Hiram Arendse for his assistance in handling animals;

Ridwaan Majiet for all the help in the tissue culture lab.; Sarah Herman, Sam Groeneveldt, Tommy Clark, Robert Samuels, Sia Samuels, and Mbuyiselo Ntshidi for the lovely conversations and providing a clean working environment;

My family and friends for encouragement and support; my daughter, Lehlogonolo, for her patience and love; Vivian Morafo, Thembisile Kumalo and Keke Tloti for your friendship and support;

This work would not have been possible without the financial support I received from the Medical Research Council, AECI, UCT Medical Awards and the Cancer Association of South Africa.

This thesis is dedicated to all the women I have known since my childhood for their encouragement, support and love.

Forward with the emancipation of women!!

ABSTRACT

The differential expression of the $\alpha2(I)$ procollagen gene in normal and transformed human fibroblasts has been correlated with differential *in vitro* DNA-protein interactions on the basal promoter region between -100 and -67. A 23 bp region of the $\alpha2(I)$ procollagen promoter encompassing the G/CBE (CCTCCATTGG) and the CME (GGAGGCCCTTTT) has previously been shown to engage in specific DNA-protein interactions that determined the transcriptional activity of the promoter. The CME forms two distinct DNA-protein complexes that might be crucial in the regulation of the $\alpha2(I)$ procollagen gene in a cell specific manner. The hypothesis was, therefore, that depending on the protein that participates in complex formation with the CME, the gene would be activated or repressed. The objective of this study was to investigate the role of this 23 bp region in the regulation of expression of the $\alpha2(I)$ procollagen gene in transformed fibroblasts. In addition, the study sought to establish the role of the proto-oncogene *c-fos* in the regulation of expression of the $\alpha2(I)$ procollagen gene.

In contrast to previous observations, this study demonstrated that only one DNA-protein complex is formed on the CME and the second complex is a specific proteolytic cleavage of the product of the larger complex. Preparation of nuclear extracts in the absence of protease inhibitors, specifically leupeptin, resulted in the formation of a smaller complex, previously shown to bind the CME. The importance of this proteolytic fragment that still retains DNA binding activity is yet to be determined. In addition, the CME binding proteins were fairly ubiquitously expressed in both $\alpha2(I)$ collagen producing and non-producing cells.

CT-1 fibroblasts (transformed by γ-irradiation) synthesise over 80% of total $\alpha2(I)$ collagen produced by its untransformed counterpart (WI-38 fibroblasts), whereas the gene is down regulated in the human embryonic lung fibroblasts transformed with SV40 (SVWI-38 fibroblasts). These cell lines are therefore ideal for studying

regulation of $\alpha2(I)$ procollagen gene. To analyse the importance of the G/CBE and CME regions of the $\alpha2(I)$ procollagen gene promoter, point mutations were introduced by site-directed mutagenesis. Mutated promoter DNA was cloned into a p8CAT reporter vector and the activity of the promoter determined in transient transfection experiments.

Mutations introduced in the G/CBE region of the $\alpha2(I)$ procollagen promoter resulted in a 3-12 fold decrease in the activity of the promoter. The decrease was observed with both proximal (-343 bp) and basal (-107 bp) promoter constructs; a significant reduction in promoter activity was observed in both CT-1 and SVWI-38 fibroblasts. These results imply that the G/CBE region of the promoter is required for the activation of transcription of the $\alpha2(I)$ procollagen gene and therefore the factor that interacts with the G/CBE functions as a transcriptional activator. Previously, this factor was shown to complex with antibodies raised against the mouse CCAAT binding factor (CBF), suggesting that the protein belongs to the CBF family of transcription factors. Furthermore, these results demonstrate that the adjacent, upstream inverted GGAGG sequence is crucial for activation of the gene through the CCAAT binding element. The inhibition of promoter activity in constructs with a mutated G/CBE element was correlated with lack of protein binding to the mutated sequence as confirmed by electrophoretic mobility shift assays.

Transfection of $\alpha2(I)$ procollagen promoter constructs containing mutations in the CME region, however, resulted in a significant increase in promoter activity in both CT-1 and SVWI-38 fibroblasts. A much higher increase, 3-fold, was observed for the SVWI-38 cell line compared to a 1.5-fold increase observed for CT-1 fibroblasts. These results suggested that the factor that interacts with the CME functions as a repressor of the $\alpha2(I)$ procollagen gene. Interestingly, the promoter activity in SVWI-38 fibroblasts transfected with mutated CME constructs was similar to that observed in CT-1 fibroblasts transfected with the wild type promoter construct. An interesting observation was that repression of the $\alpha2(I)$ procollagen gene via the CME required upstream elements since transfection of the basal mutated promoter did not result in increased promoter activity. From these results, it can be concluded that the CME binding protein is involved in cell-specific repression of the $\alpha2(I)$ procollagen gene and that the mechanism of repression appears to be dependent on the presence of upstream elements.

Mutations in the G/CBE and CME pointed out the significance of these elements in the expression of the $\alpha2(I)$ procollagen gene and since a number of studies have characterised the mouse CCAAT binding protein, this study focused on purification and identification of the CME binding protein(s). Purification was performed by conventional biochemical techniques using heparin-agarose and sequence-specific DNA affinity chromatography, as well as separation on SDS-polyacrylamide gels. Two cycles of DNA affinity chromatography yielded two polypeptides with apparent molecular weights of 50 and 67 kDa. Automated N-terminal sequencing of the polypeptides indicated that they were blocked and therefore no sequence could be obtained. In addition, these polypeptides failed to raise an immune response in mice and rabbits. Subsequently, polypeptides were digested with trypsin *in situ* in polyacrylamide gels and the eluted peptides were analysed by MALDITOF-mass spectrometry. The mass:charge ratios (*m/z* ratios) obtained were used to search the database using a mass tolerance of 1.5 Da and only one hit was obtained. The match obtained was that of a mouse zinc finger protein of which not much is known, except that it might be a transcription factor. This result supports previous observations of Collins *et al* (**J Cell Biochem** 1998, 70: 455-467) that complex formation requires the presence of zinc. The primary structure of the CME binding protein remains to be determined.

Transformation of fibroblasts is normally accompanied by changes in the expression of extracellular proteins, including type I procollagen. Although CT-1 fibroblasts, show very little change in $\alpha2(I)$ procollagen gene expression, the *c-fos* gene is drastically down-regulated. This study sought to establish if there is any relationship between the unusually high levels of the $\alpha2(I)$ procollagen gene in this transformed cell line and failure of the cells to stimulate *c-fos* expression in response to serum.

CT-1 fibroblasts that overexpressed wild type Fos were established and changes in the expression of the $\alpha2(I)$ procollagen gene were measured. Overexpression of Fos down-regulated the $\alpha2(I)$ procollagen gene, which was not due to increased turnover of the $\alpha2(I)$ procollagen mRNA. Analysis of promoter activity showed that the promoter and first intron, which has been reported to contain negative regulatory elements, did not harbour any Fos-responsive elements. The -343 bp and the -2300 bp promoter constructs were transactivated in cells overexpressing Fos. Thus, although overexpression of Fos resulted in a significant decrease in the levels of the $\alpha2(I)$ procollagen RNA, it does not involve the region between -2300 bp and +1800 bp of the $\alpha2(I)$ procollagen gene. Furthermore, there was no change in the stability of the message, indicating that constitutive expression of Fos did not activate a factor that could play a role in altered turnover of the $\alpha2(I)$ procollagen mRNA. It is therefore possible that constitutively high levels of Fos may trigger the expression of a number of other genes, which have a negative impact on the expression of the $\alpha2(I)$ procollagen gene.

CHAPTER 1

INTRODUCTION

The collagens belong to a family of structural glycoproteins that are crucial in maintaining the structure of most organs; they are the most abundant components of the extracellular matrix (ECM). They are important in a number of physiological processes such as transmission of light, cell adhesion and movement, homeostasis, tissue remodelling and wound healing (1). They are classified as either fibrillar (e.g. type I collagen) or non-fibrillar (e.g. type IV collagen). Type I collagen is the most abundant protein of the body; it is synthesised in bone, cornea, skin, ligaments and tendons predominantly by osteoblasts or fibroblasts, although odontoblasts, mesenchymal and smooth muscle cells have also been shown to synthesise the protein (2-4). Type I collagen is a triple helical molecule composed of $\alpha 1(I)$ and $\alpha 2(I)$ chains in a 2:1 ratio; this 2:1 stoichiometry is also observed for the steady state mRNA levels in human, mouse and chicken fibroblasts (1-4). The expression of the two type I procollagen genes has been shown to be coordinately regulated and it is thought to be achieved through the presence of common *cis*-acting regulatory elements in the promoter and/or the first intron of these genes (5-7).

The 2:1 ratio of the $\alpha 1(I)$ and $\alpha 2(I)$ chains in cells allow formation of functional triple helical structures, and this ratio can be altered under certain pathological conditions. Alteration in the expression of type I procollagen genes occurs during embryonic development, wound healing and certain pathological conditions such as scleroderma, cancer and fibrotic diseases of the liver, kidney and lung (4,8-9). The collagen synthesised in normal livers is predominantly type I, however, in cirrhotic livers, equal amounts of type I and III collagen are synthesised (10). Type I collagen from cirrhotic livers has been shown to be composed of two $\alpha 2(I)$ and one $\alpha 1(I)$ chains and this has been associated with a collapse in the architecture characteristic of cirrhotic livers. Fibroblasts derived from scleroderma patients, on the other hand, have been shown to exhibit a 2-7 fold higher level of α1(I) procollagen mRNA, which accounts for the 2-7 fold increased synthesis of collagen (11).

Transformation of cells by either DNA or RNA viruses, oncogenes, chemical carcinogens or ultraviolet radiation has been shown to affect the expression of the gene in a wide variety of cell lines (section 1.1). In addition, type I procollagen gene expression is also modulated by hormones, cytokines and vitamins (section 1.2.2). Most of these changes occur at the level of transcription of the gene, implying that the modulators either activate or inactivate transcription factors which regulate the expression of these genes. A number of regulatory *cis*-acting elements, and their cognate *trans*-acting factors, have been identified both upstream and downstream of the transcription start site (section 1.2.1). A full understanding of the mechanisms involved in altered expression of the type I procollagen genes relies on the purification and characterisation of the transcription factors that interact with the regulatory elements on the gene. Since degradation of ECM components such as type I collagen is important in tumour invasion and metastasis, these studies will lead to a better understanding of these processes.

1. The type I procollagen genes

Type I collagen molecule contains two α1(I) and one α2(I) chains that forms a semiflexible rod that assembles into microfibrils (3). The molecules overlap to produce the characteristic 67-nm periodicity of collagen fibrils. These properties arise from the triple helical conformation formed from the uninterrupted sequence of $(\text{glycine-X-Y})_{338}$ triplets where X is often proline and Y is often hydroxyproline. The helix is formed by glycine residues whilst hydroxyproline residues stabilise the helix. The protein is synthesised as a procollagen molecule, secreted and the amino and carboxy propeptides are removed by specific proteinases to yield type I collagen, which self-assembles into fibrils. The 18 kb COL1A1 gene encodes the α1(I) chain located on chromosome 17q21.3-q22, whilst the α2(I) chain is encoded by the 38 kb COL1A2 gene located on chromosome 7q21.3-q22. Each gene contains 52 exons, of which exons 7 to 48 encode the main triple helical domain of 1014 amino acids and 338 glycine-X-Y repeats (3).

1.1 Cellular transformation and type I procollagen gene expression

The expression of type I and other procollagen genes has been reported to be altered by transformation with chemical carcinogens, ultraviolet radiation, RNA/DNA viruses and oncogenes. Chick embryo fibroblasts (CEF) transformed by the Rous sarcoma virus (RSV), an RNA tumour virus that codes for the *v-src* oncogene, showed a 10-fold reduction in type I procollagen mRNA levels. To investigate the mechanism of the observed decrease, the study employed temperature-sensitive mutants in the transforming *src* gene and changes in the synthesis of type I procollagen mRNA was determined at 35 °C and 41 °C. From these studies it was concluded that the decrease in the levels of type I procollagen mRNA was a result of a reduction in the half-life of the transcript from 18 hours to about 10 hours, although the decrease could also reflect the direct effect of *src* on procollagen RNA stability (12). In a study by Parker *et al* (13) transformation of human embryonic lung fibroblasts by the simian virus 40 (SV40) resulted in a shut down in the expression of the α2(I) procollagen gene, an effect which was mediated at the transcriptional level. Whilst transformation of fibroblasts is often associated with changes in the synthesis of type I collagen, transformation of 11-day old chick embryo vertebral cartilage or vertebral chondrocytes with RSV resulted in a dramatic decrease in the synthesis of type II collagen and an increase in type I collagen and fibronectin-like molecules (12,14,15). Similarly, immortalisation of rabbit articular chondrocytes by SV40 (SVRAC) resulted in the loss of the ability to synthesise type II collagen but these cells synthesised predominantly type I collagen, the SVRAC displayed an apparently irreversibly dedifferentiated phenotype (16). It would therefore appear that transformation is modulated by host cell-specific factors. In contrast, a study by Sandmeyer *et al* (17) demonstrated very little changes in the synthesis of collagen in rat embryo fibroblasts transformed with SV40, although transformation of the same cells with the Polyoma virus (DNA virus) and RSV resulted in a 30-48 % and 12-25

% decrease in the synthesis of collagen, respectively. These observations suggested that the extent of transformation determines the degree to which the synthesis of collagen is affected. It therefore appears that transformation by RNA or DNA viruses will, in one way or another, lead to changes in the expression of type I procollagen genes in a cell specific manner; whilst a decrease often accompanies transformation of fibroblasts, an increase has been observed consistently in the case of chondrocytes.

In a study by Schmidt *et al* (18), transformation and overexpression of *v-mos*, the transforming gene of the Moloney murine sarcoma virus that codes for a protein serine/threonine kinase, resulted in a decrease in the transcriptional activity of the α2(I) procollagen gene. Interestingly, *mos* does not require to be altered by mutations to become a transforming gene, elevated expression of the mouse or chicken *mos* proto-oncogene is sufficient to cause transformation of NIH3T3 fibroblasts (19). Whilst overexpression of wild type *mos* results in transformation, immortalisation by the p53 tumour suppressor gene requires overexpression of the mutant p53 protein and immortalised cells have reduced levels of α1(I) procollagen mRNA, but, the effects appear to be indirect (20). A more direct effect was obtained in cells transformed by activated *ras* whereby the reduction in the α1(I) procollagen mRNA was accompanied by altered cell shape and cell proliferation (20). A study by Slack *et al* (21) demonstrated that transformation of Rat-1 fibroblasts with oncogenic *ras* reduces the expression of both α1(I) and α2(I) procollagen genes, the decrease in the α1(I) procollagen gene was correlated with both an unstable procollagen message and increased activity at the intronic AP-1 regulatory element (21). Similarly, transformation of Rat-1 fibroblasts by the *v-fos* oncogene resulted in a decrease the levels of both α1(I) and α2(I) procollagen mRNA; this effect could be reversed with the use of revertant cell lines that expressed a functional *c*-Fos oncoprotein (22). NIH3T3 fibroblasts transformed by *v-fos* (FBR virus), however, showed a selective increase in the levels and synthesis of the α1(III) procollagen mRNA and protein (23). Transformation of normal diploid fibroblasts is often accompanied by changes in the morphology of the cells, however, in the case of human corneal fibroblasts expressing the E6 and E7 proteins of the human papilloma virus type 16, changes in cell morphology were observed in some of the E6/E7 expressing cell lines and all cells produced type I, III and V collagens at levels similar to the parental corneal diploid fibroblasts (24). Thus it is possible to expand the lifespan of cells without altering either the growth rate or the expression of collagen in cells expressing viral proteins. Although this is true for E6 and E7 human Papilloma viral proteins, it would appear that most viral proteins, especially oncogenes, result in changes in the expression of procollagen genes.

The synthesis of ECM proteins in human osteosarcoma cells and their viral and chemical transformed phenotypes were significantly reduced as a result of decreased levels of $\alpha1$(I) procollagen mRNA (25). In addition to a decrease in cellular mRNA, the synthesis of the collagen receptor, $\alpha\beta1$ integrin, could be induced with cytokines in the chemically transformed cells. Some of the studies that have reported altered expression of type I procollagen genes after chemical transformation include transformation of mouse 3T3 cells with a phorbol ester where a 3-5 fold decrease in the levels of ECM proteins including type I procollagen mRNA was observed, interestingly, the decrease occurred in the absence of protein synthesis and was accompanied by the induction of *c-myc* and *c-fos* (26). In W8 cells, rat liver epithelial-like cells transformed chemically with 2-N-(acetoxyacetyl)aminofluorene, there was no $\alpha2$(I) procollagen transcription, this effect was shown to result from DNA methylation of the promoter of this gene (27,28). Instead, these cells produced an $\alpha1$(I) collagen trimer, thus implying that chemical transformation interfered with the co-ordinated control of the type I procollagen genes. Transformation of Syrian hamster embryo fibroblasts with 4-nitroquinoline-1-oxide resulted in a 30 % reduction in total collagen synthesis and the synthesis of the $\alpha1$(I) collagen chain was completely suppressed (29). Treatment of rats with 3-methylcholanthrene lead to fibrosarcoma that was characterised by excessive production of type V, type I trimer and type III collagens, however, levels of the total collagen were decreased (30). Chemical transformation of fibroblasts results in an overall decrease in the expression of ECM proteins but the mechanisms range from induction of early immediate genes such as *c-fos* and *c-myc* independent of new protein synthesis, altered expression of the collagen receptors and inhibition of transcription by methylation of regulatory elements in the gene.

Exposure to high doses of ultraviolet radiation is known to alter the expression of a number of genes resulting in cancers such as skin cancer and some haemopoietic cancers. Chen and Davis (31) demonstrated that UV irradiation of hepatic stellate cells results in increased endogenous $\alpha1(I)$ procollagen mRNA levels due to stimulated transcriptional activity of the $\alpha1(I)$ procollagen gene as determined from transient transfection experiments. Human embryonic lung fibroblasts transformed by γ-radiation, CT-1 fibroblasts, were able to grow in soft agar and produce tumours in nude mice but collagen synthesis in these cells was reduced by only 20 % compared to the parental cell line (32,33). In a separate study by Hatamochi *et al* (34), transformation of normal diploid human foetal fibroblasts with Co-60 γ-rays resulted in a 30 % decrease in the steady state collagen protein and a 6-fold decrease in the rate of collagen synthesis. This was associated with a 20 % decrease in type I collagen and a 9 % decrease in type III collagen. Radiation might be one of the potent transforming agents, however, its effect on type I collagen synthesis appears to be less dramatic.

In addition to transformed cells exhibiting altered type I procollagen gene expression, changes in the expression of procollagen genes may also result from genetic disorders associated with mutations in the type I procollagen genes. Diseases that result from mutations in the procollagen genes include osteogenesis imperfecta [$\alpha1(I)$ and $\alpha2(I)$ mutations], some forms of osteoporosis [$\alpha2(1)$ mutations], Ehlers-Danlos syndrome type VIIA, VIIB [$\alpha2(I)$ mutations], chondrodysplasias [$\alpha1(II)$ mutations], osteoathritis [$\alpha1(II)$ mutations] and Alport syndrome, autosomally inherited forms [$\alpha3(IV)$ and $\alpha4(IV)$] or X-linked form [$\alpha5(IV)$] (3-4). It is possible to introduce point mutations in the affected exons to simulate some of these disease states and the effect of these mutations studied in transgenic animals. With this approach, a link between certain diseases with specific mutations has been established, however, it remains essential to establish transcriptional mechanisms that regulate the expression of these genes to allow therapeutic interventions at the transcriptional level (3-4).

1.2 Regulation of type I procollagen gene expression

Characterisation of the regulatory elements within the α1(I) and α2(I) procollagen genes will enhance the understanding of the molecular mechanisms controlling type I procollagen gene expression in normal and pathological conditions. A number of studies have identified different *cis*-acting elements in the promoter and first intron of the type I procollagen genes. Most of the data gathered to date involved identification and characterisation of regulatory elements in the mouse gene; there is about 86 % sequence similarity in the proximal promoters (-400) of the mouse and human α2(I) procollagen genes.

1.2.1 *Cis*-acting elements involved in the regulation of α2(I) procollagen gene expression.

Multiple transcriptional activators have been shown to regulate the expression of eukaryotic genes, many of these activators bind to specific sequences in the regulatory elements of genes to promote initiation of transcription by RNA polymerase II through interactions with other components of the transcriptional apparatus. The activity of a specific promoter is determined both by its architecture, consisting of the type of control elements in the promoter and the context in which they appear, and whether the proteins that recognise these elements are modified or not.

The 3500 bp of DNA located immediately upstream of the transcription start site of the α2(I) procollagen gene contains all the sequences necessary for the cell and tissue specific transcription of the gene (35). Three areas of Sp1 binding were shown to be located between -568 and -453 of the 5′ flanking region of the murine α2(I) procollagen promoter (36). Recently, two additional Sp1 binding sites located at -371 to -351 and -690 to -613 have been identified in the murine α2(I) procollagen promoter (36). Mutation of both regions resulted in an increase in promoter activity in transfected rat hepatic stellate cells, indicating that these regions contain negative regulatory elements. The -371 to -351 element was also shown to bind AP-2 resulting in a decrease in Sp1 binding, further studies demonstrated that low AP-2 expression inhibited the stimulatory effect of a transfected Sp1 expression vector on the α2(I) procollagen promoter in Drosophila cells whilst high expression of AP-2 enhanced the stimulatory effect of Sp1 (36). The region between -690 to -613 was shown to bind YY1 and it has a positive effect on Sp1 binding. Thus it would appear that the two negative regulatory elements in the murine α2(I) procollagen promoter involve interaction of Sp1 with AP-2 and YY1 transcription factors to inhibit activation of the gene (36). Downstream of this negative regulatory region is located a positive *cis*-acting element between -420 and -399 in the mouse α2(I) procollagen promoter which has been shown to bind COLF1, the DNA binding activity of this factor is 4x higher in NIH3T3 fibroblasts compared to cells transformed by either *v-mos* or *v-ras*, and in lymphocytes and epithelial cells (36). The decreased DNA binding activity of this factor in extracts from lymphoid and epithelial cells was correlated with the inability of these cells to synthesise type I collagen, implying that the factor could have a positive effect on the expression of type I procollagen in NIH3T3 fibroblasts. Furthermore, it would appear that COLF1 is down-regulated or inhibited in cells transformed by *v-mos* or *v-ras* since nuclear extracts from these cells showed a significant reduction in DNA binding activity; there were no differences in the DNA binding activity of the CCAAT binding factor (CBF) from these cells indicating the specificity of the effect (37).

Cis-acting DNA elements that direct high activity and tissue specific transcription of the human COL1A2 promoter are located between nucleotides -376 and -108 as illustrated in figure 1.1. A 131 bp region between -378 and -255 of the proximal human α2(I) procollagen promoter has been shown to contain a transforming growth factor β responsive element (TbRE) and three Sp1/Sp3 recognition sites are located between -313 and -286 (38-40). Subsequent studies have shown that these Sp1/Sp3 sites are required for the early response of the human α2(I) procollagen gene to

TGFβ1 (41). In a conflicting report, Chung *et al* (42) demonstrated that the TbRE is located between -265 and -241 of the human α2(I) procollagen promoter, this 25 bp fragment contains one Sp1 binding site (between -304 and -273) and an AP-1 like consensus sequence (between -258 and -244), only the AP-1 element, however, was shown to mediate TGFβ effects. The element between -265 and -241 of the human α2(I) procollagen promoter was also shown to mediate TNFα effects; TNFα is antagonist to the effects of TGFβ on the expression of the human α2(I) procollagen gene (42). Studies using transient transfection of human dermal fibroblasts with a number of COL1A2 5′ deletions were carried out to characterise the TNFα response element (TaRE) and it was evident that a 2 bp substitution mutation in the NF-κB1 like binding site between -251 and -241, eliminated the inhibitory effect of TNFα (43). This element was shown to be recognised by antibodies against NF-κB (p50/p65) and not antibodies against AP-1, nuclear factor of activated T cells (NF-AT) or Jun family members (43). These studies suggest that in fibroblasts TNFα activates and initiates the nuclear translocation of NF-κB which can interact with the α2(I) procollagen promoter at the NF-κB like sequence located between -265 and -241, downstream of the AP-1 binding site, with a consequent decrease in the transcription of the α2(I) procollagen gene (43). Also interesting is the observation that this region harbours a CAGA sequence (between -250 and -247) although the factor that binds here has yet to be identified (1). The mouse α2(I) procollagen promoter contains a nuclear factor 1 (NF-1) recognition sequence between -315 and -295, which is absent in the human promoter (44,45). Interestingly the region between -350 and -300, overlapping the NF-1 site, of the mouse α2(I) procollagen promoter is required for TGFβ induced stimulation of the gene (46).

A pyrimidine-rich sequence located between -178 and -155 has been shown to bind a factor with a negative effect on the activity of the human α2(I) procollagen promoter activity; mutations introduced in the TCCCCC motif between -164 and -159 increased promoter activity by 4 fold in human foreskin fibroblasts (47). The authors suggested that this factor may be related to either c-Krox, a transcription factor which recognises GC-rich sequences or BFCOL transcription factors, these factors have been shown to interact with the corresponding *cis*-regulatory element in the mouse α2(I) procollagen promoter (7,48). Studies by Ihn and Trojanowska (40) have demonstrated that this region binds Sp3 with the same potency as Sp1. Thus it appears that both Sp1/Sp3 can act as either transcriptional activators or repressors of the human α2(I) procollagen gene. Interestingly, a c-Krox binding consensus sequence has been mapped to a GC-rich region in the mouse α2(I) procollagen promoter between -165 and -155, this factor has been shown to have a negative effect on the transcriptional activity of the mouse α2(I) procollagen promoter (7). In contrast, Hasegawa *et al* (48) have demonstrated that a protein factor, BFCOL, is required for the transcriptional activation of the mouse α2(I) procollagen gene and it binds to the promoter between -180 and -152, the recognition sequence is characterised by a 5′-CCTCCCCCCTC-3′ motif between -176 and -152 and mutations within this sequence abolished binding of BFCOL and inhibited transcriptional activity. Adjacent to the BFCOL binding site in the mouse α2(I) procollagen promoter is a repressor binding region between nucleotides -154 and -145 termed an inhibitory factor 1 (IF-1) binding region (5). IF-1 has been shown to repress the activity of the mouse promoter in transient transfections experiments in NIH3T3 cells; a 3 bp substitution mutation in this region of the promoter resulted in a 4 fold increase in promoter activity. It would therefore appear that the region between -180 and -140 of the human and mouse α2(I) procollagen promoter is characterised by both positive and negative regulatory elements, however, deletion of the region between -180 and -136 decreased promoter activity implying that the overall activity of this element is that of an enhancer and it therefore binds a number of transcription activators (48).

A report by Higashi *et al* (49) identified a proximal element located between -161 and -125 within the human α2(I) procollagen promoter that mediates the transcriptional repression by interferon-γ (IFNγ), termed the IFNγ response element (IgRE); this element is distinct from either the TbRE or TaRE located between -265 and -241.

Competition experiments with oligonucleotides spanning discrete segments of this promoter region mapped the binding element within a distinctive pyrimidine-rich sequence and point mutations revealed that this element plays a crucial role in both the IFNγ response and basal activity of the proximal promoter (49). Although the factor that binds this element has not yet been identified, UV-cross linking experiments indicated that the DNA-protein complexes at the IgRE have molecular weights of 55 and 30 kDa, which corresponds to proteins of approximately 30 and 6 kDa, respectively (49).

Additional Sp1/Sp3 binding sites with a positive effect on transcription were mapped further downstream between -128 to -123 region of the human α2(I) procollagen promoter (39,40). This region is characterised by a novel positive *cis*-acting element containing a TCCTCC motif and using base substitution mutagenesis and DNA binding in the presence of antibodies against Sp1 and Sp3, it was evident that this element was required for the transcriptional activation of the α2(I) procollagen promoter by Sp1/Sp3 (39,40). The region immediately 100 bp upstream of the transcription start site is often associated with basal transcription, however, it contains transcription factor binding sites. The region between -84 and -80 of both the human and mouse α2(I) procollagen promoter contains an inverted CCAAT sequence, which has been shown to bind a protein belonging the CCAAT binding factor (CBF)/NF-Y family of proteins (1,6,7,13,47,50-57). The mouse CBF has been extensively characterised; the protein is a heterotrimer of CBF-A, CBF-B and CBF-C subunits and it requires all subunits for DNA binding and therefore its transcriptional activity depends on the correct assembly of these subunits (53,58). It has been shown that the histone-fold motifs of CBF-A and CBF-C interact with each other and then with CBF-B to form the CBF heterotrimer, which in turn binds to a specific DNA sequence (58). Studies using base substitution mutations and transient transfection experiments demonstrated that the CCAAT motif binds a protein that is involved in transcriptional activation of a number of genes (reviewed in 53). Adjacent to the CCAAT sequence is a novel collagen modulating element (CME) present only in the human α2(I) procollagen promoter between -80 and -72 (13,51,56,57,59). The protein that recognises the CME has been suggested to have a negative effect on the transcription of the α2(I) procollagen gene. In a recent report, Sengupta and Smith (60) demonstrated that the region between -41 and +54 of the mouse α2(I) procollagen gene is inactivated by methylation in transient transfection assays. A DNA-protein complex was identified in protein extracts prepared from nuclei of chemically transformed rat epithelial-like cells (W8) in the presence of 500 mM NaCl, this complex was shown to have a high affinity for methylated sequences and it did not involve the TATA sequence between -30 and -25. Mutations at the methylation sites (+7, +23) in the first exon increased DNA-protein complex formation and inhibited α2(I) procollagen transcription as confirmed by transient transfection and *in vitro* transcription assays. Taken together these studies indicate that although type I collagen is one of the abundant proteins in the ECM, its expression is regulated by a number of transcription factors which interact directly with the promoter of the gene; both negative and positive elements are present in the promoter and through co-ordinated functions, the expression of type I procollagen genes is kept in balance in type I collagen producing cells; and perhaps some of these promoter elements might be important in cell specific expression of the α2(I) procollagen gene.

Although most reports focused on transcription factor binding sites upstream of the transcription start site, a number of other regulatory sequences have been identified in the first intron of the α2(I) procollagen gene. The mouse α2(I) procollagen gene has been shown to contain two distinct *cis*-acting elements between +948 and +1040 that are essential for the transcriptional activity (61,62). The region between +871 and +880 contains a GT box which binds an Sp1 like protein and the +983 to +1007 region contains a CIBF-I binding site, which is related to a protein(s) that binds to the rat *c-mos* enhancer sequence, 5′-TGT TTT AA-3′ (62). The first intron of the human α2(I) procollagen gene, on the other hand, has an overall negative effect on the transcriptional activity of the α2(I) procollagen promoter in transient transfection experiments (63). Some of the sequences with potential transcription factor binding activity include 1) a NF-1 like consensus sequence between +192 and +205, 5′ to the

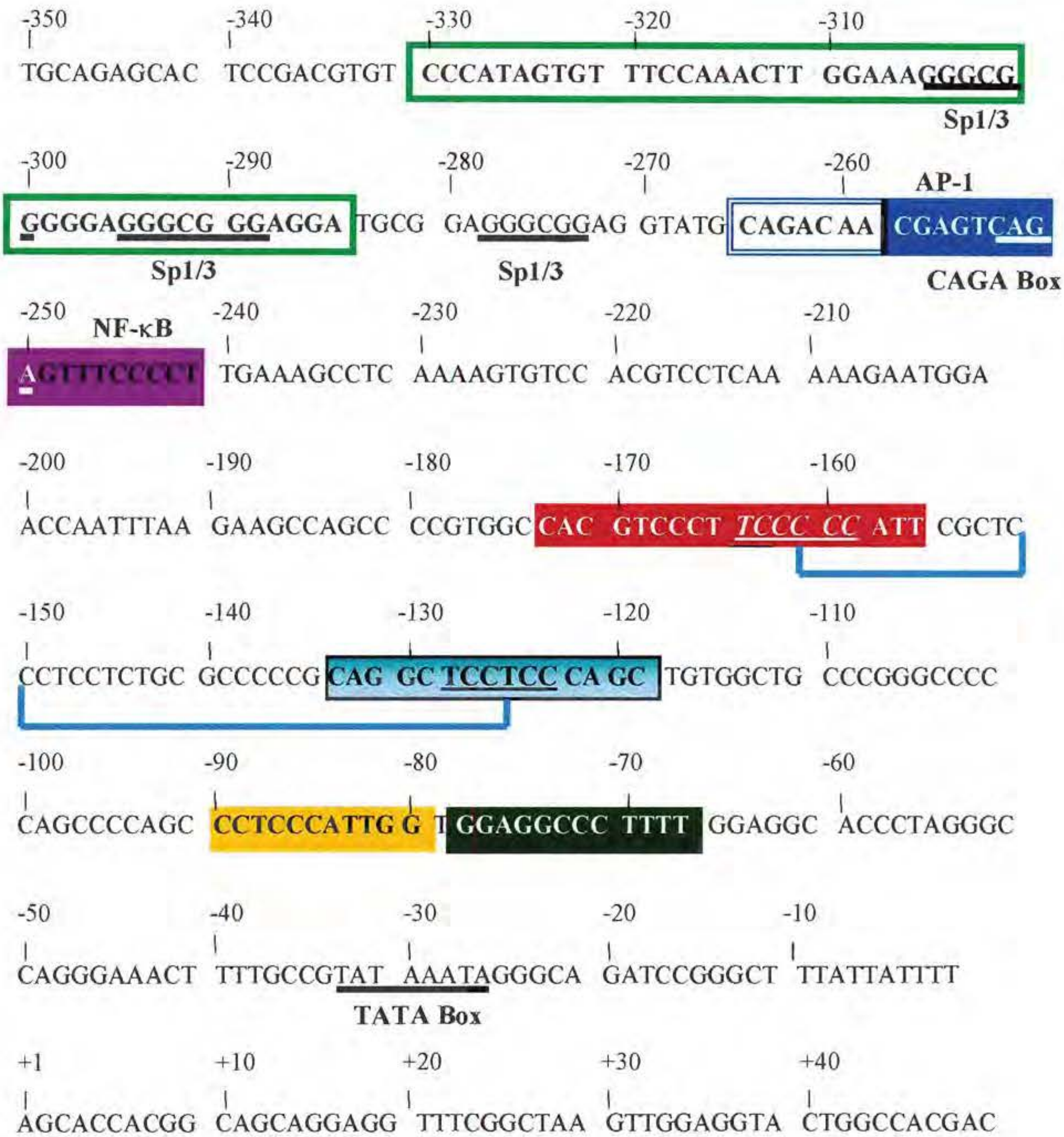
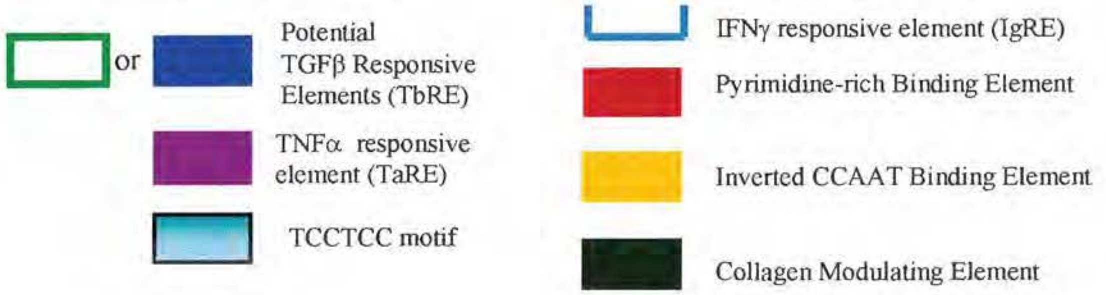


Figure 1.1: Nucleotide sequence of the proximal promoter region of the human α2(I) procollagen gene. Regulatory elements are indicated below (see text for references).



intron/exon junction, 2) two putative AP-1 binding sites at +1112 and +1085, 3) CCCTG pentanucleotide repeats in the first exon (a single repeat) and the first intron (five repeats), and 4) an alternating GT stretch located between +1421 and +1460, this sequence has the potential to form Z-DNA. Interestingly, the intronic AP-1 site has been shown to partly mediate the transcriptional repression of the mouse α2(I) procollagen gene in Rat-1 fibroblasts transformed by mutant H-*ras* (64). Figure 1.2 summarises some of the transcription factor binding sites identified in the human α2(I) procollagen gene between -400 and +1800.

Although a number of transcription factor binding sites have been identified in both the mouse and human α2(I) procollagen genes, a better understanding of regulation of the gene will depend on purification and molecular cloning of the factors that bind to these elements.

1.2.2 Modulators of type I procollagen gene expression

The expression of type I procollagen genes occurs as a normal physiological event, but expression can be enhanced or repressed by modulators such as cytokines (TGFβ1, TNFα and IFNγ), vitamins and glucocorticoids. Most of these modulators are secreted under certain pathological conditions; the release of cytokines, which induce the expression of a number of genes that modulate the expression of connective tissue proteins by connective tissue cells, is a major factor in the pathophysiology of fibrosis. Some of these modulators have been shown to alter the levels of type I procollagen mRNA and to a certain extent, some of their response elements have been mapped in the α1(I) and α2(I) procollagen genes. Although the focus of this study is on the regulation of the expression of the α2(I) procollagen gene, this section will also describe findings reported for the α1(I) procollagen gene.

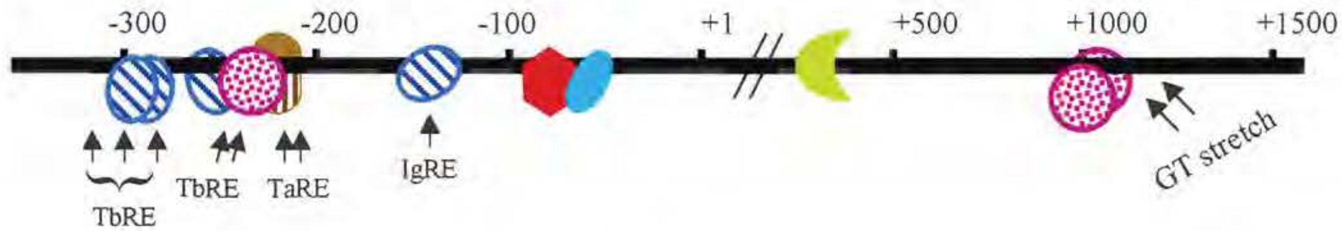
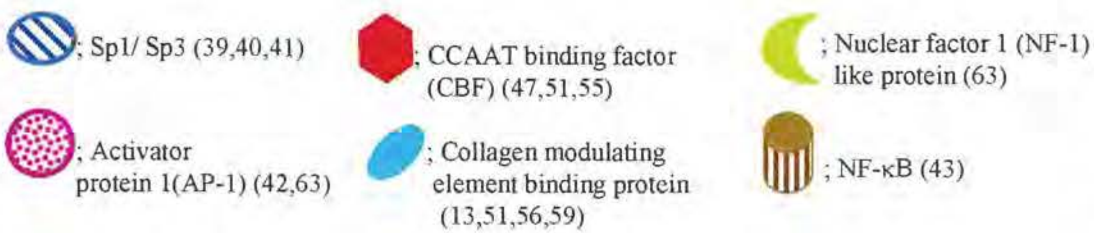


Figure 1.2: ***Cis*-acting elements and their cognate transcription factors that bind to the human α2(I) procollagen gene.** Data taken from references indicated in brackets.



TbRE~ Transforming growth factor β1 response element (38,41,42)

TaRE ~ Tumour necrosis factor α response element (42,43)

IgRE ~ Interferon γ response element (49)

1.2.2.1 The effect cytokines on the expression of type I procollagen genes

The formation of the ECM is regulated by a number of cytokines, most of which are secreted by mononuclear inflammatory cells. These cytokines have also been implicated in the onset and progression of certain pathological conditions such as scleroderma, keloids and pulmonary fibrosis. TGFβ is a member of the epidermal growth factor (EGF) family of proteins that modulate a variety of cellular activities and morphogenetic processes, including ECM remodelling (65,66). It activates signal transduction pathways leading to cell proliferation through the interaction with the cell surface EGF receptor (EGFR). TGFβ stimulates the expression of ECM proteins such as procollagen types I, III, VI and X, fibronectin, matrix glycoproteins such as osteopontin, osteonectin and thrombospondin, and proteoglycans such as decorin (65). This is achieved by increasing the steady state mRNA levels in a number of cell types, including avian, rodent and human fibroblasts from skin, lung, bone marrow, hepatic fat-storing cells, rodent osteoblasts, smooth muscle cells and hepatocytes (6). The increase in type I collagen production and deposition in the ECM is either associated with a stable transcript or an increase in the rate of transcript synthesis depending on the cell line under study. Ironically, TGFβ has been implicated in the pathogenesis of scleroderma, keloids and fibrosis due to excessive production of collagen (66,67). A TGFβ response element has been identified in both the mouse and human α2(I) procollagen promoters, the mouse TbRE appears to require an NF-1 consensus sequence whilst the human TbRE contains a number of Sp1 consensus sequences, although as mentioned above, the human TbRE could also involve the AP-1 transcription factor (38, 41,42).

The effects of tumour necrosis factor α (TNFα) are variable, but in the majority of studies, treatment of cells with TNFα results in decreased type I collagen production (6). The effect of TNFα is also dependent on the state of differentiation of cells, for example, a decrease in the α2(I) procollagen gene is observed in murine 3T3-L1 fibroblasts that are in the preadipocyte stage, while differentiation of these cells into adipocytes is accompanied by a significant increase in α2(I) procollagen gene transcript (68). The effect of TNFα on expression of the human $\alpha2(I)$ procollagen was thought to be mediated through the same *cis*-acting element (AP-1) between -265 and -241 in the proximal promoter identified as the TbRE, however, in a recent report the TNFα response element (TaRE) was shown to involve an NF-κB binding site within the -265 and -241 region of the human $\alpha2(I)$ procollagen promoter (38,42, 43). In pulmonary fibrosis there is an increase in the levels of TNFα that is derived mainly from macrophages and it has been shown to increase collagen deposition, an effect that could be inhibited with antibodies against TNFα (69).

Interferon-γ (IFNγ) is the primary negative regulator of type I procollagen gene expression through both transcriptional and post-transcriptional mechanisms (6). Its effects have been reported from *in vitro* studies using human fibroblast, chondrocytes, osteoblasts and rheumatoid synovial cells in culture, and *in vivo* by direct administration of IFNγ to mice (6). All these studies have demonstrated a decrease in the production of type I collagen and the steady state mRNA levels were down by 25-50 %. The deposition of collagen in wounds is inhibited by treatment with IFNγ and the fibroblasts appear smaller and less active (6). Recently, a study by Higashi *et al* (49) identified an IgRE in the proximal promoter of the human $\alpha2(I)$ procollagen gene, this element was shown to be distinct from the TbRE or TaRE.

In addition, interleukins, IL-1 and IL-10, have been reported to modulate the synthesis of collagen and matrix metalloproteases in cultured fibroblasts (6). The effects of IL-1 are tissue specific; an increase in the steady state mRNA levels of procollagen gene is observed in cultured human skin, synovial and lung fibroblasts and in chondrocytes. In contrast, a decrease in collagen synthesis occurs in a murine osteoblastic cell line. Although the mechanism is still not clear, it appears to involve the AP-1 transcription factor (6).

1.2.2.2 Modulation of type I procollagen gene expression by hormones

Glucocorticoids have inhibitory effects on the formation of bone, they inhibit cell replication and the capability of cells to synthesise type I collagen (6). In addition, they downregulate procollagen gene expression in osteoblasts. Dexamethasone has been shown to accelerate the rate of $\alpha1(I)$ procollagen mRNA turnover resulting in an overall reduction in the levels of $\alpha1(I)$ mRNA in MG-63 osteosarcoma cells (70). In contrast, estradiol had no effect on type I collagen synthesis in MG-63 cells but decreased synthesis in the less differentiated SaOs-2 osteoblast-like osteosarcoma cells (70). One of the primary effects of estradiol on osteoblasts is thought to be in the regulation of the circuitry of cytokine actions controlling bone remodelling (70). Although oestrogen enhanced type I procollagen mRNA levels in rat osteosarcoma cells it has been shown to cause a decrease in mRNA levels of type I procollagen, osteocalcin and osteopontin in bone cells derived from ovariectomised rats (71,72). In a study by Silbiger *et al* (73) estradiol was shown to suppress type I collagen synthesis in murine mesangial cells, however, neither the $\alpha1(I)$ nor the $\alpha2(I)$ procollagen genes contain an oestrogen response element. Estradiol, however, modulates the transcription of several genes that lack an oestrogen response element but contain a regulatory AP-1 binding motif. It has been shown to increase the steady state levels of *c-fos* mRNA and this was associated with increased binding of nuclear extracts to an AP-1 consensus binding site oligonucleotide (73). To confirm the involvement of AP-1, cells were treated with either PMA (activator of AP-1) or curcumin (inhibitor of AP-1) and it was shown that PMA decreased type I collagen synthesis whilst curcumin resulted in an increase and the conclusion was that the suppression of type I collagen synthesis by estradiol was mediated via enhanced AP-1 activity (73).

1.2.2.3 The role of vitamins in the synthesis of type I collagen

Vitamin C is required in the synthesis of type I collagen, however, it has variable effects on the expression of the type I procollagen gene (6). Increased type I collagen synthesis occurs in cultured fibroblasts and primary avian tendon cells exposed to ascorbic acid. Retinoic acid, on the other hand, has a general inhibitory effect on type I collagen levels; its effects have been documented in human skin fibroblasts and in MG-63 osteosarcoma cells (6,70). The decrease in either $\alpha1(I)$ or $\alpha2(I)$ procollagen gene expression that occurs after treatment with retinoic acid is not associated with a decrease in message stability, but due to a reduced rate of transcription. Vitamin D (1,25-dihydroxyvitamin D_3) is important in bone deposition but its effects have been shown to be variable, depending on the clonal origin, methods of culture and the extent of differentiation of the cell line (6). Treatment of primary human osteoblast cultures or rodent osteoblast-like cells with Vitamin D results in increased type I collagen, whilst a decrease was observed in ROS 17/2.8 rat osteosarcoma cells (74). Interestingly, Vitamin D has been shown to reverse the inhibition of type I collagen synthesis in MC3T3-E1 osteoblastic cells overexpressing *c-fos* (75). Calcitriol, a Vitamin D metabolite, stimulated the synthesis and secretion of type I collagen in MG-63 osteoblast-like osteosarcoma cells, by increasing the $\alpha 1$(I) procollagen mRNA levels; this effect appears to involve the induction of a regulatory factor since treatment with calcitriol induced a rapid but transient increase in the expression of proto-oncogenes *c-fos* and *c-jun* (70).

1.2.3 The role methylation in the expression of type I procollagen gene

The methylation of cytosines in the dinucleotide CpG is essential for mouse development and it has been implicated in differential regulation of gene expression, suggesting a possible link between methylation and suppression of transcription of a number of tissue specific genes, however, the precise mechanism is not fully understood (76-79). The vertebrate globin genes were among the first examples in which an inverse correlation was shown between CpG methylation and transcriptional regulation (80). Methylation of the minimal promoter, 235 bp, of an embryonic ρ-globin gene silenced transcription in primary chicken erythroid cells (80). There was a 20-30 fold inhibition of transcription and this effect was not overridden by the presence of potent erythroid specific enhancers. A fully methylated but not unmethylated 235 bp fragment of the ρ-globin gene promoter fragment could compete efficiently for the DNA methyl cytosine binding protein complex (MeCPC) binding (80). The methylation of cytosine residues at the CpG sites has been shown to provide a binding site for MeCP2, an abundant mammalian protein capable of transcriptional repression by binding to DNA in a non-sequence specific manner (81). In a study by Umezawa *et al* (82) the optimal activity of the keratin 18 (K18) gene was shown to be dependent upon an ETS binding site within an enhancer region located in the first intron; methylation of the ETS site was correlated with the repression of the K18 gene in normal human tissues and in K18 transgenic mouse tissues (82). The common feature of these studies is the significance of methylation in the regulation of gene expression.

A number of reports have examined the role of DNA methylation in the expression of collagen genes. Analysis of the methylation pattern of the $\alpha2(I)$ procollagen gene showed no differences in the differentiated (expressing type II collagen) and dedifferentiated (expressing type I) chick embryo chondrocytes (83). Thus, it appears that the DNA methylation pattern during cell differentiation is not sufficient to explain activation and inactivation of the procollagen genes. Transient transfection of *in vitro* methylated promoter and enhancer reporter constructs of the $\alpha1(I)$ procollagen gene has been shown to lead to transcriptional inactivation of the gene (84). A similar observation was made in human rhabdomyosarcoma cells that do not produce $\alpha1(I)$ collagen whereby inhibition of DNA methylation by treatment with 5-azacytidine resulted in transcriptional activation of the gene, providing additional evidence that DNA methylation is an important mechanism of transcriptional inactivation of the interstitial procollagen genes (84). In another study, DNA methylation was reported to play a role in transcriptional regulation of collagen IV genes during differentiation of F9 teratocarcinoma cells (85). F9 cells express low levels of mRNA for type IV collagen and treatment of these cells with 5-azacytidine activated transcription of an integrated $\alpha1(IV)$ procollagen gene promoter-enhancer-CAT construct and it was further shown that 5-azacytidine acted in synergy with retinoic acid and cAMP. Analysis of DNA isolated from F9 cells revealed that there was a specific demethylation of the DNA within the 5′-flanking region of the collagen IV genes following treatment with retinoic acid and cAMP (85).

DNA methylation of the murine $\alpha1(I)$ collagen promoter and part of the first exon linked to a luciferase gene construct resulted in repression of transcription in fibroblasts and to a lesser extent in embryonal carcinoma cells (86). Interestingly, Sp1 and NF-1, factors that both bind to the $\alpha1(I)$ collagen promoter, were shown to transactivate the unmethylated and methylated luciferase reporter gene in Drosophila SL2 cells, confirming that the factors can bind to methylated DNA indicating that DNA methylation represses the $\alpha1(I)$ collagen promoter via a mechanism that is independent of Sp1 or NF-1 binding. Kopp *et al* (87) observed that expression of type VI collagen is completely inhibited in virally transformed fibroblasts and cells derived from spontaneous mesenchymal tumours; the decreased type VI collagen mRNA levels were accompanied by a significant increase in the mRNA levels of the DNA methyl transferase. Furthermore, *in vitro* methylation of the $\alpha2(VI)$ collagen promoter abolished promoter activity in transient transfection assays, and sequence analysis of the genome showed that the promoter was extensively methylated in transformed fibroblasts and virtually no methylation was observed in normal cells (87). Thus implying that DNA methylation is involved in the silencing of the $\alpha2(VI)$ collagen gene and it is possible similar mechanisms are responsible for the repression of other transformation sensitive proteins (87).

The $\alpha2(I)$ procollagen gene is repressed in W8 cell line as a result of DNA methylation (28,29,88). These cells contain a transcriptionally inactive $\alpha2(I)$ gene with extensive hypermethylation in the promoter and first exon (88). Methylation of specific sites in the $\alpha2(I)$ procollagen gene were responsible for inactivation of transcription in the W8 cells (88). In a recent study, methylation in the first exon was shown to suppress $\alpha2(I)$ procollagen gene transcription (60). Specific methylation of the CpG sites at +7 and +23 in the first exon is accompanied by increased protein-DNA complex formation and inhibition of the transcriptional activity as confirmed by transient transfection and *in vitro* transcription (60). Based on these observations, it would appear that methylation plays a role in the regulation of expression of a number of genes including type I procollagen genes. Further studies, however, are required to correlate methylation of specific *cis*-elements with the activity of transcription factors that have been shown to regulate the expression of type I procollagen genes.

1.3 Signal transduction from the cell surface to the nucleus

The regulation of gene expression is often associated with events that take place inside the nucleus, but the extracellular environment plays a crucial role in the expression of almost all eukaryotic genes. A number of physiological and non-physiological agents that modulate expression of type I procollagen genes are present in the extracellular milieu of cells to allow cells to perform their required functions. The important question is how do signals arising from the extracellular environment ultimately transcend into discrete changes in gene expression in normal and transformed cells. Previous studies have demonstrated that transformation of normal human embryonic lung fibroblasts (WI-38 fibroblasts) by γ-irradiation (CT-1 fibroblasts) does not alter the ability of the transformed phenotype to express $\alpha2(I)$ procollagen, however, these cells have lost the ability to trigger the expression of a very central gene in the signal transduction cascade, namely *c-fos* (32,33,89). Induction of these transformed cells with either serum or phorbol esters failed to induce *c-fos* at both the RNA and protein levels (89). This effect could not be correlated with a specific alteration in the structure of the promoter of the *c-fos* gene and no obvious deletions or insertions were observed in the gene (89). A study by Mizuki *et al* (90), however, demonstrated that the *c-Ki-ras* 2 gene is amplified in CT-1 cells and is accompanied by an elevated level of *kis* specific RNA. In addition, the report showed that these cells expressed 10 times more *c-myc* mRNA compared to normal cells which is associated with their immortal phenotype (90). No signals were obtained for the *c-yes* gene, codes for a non-receptor protein tyrosine kinase, and for the *c-mos* gene, which codes for a serine/threonine protein kinase (90). Thus it is apparent γ-irradiation has diverse effects on the expression of a number of oncogenes, that is, some are shut down, like *c-fos*, whilst others, such as *c-myc*, are overexpressed. Interestingly, transformation of WI-38 fibroblasts by SV40 (SVWI-38 fibroblasts) resulted in a shutdown of the $\alpha2(I)$ procollagen gene and cells retained their ability to trigger *c-fos* in response to appropriate stimuli (13). These observations suggest that changes in the expression of a specific gene is determined to a certain degree by intricate events originating from the extracellular environment and since type I collagen is an ECM protein, it has the potential to modulate its own expression in both normal and transformed cells. A schematic representation of some of the signal transduction pathways that have been reported to modulate the levels of type I procollagen genes is illustrated in figure 1.3.

The accumulation and organisation of ECM components play a critical role in development, maintenance and pathogenesis of most organ systems (1). These processes are regulated by the precisely orchestrated expression of ECM components, their receptors and matrix proteases. Growth factors such as TGFβ, basic fibroblast growth factor (bFGF), insulin-like growth factor (IGF) and angiotensin II have been shown to modulate elastin and collagen production in porcine vascular smooth muscle cells and skin fibroblasts as well as stimulate collagen gel contraction in an *in vitro* model for wound healing (91,92). The stimulatory effect of angiotensin II could be blocked by *src*-related tyrosine kinase inhibitors, such as genistein and herbimycin (92). Thus, it would appear that tyrosine phosphorylation is crucial for the response of cells to angiotensin II and the JAK2/STAT pathway has been implicated in this process (92). The accumulation of collagen in burn hypertrophic scars, on the other hand, appears to be TGFβ dependent (93). The stimulation of collagen production was inhibited by nitric oxide (NO) which acts via the adenylate cyclase pathway involving cGMP (93). In hepatic stellate cells stimulation of G proteins, adenylate cyclase or addition of 8-bromo-cAMP, with subsequent activation of protein kinase A (PKA), resulted in a decrease in the $\alpha 1(I)$ procollagen mRNA levels (94). Blocking G_i protein (the G protein that inhibits adenylate cyclase activity with a resultant decrease in the levels of cAMP), phospholipase A_2 (signalling involves prostaglandins) or phospholipase C (involves the phosphoinositide pathway) as well as blocking the accumulation of intracellular calcium and calmodulin levels, increased the expression of the $\alpha 1(I)$ procollagen transcript in hepatic stellate cells (94).

Fibroblasts isolated from systemic sclerosis (SSc) lesions have been shown to synthesise increased amounts of type I and III collagens and the increase was shown to be a result of increased transcriptional activity (95). These fibroblasts have a proliferative response to TGFβ, platelet derived growth factor (PDGF) and bFGF, however, SSc fibroblasts are less responsive to bFGF compared to normal fibroblasts. Basal fibroblast growth factor inhibited basal and TGFβ induced α2(I) procollagen gene expression in scleroderma and normal fibroblasts and at this stage the mechanism remains unclear (95).

Signals that arise from TNFα and IL-1 are transmitted via the activation of protein kinase C (PKC) or PKA and a subsequent increase in the levels of *c-fos* and *c-jun*; in addition, the pathway leads to the activation of NF-κB, a member of the *rel* family of proteins (96,97). Activation is achieved via the proteolytic cleavage of the I-κB subunit from the p65/p50 protein complex (98). The active NF-κB (p65/p50) is translocated to the nucleus where it participates in a number of DNA protein interactions in genes regulated at NF-κB sites (99). Although NF-κB is activated by a number of effectors that activate PKC, such as TNFα, IL-1 and PMA, which also induce the expression of *c-fos*, there are no NF-κB sites in the *c-fos* promoter, NF-κB/p65 has been reported to cross-couple with Fos or Jun and in that way activate κB and AP-1 enhancer-dependent promoters (100). In a recent report, Ogretmen and Safa (101) demonstrated that negative regulation of the multidrug resistance (MDR) gene promoter activity is inhibited in MCF-7 human breast cancer cell line via cross-coupling between NF-κB/p65 and Fos and this complex was shown to interact with a CAAT sequence in the MDR promoter. Both TNFα and IL-1 have been reported to have a negative effect on the expression of type I procollagen genes in a number of studies (6). NF-κB was shown to mediate TNFα inhibitory effect on α2(I) procollagen gene transcription in human dermal fibroblasts (43). Although the mechanism requires further elucidation, it is possible that TNFα activates the IκB kinase cascade that leads to translocation of NF-κB from the cytoplasm to the nucleus to modulate gene expression.

Studies using the tumour promoter phorbol-12-myristate-13-acetate (PMA; also called 12-o-tetradecanoylphorbol-13-acetate, TPA), a potent activator of protein kinase C, have generated contradicting data (102). Treatment of JB-6 mouse epidermal cells and NIH3T3 mouse fibroblasts with PMA has been shown to inhibit the expression of the type I procollagen gene (103,104) while Stuiver *et al* (105) reported that PMA increased the steady state mRNA levels of the $\alpha2$(I) procollagen gene by 6 fold in mouse 3T3-L1 cells. This effect was also observed at the protein level. Interestingly, PMA treatment of mouse NIH3T3 cells resulted in an increase in *c-myc* and *c-fos* mRNA levels (104).

Ultraviolet radiation (UV) from the sun is known to damage the human skin resulting in an old and wrinkled appearance through a process called photoaging (106). Photoaging results partly from alterations in the composition, organisation and structure of the collagenous ECM in the dermis (106). UV irradiation triggers the activation of a number of cytokine and growth factor receptors on keratinocytes and dermal cells and the downstream events in the signal transduction pathway involve activation of the MAP kinase pathway, specifically the JNK/SAPK cascade and an ultimate increase in the expression of *c-jun* (106). Jun in turn forms heterodimers with Fos to form the AP-1 transcription factor complex. The AP-1 complex binds to AP-1 consensus sequences present in promoters and introns of target genes of a number of matrix metalloproteinases (MMP) such as interstitial collagenase (MMP-1), 92 kDa gelatinase B (MMP-9) and stromelysin genes (MMP-3, MMP-10 and MMP-11), with a concomitant increase in their expression and degradation of collagen and other ECM proteins in the dermis (106-108). Although the degraded matrix may be repaired, the process generates an "imperfect" solar scar, which upon subsequent exposure to UV light will result in accumulation of solar scarring that is a manifestation of photoaging. This process has been shown to be inhibited by all-trans-retinoic acid via the inhibition of *c-jun* induction (106).

In hepatic stellate cells, the primary cell responsible for the dramatic increase in the synthesis of type I collagen in cirrhotic liver, inhibition of Ras and Raf activity results in an increase in the expression of the type I procollagen gene, whilst, inhibition of the MAP kinase activity was accompanied by a decrease in the type I procollagen mRNA (109,110). These results suggest a possible branch point between the Ras→Raf and MAPK pathways. Overexpression of oncogenic *ras* in Rat-1 fibroblasts has been shown to down-regulate type I procollagen, the mechanism of action appeared to involve both enhanced turnover of the message and decreased transcriptional activation mediated via an intronic AP-1 site in the $\alpha 1$(I) procollagen gene (21). Ledwith *et al* (111) demonstrated that antisense-*fos* RNA caused a partial reversion of the transformed phenotype induced by the c-Ha-*ras* oncogene. These results support the idea that *c-fos* is a downstream mediator of *ras* in the growth regulatory signal transduction pathway. Transformation of fibroblasts with Ras is characterised by reduced adhesion to the substratum due to reduced expression or increased degradation of ECM proteins including type I collagen (112,113). Transfection of *ras* transformed fibroblasts with the $\alpha 2$(I) procollagen expression vector construct increased their adherence to the substratum and suppressed tumourigenicity of these cells in nude mice (114). These observations suggest that ECM proteins could subvert oncoprotein signalling pathways associated with the malignant phenotype. In other reports, the treatment of cultured fibroblasts with okadaic acid (OA), the phosphoserine/threonine specific phosphatase 1 and 2A (PP1/2A) inhibitor, was shown to suppress type I procollagen gene expression (115). Protein phosphatases form an important part of the signal transduction pathway in that activation or inactivation of a specific pathway might require removal of phosphate moieties, some of the proteins involved in the signal cascade are activated by dephosphorylation (116). The selective inhibition of PP1/2A offers an interesting approach for treating disorders related to overexpression of type I procollagen.

It is evident that several signal transduction pathways are involved in the regulation of expression of the type I procollagen genes in normal and pathological conditions, perhaps there is a lot of cross-talk in these cascades. Interestingly most signals appear to result in an increase in the expression of *c-fos* and *c-jun* early immediate genes (117,118).

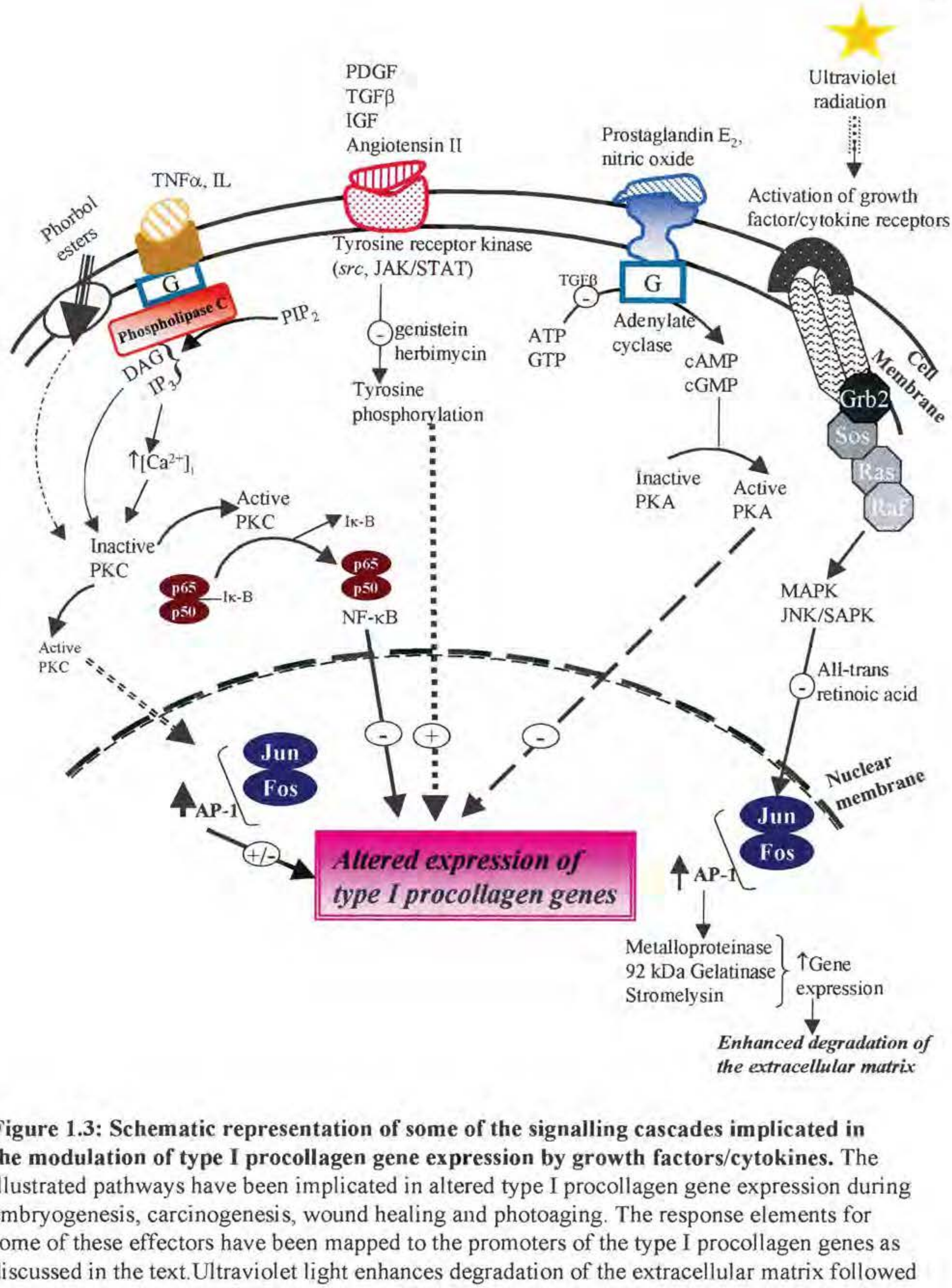


Figure 1.3: Schematic representation of some of the signalling cascades implicated in the modulation of type I procollagen gene expression by growth factors/cytokines. The illustrated pathways have been implicated in altered type I procollagen gene expression during embryogenesis, carcinogenesis, wound healing and photoaging. The response elements for some of these effectors have been mapped to the promoters of the type I procollagen genes as discussed in the text.Ultraviolet light enhances degradation of the extracellular matrix followed by the process of "imperfect"repair which yields inadequacies in the structural integrity of the dermis creating a solar scar. AP-1 proteins together with other early response genes convert the extracellular stimuli into changes in the genetic programme of a number of genes, including type I procollagen genes. It also possible for Fos or Jun to cross-couple with NF-κB/p65 protein to modulate gene expression at sites other than the AP-1 or NF-κB binding sites.

1.4 The proto-oncogene *c-fos*

The *c-fos* proto-oncogene is a cellular homologue of the transforming gene, *v-fos*, of the murine Finkel, Biskis and Jinkins (FBJ) and Finkel, Biskis and Reilly (FBR) retroviruses that cause osteogenic sarcomas (119,120). Both *c-fos* and *v-fos* genes are capable of transforming fibroblasts in culture, although certain changes in the *c-fos* gene are required for a successful transformation (121-125). Changes in the regulatory upstream elements such as insertion of a long terminal repeat to drive expression of the gene or removal of the RNA destabilising sequence in the 3′ untranslated region (UTR) suffices to activate the transforming potential of *c-fos* (122). Transcripts of 2.2 kb which translate into a protein of 380 amino acids with an apparent molecular weight of 42 kDa have been identified in human cells (126,127). The protein, however, migrates as a broad band between 54 and 62 kDa in sodium dodecyl sulphate-polyacrylamide gels due to the high proline content and extensive phosphorylation of the protein. The *c-fos* mRNA and protein are undetectable in quiescent cells and are elevated in response to stimulation by hormones, serum, mitogens and phorbol esters (117,127). The mRNA has a half-life of 15-30 minutes whilst the protein has a half-life of about 2 hours; the half life of the protein depends on the extent of post-translational modification of the protein (121,124,128,129). The rapid turnover of the message is associated with a 100 nucleotide long AU-rich region located at the end of the 3′ UTR (124,130,131). In addition to Fos, there are three other proteins which belong to the Fos family of proteins, namely FosB (46 kDa), Fra-1 (35 kDa) and Fra-2 (30 kDa) (117). Immunoprecipitation of the *v-fos* and *c-fos* proteins from fibroblasts generates a protein complex, which contains the Fos protein in tight but non-covalent association with a basic cellular phosphoprotein of between 39 and 40 kDa (132-134). These Fos associated proteins were later identified as proteins belonging to the Jun family (135,136).

1.4.1 Regulation of *c-fos* expression

The expression of the *c-fos* gene takes place during a number of biological events such as cell proliferation, differentiation and transformation as well as during apoptosis (117,122,124,127,137,143). The *c-fos* proximal promoter is contained within a 350 bp region upstream of the transcription start site, a number of *cis*-acting regulatory elements have been mapped to this region of the promoter (117). The *c-fos* promoter is a TATA box-containing promoter with several upstream elements, including a calcium-cAMP response element (CRE), a serum response element (SRE), a binding site for AP-1 or TPA-response element (TRE) directly adjacent to the SRE, and a *c-sis*-PDGF inducible factor element (SIE) (144-147). The CREB transcription factor has been associated with calcium and cAMP induction of *c-fos* promoter through the CRE, and SIE element binds STAT proteins and can regulate the responses to PDGF and tyrosine kinases (148,149). The SRE binds SRF and can form complexes with a number of other transcription factors, most notably ternary complex factors (TCFs) including Elk-1 and SAP-1, which are responsive to MAP kinase and related pathways (98,117,118). Fos can, on the other hand, negatively autoregulate its expression and this trans-repression occurs at either the TRE or other elements (150,151). In a recent review, Bateman indicated that the negative autoregulation of Fos requires the SRE, although the mechanism is not entirely clear, Bateman suggests Fos and Jun proteins may both be involved in repression which does not seem to involve DNA binding by AP-1 (152). Studies by Lucibello *et al* (153) have demonstrated that overexpression of *c-fos* results in transrepression of the *c-fos* promoter via the TRE. Trans-repression was shown to be enhanced by co-expression of *c-jun* and it did not require AP-1 or ATF (CRE) sites in the mouse *c-fos* promoter (153). These studies also demonstrated that the SRE is a major target of repression. Structurally, the Fos protein required the bZIP motif, important in dimerisation, for trans-repression and transformation (117,127,153). In addition to regulation via promoter *cis*-elements, the localisation of the Fos protein plays a role in the function as a transcription factor (154). The protein is synthesised in the cytoplasm and it is subsequently translocated to the nucleus. This translocation process, however, is not spontaneous but depends on the continuous stimulation of cells by serum factors while serum starvation results in the retention of the already synthesised *c-fos* protein in the cytoplasm until it is degraded (154). Whilst this is true for the *c-fos* proto-oncogene product, the *v-fos* protein has been shown to evade the translocation control and by maintaining a high nuclear presence, it can function potently in the process of tumourigenesis (154).

1.4.2 Regulatory domains of Fos

The ability of Fos to engage in the transcriptional regulation of other genes lies in the observation that the protein participates in heterodimer complex formation with members of the Jun family of proteins and subsequently engage in sequence specific DNA binding to AP-1 consensus sequences in the regulatory regions of target genes (117,155-157). The protein contains three main domains, namely, a) an N-terminal domain involved in transactivation, transrepression and a basic region which harbours the DNA binding residues, b) the central domain contains a leucine zipper structure that is crucial for heterodimerisation with the Jun family of proteins to constitute the AP-1 transcription factor, and c) adjacent to the leucine zipper on the C-terminus is another domain involved in transactivation and transrepression (116,117,158). The N-terminus is regulated by phosphorylation involving the cdc2 kinase and protein kinase C, whilst the C-terminus has been shown to contain residues phosphorylated by cdc2 kinase, protein kinase A and double-stranded DNA protein kinase (116,158,159). The C-terminus has been reported to be crucial for the transformation ability of Fos; mutation of residues 362-364 gives rise to a Fos protein which is activated for cellular transformation, supporting the idea that loss of this region in the virally-transduced *v-fos* oncogene contributes to its oncogenicity (123-125). Fos is therefore a very good example of a nuclear protein with different activities, which is regulated independently by phosphorylation.

In order for Fos to function as a transcription factor, it has to engage in dimerisation with the Jun family of proteins involving the leucine zipper domain (117,158,160). The dimerisation region, however, extends beyond the leucine zipper to include a histidine residue located 7 amino acids from the last leucine on the C-terminus (160,161). Involvement of additional regions of the protein in heterodimerisation ensures formation of different homodimeric and heterodimeric leucine zipper complexes which have subtle differences in their DNA binding properties; the overall transcriptional activity of the dimer results from a combination of the effects of the positive and negative domains, thus giving each protein dimer a unique transcriptional profile (158,160-162). Co-transfection of *c-jun* and *c-fos* expression vectors into F9 embryonal carcinoma cells results in more potent transactivation of AP-1 dependent indicator genes than what is achieved using a *c*-Jun expression vector alone; this increase has been ascribed to the formation of *c*-Jun:*c*-Fos heterodimers which are more efficient DNA binding protein complexes than *c*-Jun homodimers (155,163). *c*-Fos, on the other hand, can not form homodimers, and since heterodimerisation is a prerequisite for DNA-binding, *c*-Fos does not on its own bind DNA hence it does not stimulate transcription in the absence of *c*-Jun (163). These associations have also been observed among other members of these families, which include JunB, JunD, FosB, Fra-1 and Fra-2 (160,164-168). All Jun proteins so far can form both homo- and heterodimers, for example JunB:*c*-Jun, and therefore they can bind to the AP-1 sites in the absence of Fos proteins (160). The Fos proteins, however, can associate with any of the Jun proteins to generate stable heterodimers that have a higher DNA-binding activity than the Jun homodimers (169). Although all these complexes seem to have very similar recognition properties and interact with the consensus AP-1 and AP-1-like sequences, they exhibit markedly different DNA binding activities (169). The differences in the protein structure might, however, contribute to differences in gene regulation; their presumed functional differences may be based on physiological activities other than DNA binding and may reside in other domains of the protein (160,169). Measurement of dimer dissociation indicated that the increased DNA-binding activity of the Jun:Fos heterodimer is due to an increased thermostability in comparison to the Jun:Jun homodimer; the heterodimer was found to dissociate between 37 °C and 42 °C whilst the homodimer dissociated between 25 °C and 37 °C, and Fos homodimers failed to form even at 4 °C (117).

1.4.3 Biological effects of Fos

Fos has a high affinity for double stranded DNA, although it only associates with DNA in the presence of Jun (170,171). In an attempt to determine whether the protein is involved in transactivation of other genes, *v-fos* constructs were cotransfected with a type III procollagen promoter linked to a reporter gene, and it was found that *v-fos* is capable of transactivating the type III procollagen promoter, however, no direct evidence of transactivation was obtained to suggest that a specific *cis*-regulatory element is involved (172). Interestingly, constitutive expression of *c-fos* has been shown to inhibit the synthesis of type I collagen in osteoblasts, however, the study did not investigate the mechanism, but the inhibition was attributed to a transcriptional event (173). A number of studies have demonstrated that overexpression of *c-fos* in transgenic or chimeric mice affects cartilage, bone and haemopoietic cell development (142,174-176). Transgenic mice over-expressing *c-fos* under the control of the human metallothionine promoter developed specific lesions in the long bones at 2-3 weeks after birth and only 15 % of the mice developed osteosarcomas after a 9-10 month latency period (142). Thus it would appear that the expression of *c-fos* is important in the formation of bone.

Cells expressing the oncogenic Ras protein exhibit a variety of changes in gene expression, including activation of the AP-1 transcription factor (21,177-179). Activation of the AP-1 transcription factor has been correlated with the synthesis of high levels of matrix degrading enzymes (106-108). The expression of Fos has been shown to be a prerequisite for the transcriptional activation of the collagenase gene by other oncogenes and phorbol esters (180). Oncogenic *ras* transformation accounts for the significant reduction in type I procollagen expression in a process that also involves increased activity of the AP-1 transcription factor (64). To further support these observations, transfection of antisense *c-fos* RNA was able to reverse transformation of cells by *ras*, thus implying that Fos is the downstream mediator of Ras effects in the nucleus (111).

Chung *et al* (42) demonstrated that treatment of human foreskin fibroblasts with TGFβ upregulated the expression of the α2(I) procollagen gene whilst TNFα had a negative effect on the transcription of the gene. These effects could be mapped to the region of the α2(I) procollagen promoter that contained an AP-1 like consensus sequence. To further support their findings, the authors showed that overexpression of *c-jun* in cells co-transfected with the α2(I) procollagen promoter reporter construct containing a putative AP-1 site blocked the TGFβ response. These results further implicated AP-1 in the regulation of the α2(I) procollagen gene expression in fibroblasts in the presence of TGFβ and TNFα. Recently, Kouba *et al* (43) demonstrated that TNFα inhibits α2(I) procollagen expression in human dermal fibroblasts through NF-κB binding sites present in the proximal promoter of the gene. TNFα is known to be a potent inducer of *c-jun* expression whilst TGFβ induces *JunB*, the two Jun proteins have opposite effects on gene expression (181). It is therefore possible that an increased expression of *c-jun* or enhanced formation of Fos dimers that involve c-Jun rather than Fos-JunB dimers, with a consequent transrepression of the α2(I) procollagen gene via AP-1 sites could account for the TNFα effect.

The abundance and wide distribution of type I collagen is a clear indication of the significance of this protein in the maintenance of structure in higher vertebrates. A number of studies have provided substantial information on the regulation of expression of type I procollagen genes by delineating regulatory elements and in certain instances factors that complex with these elements either *in vitro* or *in vivo*. A number of models have been used to address this question, such as cell lines derived from scleroderma patients and fibrotic tissue, transformed cells which display altered expression of type I procollagen genes and transgenic animals; although the approaches and model systems might be distinct, the primary objective of these studies has been elucidation of mechanisms that lead to changes in the expression of these genes and ultimately designing drugs that can be used to alleviate symptoms associated with collagen diseases. The data presented in this study will thus further advance the current knowledge and understanding of regulation of the human type I procollagen gene expression.

CHAPTER 2

CHARACTERISATION OF PROXIMAL PROMOTER ELEMENTS IN THE HUMAN α2(I) PROCOLLAGEN GENE.

2.1 INTRODUCTION

Changes in the synthesis of type I collagen occur as a normal physiological event during embryonic development and in certain pathological conditions, such as during wound healing and fibrosis of the liver and lung (8). A number of other events where modulation of synthesis of type I collagen occurs include: (a) cell transformation by viral oncogene products, chemical carcinogens or ultraviolet light, (b) cytokines, (c) hormones and (d) vitamins (6,182). Most of these changes in the expression of this gene have been shown to occur at the transcriptional level. (See Chapter 1 for details).

Expression of eukaryotic genes is regulated by multiple transcription factors that act either as activators or repressors/silencers. Many of these factors bind to specific DNA sequences in the regulatory elements of genes and promote initiation of transcription by RNA polymerase II through interactions with other components of the transcription machinery. To understand procollagen gene transcription under different physiological and pathological conditions, a number of studies have focused on the identification and characterisation of *cis*-acting regulatory elements in mammalian procollagen genes. The 3500 bp region located immediately upstream of the transcription start site of the human α2(I) procollagen gene contains most of the sequences necessary for tissue specific transcription (35). *Cis*-acting DNA elements that direct high and tissue specific transcription of the human COL1A2 gene, however, are located between nucleotides -376 and -108, a region referred to as the proximal promoter. Some of the *cis*-acting elements mapped to the human proximal promoter region include an 11 bp element located between -67 and -78 that binds a novel factor and is a focus of this study (13,51), an inverted CCAAT box located between -80 and -84 (47,51), a novel positive *cis*-element containing a TCCTCC motif between -119 and -133 (47), a pyrimidine-rich repressor binding element between -159 and -164 (47), and a CAGA sequence between -247 and -250 which binds an as yet unidentified factor (1). Additional elements have been mapped on the mouse α2(I) procollagen promoter including the inhibitory factor-1 (IF-1) binding site between -145 and -154 and an NF-1 binding element between -295 and -315 (46,182). The novel element between -67 and -78 is, however, absent in the mouse promoter. Studies by Inagaki *et al* (38) and Greenwel *et al* (41) mapped a TGFβ-responsive element (TbRE) between -378 and -255, this region contains Sp1 binding sites at -313 and -286. In a conflicting report, however, Chung *et al* (42) narrowed the TbRE between -265 and -241, this region contains an AP-1-like binding site and the authors excluded binding of either Sp1, NF-1 or NF-κB transcription factors to the TbRE. In a recent report, a TNFα and IFN-γ responsive elements were mapped to the proximal promoter of the human α2(I) procollagen gene (43,49)

In the present study, normal and transformed WI-38 human embryonic lung fibroblasts were used as models to study the transcriptional regulation of the human α2(I) procollagen gene. Transformation of fibroblasts is generally accompanied by changes in the expression of the genes encoding type I procollagen. Some of the documented changes in the expression of type I procollagen that occur as a result of transformation include shutdown in the expression of only the α2(I) procollagen gene after transformation of WI-38 fibroblasts by SV40, SVWI-38 fibroblasts (183), transformation of Rat-1 fibroblasts by mutant *ras* down-regulates the expression of both α(I) and α2(I) procollagen genes (21), and transformation of Rat-1 fibroblasts by *v-fos* results in the decreased expression of both type I procollagen genes (22). Surprisingly, transformation of WI-38 fibroblasts by γ-irradiation (CT-1) has been shown to have very little effect on the expression of the type I procollagen genes; these cells have been shown to express more than 80 % of total type I collagen (32,33). Thus, the two transformed WI-38 fibroblast lines, viz.: CT-1 and SVWI-38, are essential tools for the characterisation of transformation related changes in the expression of type I procollagen and these differences were exploited in this study. Differences in the expression of the α2(I) procollagen gene between CT-1 and SVWI-38 fibroblasts have been shown to occur at the transcriptional level (13,33,57,59). Subsequent studies concentrated on the identification of *cis*-elements that are important in the regulation of the α2(I) procollagen gene in the two cell lines. Transient transfection studies using α2(I) procollagen promoter-chloroamphenicol acetyl transferase constructs (COL1A2-CAT) delineated differences in the promoter activity to the proximal promoter region (13). Although the basal promoter fragment (-107 to +1) generally shows very low promoter activity, there were also differences in the activity of this promoter fragment in the two cell lines, implying that the regulatory elements for the differential expression of the gene could reside in the proximal promoter. Further analysis of the -107 promoter identified two *cis*-acting elements, viz.: an inverted GGAGG/CCAAT binding element (G/CBE) and a procollagen modulating element (CME), between -80 and -67 (13,51). Both these elements were shown to form distinct DNA-protein complexes *in vitro* in the presence of nuclear proteins, produced a DNase I footprint and DNA methylation interference analysis confirmed the guanines involved in binding. The possible function of proteins that complex with these elements was extrapolated from studies based on DNA binding assays *in vitro*.

The main objective of this study was to demonstrate that G/CBE and CME binding proteins play a regulatory role in the expression of the α2(I) procollagen gene and to establish whether a similar mode of regulation applies to fibroblasts that differ in their ability to express the α2(I) procollagen gene. The study also investigated the effect of proteases on DNA-protein complexes formed on the -107 to -60 α2(I) procollagen promoter. To realise the main objective of this chapter, point mutations were introduced into the α2(I) procollagen promoter and the effect of these mutations on the transcriptional activity of the α2(I) procollagen promoter was analysed by transient transfection assays. The use of mutants has become an essential tool in the study of gene structure and function since mutations in *cis*-acting elements interfere with transcriptional activity.

2.2 RESULTS

2.2.1 Characterisation of DNA-protein complexes formed on the proximal α2(I) procollagen promoter

Previous studies reported differences in transactivating factor binding to the α2(I) procollagen gene in CT-1 and SVWI-38 nuclear extracts (13,51,57). Three specific DNA-protein complexes were shown to form on the -107 to -60 α2(I) procollagen promoter as illustrated in figure 2.1. Complex I proteins were shown to bind to the inverted CCAAT box and the factor that binds to this region was concluded to belong to the CBF family of proteins (51,59). Complex III, on the other hand, formed in the presence of nuclear proteins from α2(I) procollagen expressing cell line, CT-1, whereas complex II proteins were present in the non-expressing cell line, SVWI-38 (59). These studies concluded that the complex II protein is associated with inactivation of the α2(I) procollagen gene whilst complex III protein is associated with activation of the gene (59).

This study shows that complex II proteins are ubiquitous and that complex III protein is a specific cleavage product of complex II proteins. To demonstrate this, nuclear proteins were extracted in the presence and absence of protease inhibitors, incubated with the -107 to -60 promoter fragment and separated by electrophoresis on a non-denaturing 5 % polyacrylamide gel (see section 6.6.2). Complex III proteins appeared in all extracts lacking protease inhibitors, with the degradation products being much more pronounced in CT-1 fibroblasts (Figure 2.2). Nuclear proteins were then extracted in the presence of various combinations of protease inhibitors, and from data presented in figures 2.3 and 2.4, it is evident that the omission of leupeptin in the extraction buffers resulted in the appearance of complex III proteins in nuclear extracts from both CT-1 and SVWI-38 fibroblasts; whilst the omission of either pepstatin A or phenylmethylsulfonyl fluoride (PMSF) did not have the same effect. The ratio of complex II to complex III was higher for SVWI-38 extracts as compared to extracts from CT-1 fibroblasts implying that complex II proteins are more stable in SVWI-38 fibroblasts, alternatively, the protein could be protected from degradation

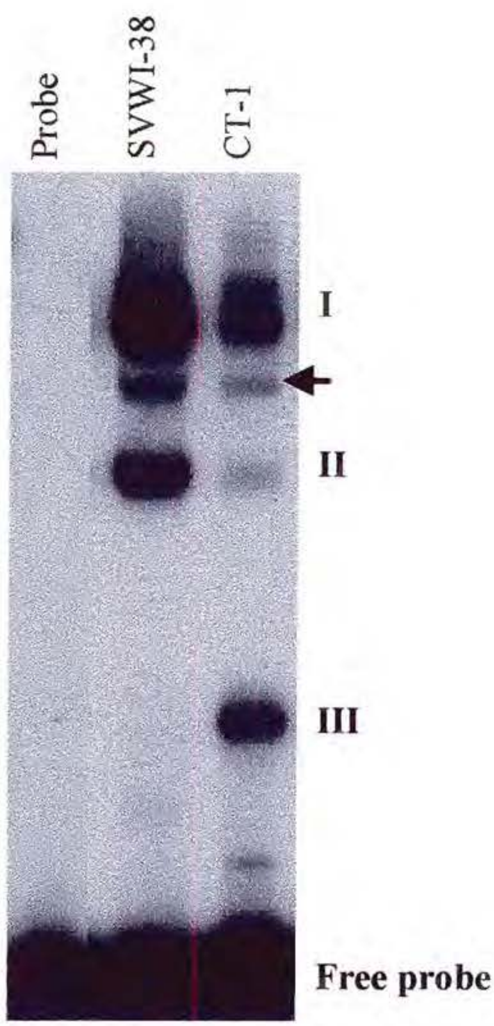


Figure 2.1: Electrophoretic mobility shift assay of nuclear proteins. Nuclear proteins were extracted from SVWI-38 and CT-1 fibroblasts using the modified Dignam method described by Parker *et al* (13), incubated with a ^{32}P-labelled fragment of the $\alpha 2(I)$ procollagen promoter between -107 and -60 and DNA-protein complexes separated by electrophoresis on a non-denaturing 5% polyacrylamide gel. The gel was dried and exposed to X-ray film at -70 °C for 16 hours. **I, II, and III** represent the three main DNA-protein complexes formed in the presence of the 47 bp $\alpha 2(I)$ procollagen promoter fragment. Complex I is present in all cell lines whilst complex II is predominantly present in nuclear extracts from SVWI-38 fibroblasts and complex III is present in extracts from CT-1 fibroblasts. The arrow indicates a complex composed of a mixture of proteins that bind to either the G/CBE or CME, however, this complex is competed mainly by G/CBE oligonucleotides.

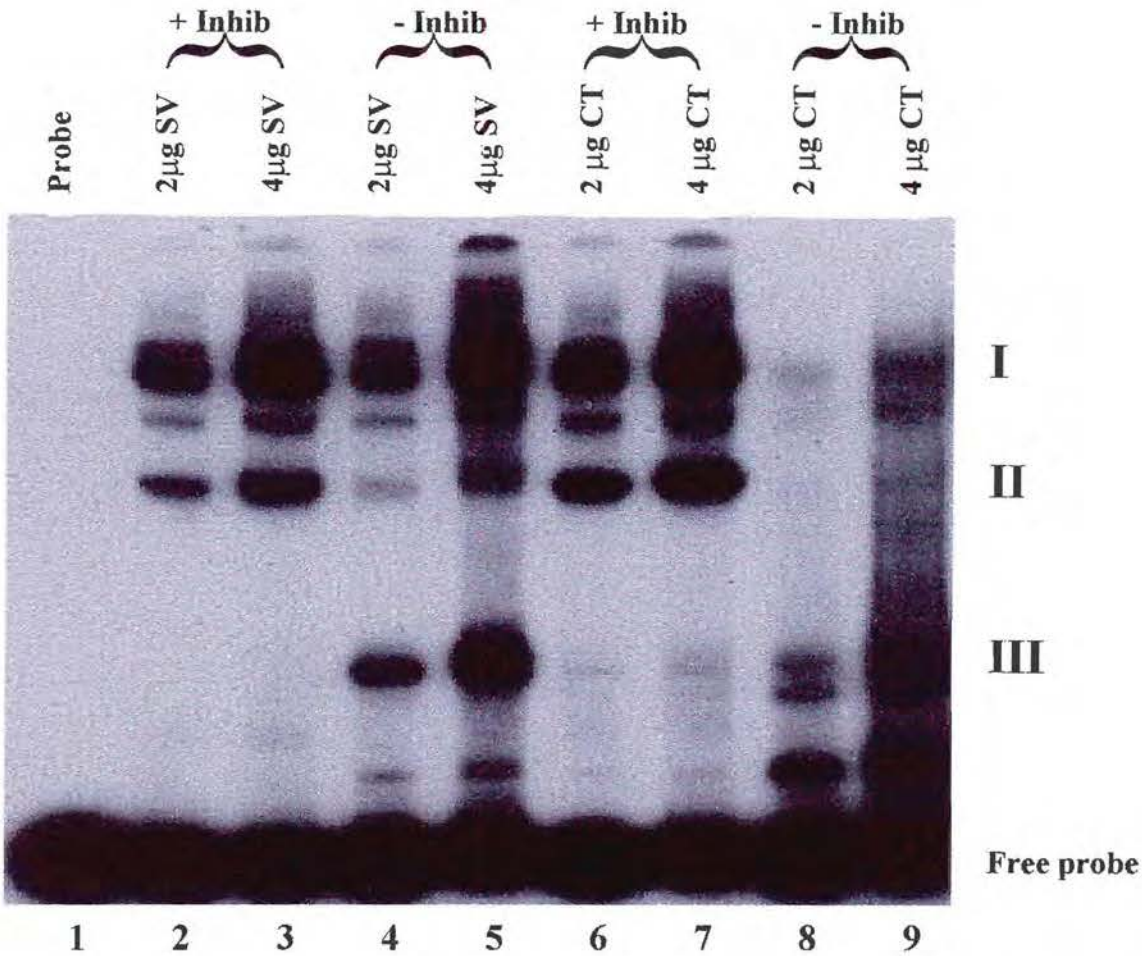


Figure 2.2: **Electrophoretic mobility shift assay of nuclear extracts from SVWI-38 and CT-1 fibroblasts.** Nuclear proteins were prepared from CT-1 (CT, lanes 6-9) and SVWI-38 (SV, lanes 2-5) fibroblasts using a modified Dignam method (245) as described in section 6.6.1 in the presence (+ inhib) or absence (- inhib) of protease inhibitor cocktail containing 0.5 mM PMSF, 1 μg/ml leupeptin and 1 μg/ml pepstatin A. 2-4 μg of nuclear proteins were incubated with a ^{32}P-labelled -107 to -60 α2(I) procollagen promoter fragment and DNA-protein complexes were separated by electrophoresis on a non-denaturing 5% polyacrylamide gel. **I**, **II** and **III** indicate the complexes formed with the α2(I) procollagen promoter fragment with complex III proteins being present in extracts prepared in the absence of protease inhibitors. Degradation of both complex I and II is more extensive in CT-1 extracts.

in SVWI-38 fibroblasts through an unknown mechanism. These observations were made from both experiments using either a probe that recognises complex I, II and III or a probe specific for complex II and its degradation products (Figures 2.2 and 2.3, respectively). Leupeptin is an inhibitor of serine proteases like trypsin, plasmin and porcine kallikrein, and cysteine proteases such as papain, cathepsins B, H and L. Pepstatin A, on the other hand, is an inhibitor of aspartic proteases such as cathepsin D, renin and pepsin. These results, however, do not provide conclusive evidence as to the specific protease responsible for the degradation of complex II proteins to complex III proteins. To confirm proteolysis, nuclear proteins were allowed to bind to the promoter fragment, the complexes formed were treated with different proteases and DNA-protein complexes were analysed by EMSA. It is evident from figure 2.5 that complex I and II proteins contain protease-resistant DNA-binding domains or at least the DNA-binding domain is protected from protease digestion in the presence of DNA. Interestingly, treatment of nuclear extracts with proteases in the absence or presence of DNA produced no significant differences in DNA-protein interactions (Figure 2.6) but it is possible that the type and concentration of the protease as well as the incubation period will determine the extent of proteolysis. These findings are consistent with studies by Hooft van Huijsduijen *et al* (184) where protease treatment of the DNA-NF-Y complex was shown not to degrade the DNA-binding domain of the protein. In a separate report, Dreier *et al* (185) demonstrated that partial proteolysis of the transcription factor Stat3 generates fragments of characteristic sizes that have retained the ability to engage in sequence specific DNA binding. Although these results suggest that the DNA-binding domain of transcription factors is resistant to certain proteases, it should be emphasised that in the cytosol, there are numerous types of proteases (mainly of lysosomal origin) that could have a much more extensive damaging effect on the integrity of transcription factors in crude nuclear extracts. Thus, the conclusion from this section is that complex III is a degradation product of complex II and complex II proteins represent DNA-protein interactions formed on the CME box of the $\alpha 2(I)$ procollagen promoter. It would be very interesting to determine the role of conversion of complex II to complex III with regards to transcriptional regulation of the $\alpha 2(I)$ procollagen gene.

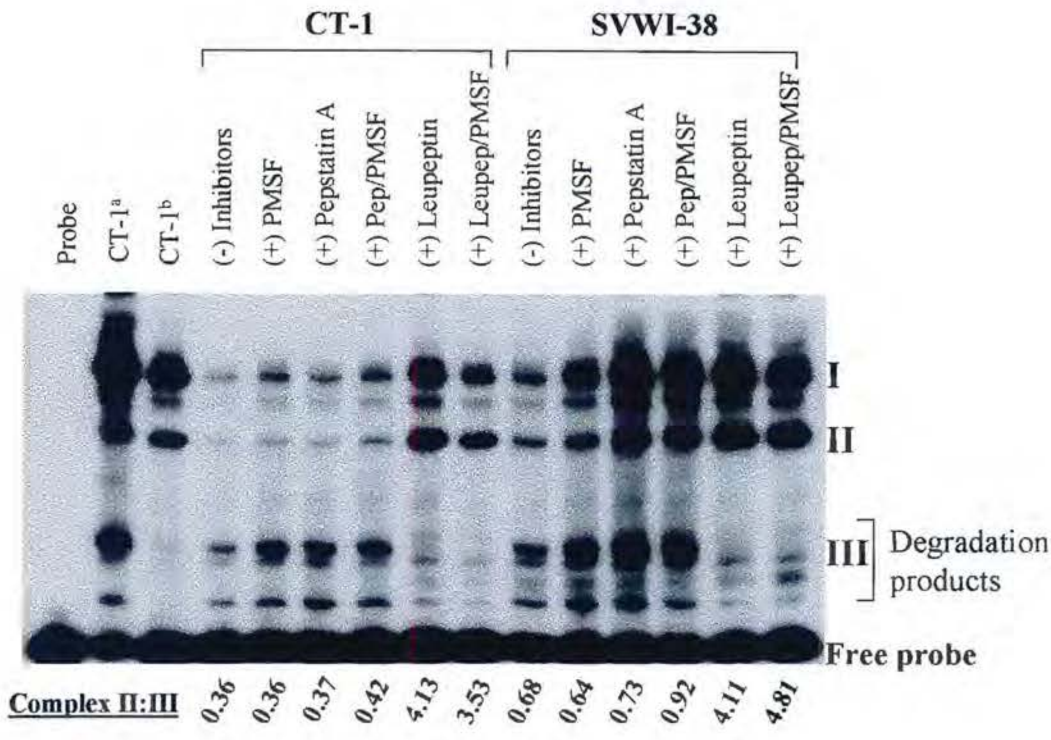


Figure 2.3: DNA binding activity of CT-1 and SVWI-38 nuclear extracts prepared in the presence of various combinations of protease inhibitors. Nuclear proteins were extracted in the absence or presence of PMSF, Leupeptin, Pepstatin A, PMSF/Leupeptin or PMSF/ Pepstatin A following a modified method of Dignam *et al* (245) as described in section 6.6.1. Proteins were incubated with a ^{32}P-labelled -107 to -60 fragment of the $\alpha 2$(I) procollagen promoter. DNA-protein complexes were resolved by electrophoresis on a non-denaturing 5% polyacrylamide gel and the gel was scanned with a Packard Instant Imager. The number of counts obtained for each band were used to calculate the ratio of complex II to complex III. **CT-1^a**, CT-1 extracts prepared as in Parker *et al* (13) in which leupeptin and pepstatin A were not included in the extraction buffers; **CT-1^b**, CT-1 nuclear extracts prepared in the presence of PMSF, Leupeptin and Pepstatin A. **I, II, and III** ~DNA-protein complexes. It is very clear that complex III is present in both CT-1 and SVWI-38 nuclear extracts prepared in the absence of leupeptin and so are other degradation products that still retained DNA binding activity. The ratio of complex II to complex III in the extracts prepared in the absence of leupeptin was found to be 2 ± 0.2 fold higher for SVWI-38 compared to CT-1.

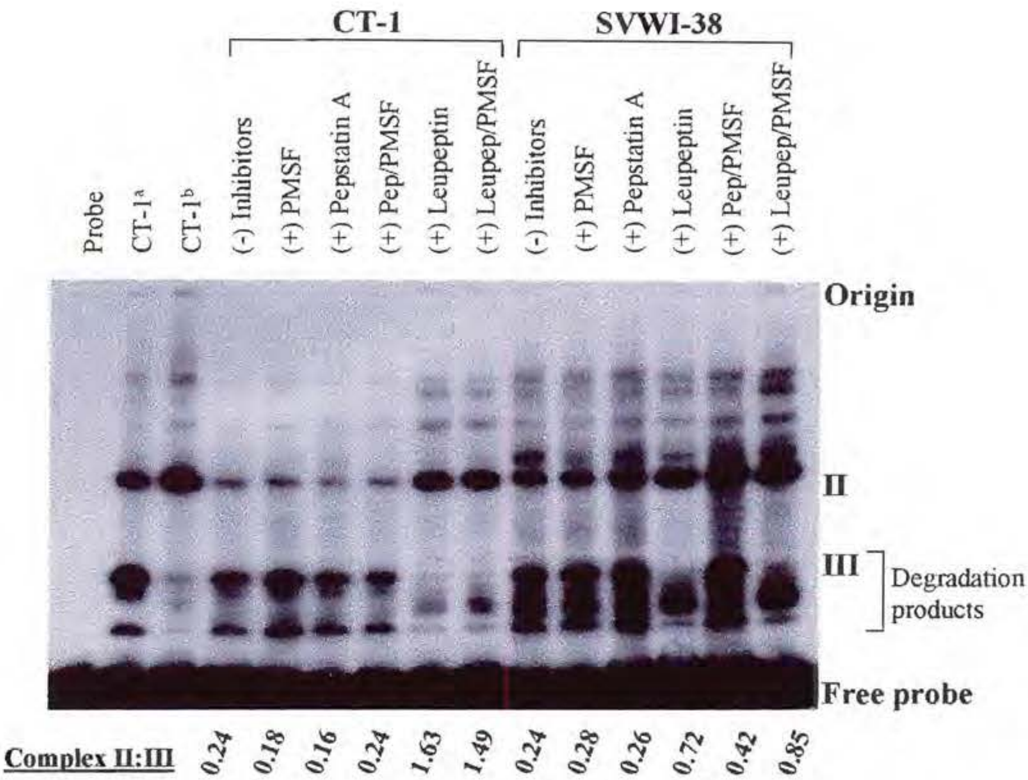


Figure 2.4: DNA binding activity of CT-1 and SVWI-38 nuclear extracts prepared in the presence of protease inhibitors. Nuclear proteins were prepared as described in figure 2.3. Proteins were incubated with a ^{32}P-labelled fragment of the α2(I) procollagen promoter that encompasses only the CME region located between -67 and -78. DNA-protein complexes were resolved by electrophoresis on a non-denaturing 5 % polyacrylamide gel, the dried gel was scanned with a Packard Instant Imager and the ratio of complex II to complex III calculated. The ratio of complex II to complex III in the absence of leupeptin was 1.5 ± 0.3 fold higher for SVWI-38 in comparison to CT-1 nuclear extracts. **CT-1^a**, CT-1 extracts prepared as in Parker *et al* (13); **CT-1^b**, CT-1 nuclear proteins were extracted in the presence of a protease inhibitor cocktail (leupeptin, pepstatin A and PMSF). The extent of degradation of complex II proteins to complex III proteins is greater in CT-1 compared to SVWI-38. Although some degradation products are present in lanes containing leupeptin, no complex III is present in these extracts.

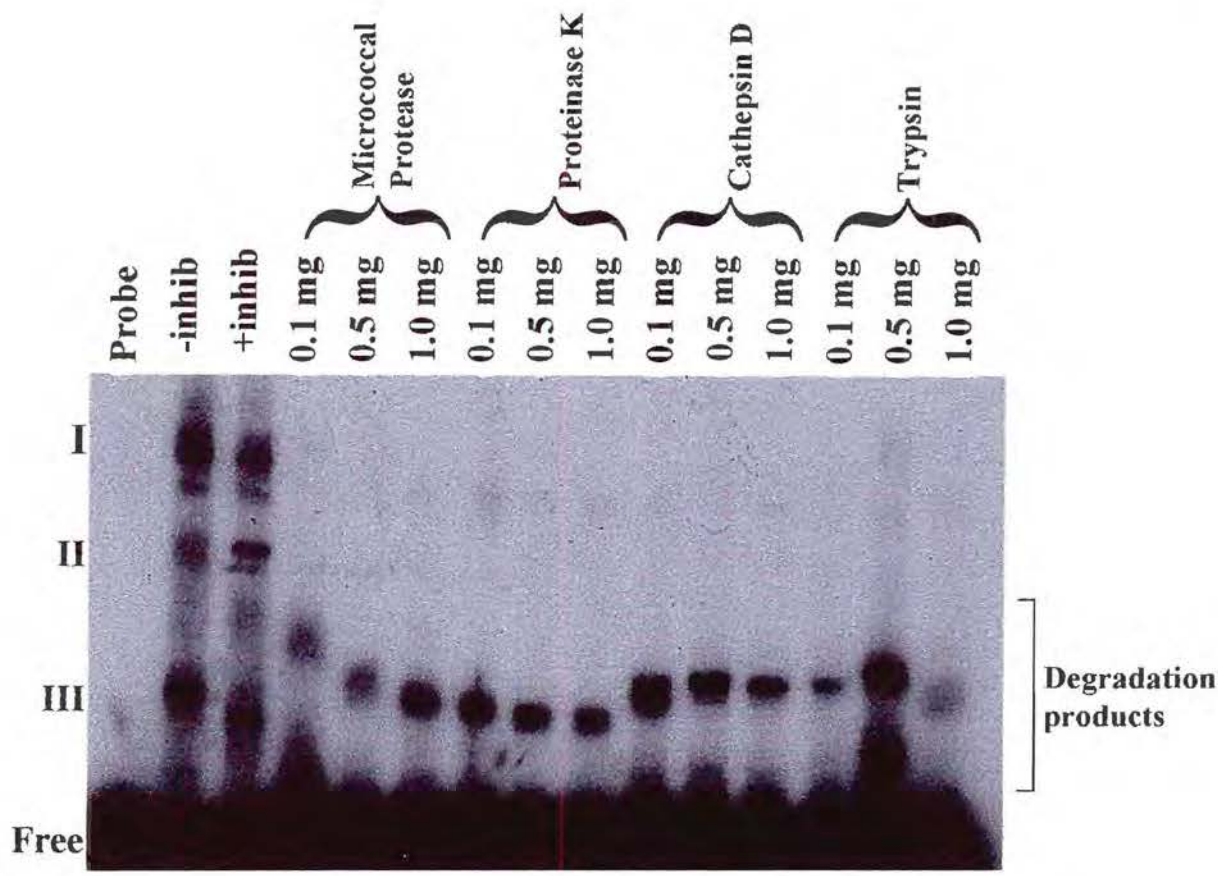


Figure 2.5: **The effect of protease treatment on DNA-protein interactions.** Nuclear proteins prepared according to the modified Dignam method (section 6.6.1) were incubated with the ^{32}P-labelled -107 to -60 α2(I) procollagen promoter on ice for 20 minutes, the reaction mixture was treated with 0.1, 0.5 or 1.0 mg of the indicated proteases, incubated for a further 5 min on ice and DNA-protein complexes were resolved by electrophoresis on a non-denaturing 5 % polyacrylamide gel. Complex III proteins, present in nuclear extracts prepared in the absence of protease inhibitors were retarded to the same extent as some of the degradation products obtained with proteinase K and cathepsin D. The molecular weight of the complex depends on the cleavage site for each specific protease.

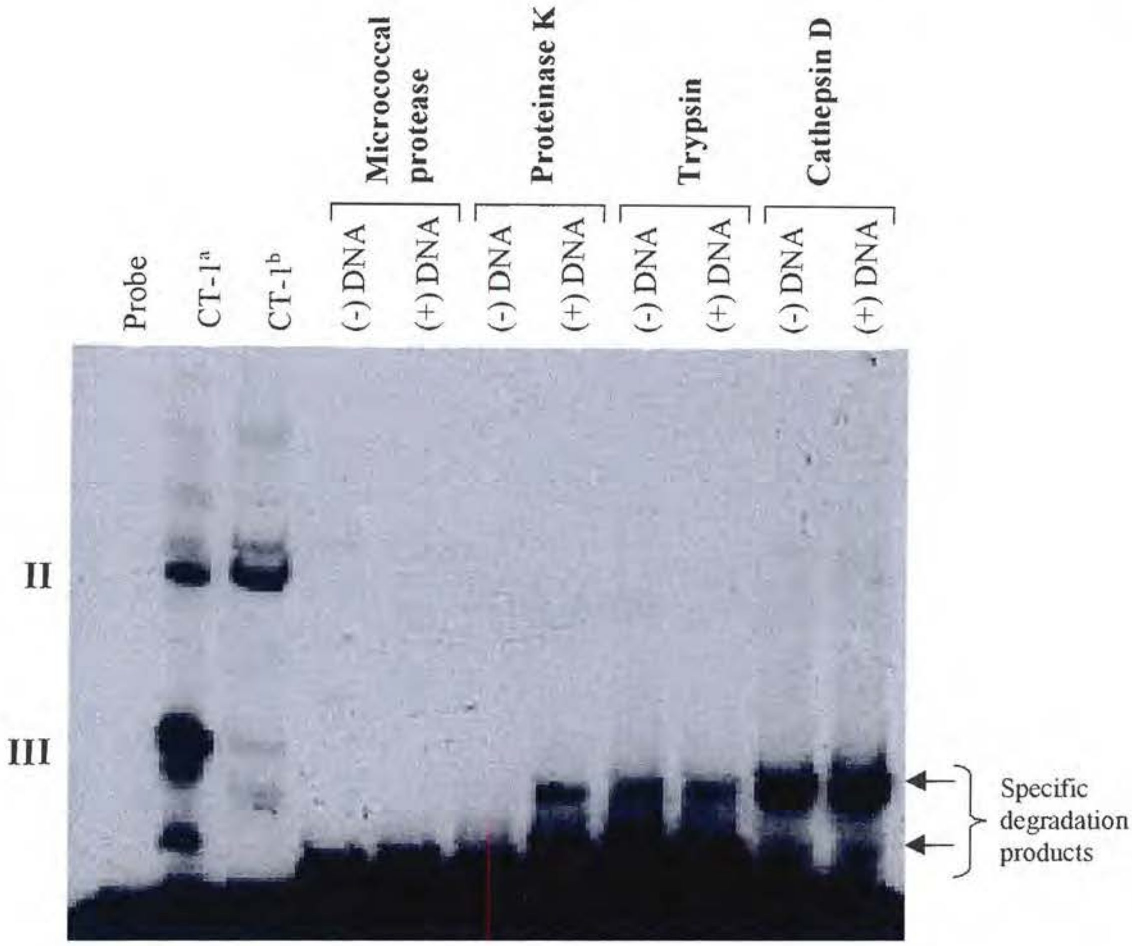


Figure 2.6: DNA-protein interactions of nuclear proteins treated with proteases. Nuclear proteins were prepared as described in section 6.6.1., treated with 0.5 mg/ml of the indicated proteases in the absence (-) or presence (+) of a ^{32}P-labelled -78 to -67 $\alpha 2(I)$ procollagen promoter fragment on ice and DNA-protein complexes were resolved by electrophoresis on a non-denaturing 5 % polyacrylamide gel; the gel was subsequently exposed to X-ray film for 16 hours at -70 °C. The probe used in this assay is specific for complex II proteins and degradation products of complex II. Treating proteins with proteases in the absence or presence of DNA generated specific products that migrate faster than complex III but they still retained DNA binding activity. It is possible that the degree of degradation depends on the concentration of and the incubation period with a specific protease. **CT-1^a**, CT-1 proteins prepared as described in Parker *et al* (13); **CT-1^b**, CT-1 proteins prepared as described in section 6.6.1.

2.2.2. Site-directed mutagenesis of the $\alpha2(I)$ procollagen promoter

Several protocols have been described in the literature ranging from the use of a single primer with single-stranded DNA as a template to the use of multiple primer sets to introduce specific mutations in a gene by PCR. The most successful and easier method seems to be the PCR-based one. The main advantage of this method is the fact that the desired mutation is obtained with close to 90-100 % efficiency. The disadvantage of this method lies in the fact that *Taq* DNA polymerase tends to introduce errors, but these can be reduced by using a *Taq* DNA polymerase with a higher fidelity. The method of Higuchi *et al* (186) was used to introduce mutations in the $\alpha2(I)$ procollagen promoter (Figure 2.7). In this protocol, two primary PCR reactions produced two overlapping DNA fragments, both bearing the same mutation in the overlap region. The overlap in sequence allows the fragments to anneal and only one of the hybrids can be extended by DNA polymerase to produce a duplex fragment. The other hybrid has recessed 5′-ends, and is thus not extended by the polymerase and is effectively lost from the reaction mixture. Templates used in the PCR reaction were pNJs400 19A and -375 COLCAT. The pNJs400 19A plasmid contains a 400 bp SphI fragment consisting of 50 bp of exon 1 and 350 bp of the promoter region of the $\alpha2(I)$ procollagen gene subcloned into the SphI site of pUC19. The -375 COLCAT plasmid contains the 50 bp region of exon 1 and a 375 bp sequence of the $\alpha2(I)$ procollagen gene upstream of the transcription start site cloned into p8CAT. The first PCR reactions were carried out with complimentary primers designated A and A′ together with the respective primers that anneal to the different vectors, namely, M13 Forward primer (B) and p8CAT Forward primer (C) for pNJs400 19A and -375 COLCAT, respectively. Table 2.1 summarises the PCR conditions used to generate the PCR products while the primer sets are listed in Table 2.2. The two products (Figure 2.7A) were resolved by agarose gel electrophoresis, bands of interest excised, and DNA from the two PCR reactions, i.e. 303 and 218 bp fragments, were co-eluted. This mixture was then used in a second PCR reaction in the presence of the flanking primers B and C only. The final PCR products were digested with Hind III and Pst I for cloning into p8CAT (see Appendix for a map of p8CAT) to generate the mutant -343 COLCAT promoter constructs (Figure 2.7C, 2.7D, Figure 2.8). To generate the -107 COLCAT constructs, the -343 COLCAT vectors were digested with Sma I which cuts the insert at -107 as well as p8CAT on the 5′ end of the insert to release a 236 bp fragment (Figure 2.7D, Figure 2.8). To confirm the sizes of the cloned fragments, plasmid DNA was digested with EcoRI, Hind III and Pst I, or Hind III and Sma I (Figure 2.8). All mutants were confirmed by dideoxy sequencing using a Col+1 primer (see Table 2.2) which anneals to the promoter region between -1 and -25 such that the sequence upstream of the promoter could be determined. Figures 2.9A-F are the sequenced regions of mutations introduced in the α2(I) procollagen promoter. The presence of deletions in Mut US and Mut CME3 (ΔCME3) might have resulted from "slipping" of the polymerase during amplification. Since these sequences could also be considered as mutations in the respective *cis*-elements, they were nevertheless used in subsequent transfection experiments.

Table 2.1: PCR conditions for amplification of the α2(I) procollagen promoter using the primers shown in Table 2.2

Cycle no.	Temp.	Time
Cycle 1	94 °C	2 minutes
Cycle 2-25	94°C	30 seconds
	55°C	30 seconds
	72°C	1 minute
Cycle 26	72°C	10 minutes

PCR reactions contained 2-50 ng plasmid DNA, 100 pmol primer and 1U Takara™ *Taq* DNA polymerase in a total volume of 50 μl; the other components are described in section 6.4.7.

2.2.3. The inverted CCAAT box and the adjacent, inverted GGAGG sequence are important for promoter activity

Two sets of mutations were introduced in the previously described G/CBE element of the α2(I) procollagen promoter, one in the CCAAT box and the other in the upstream GGAGG box. The mutated sequences were cloned into p8CAT and transfected into

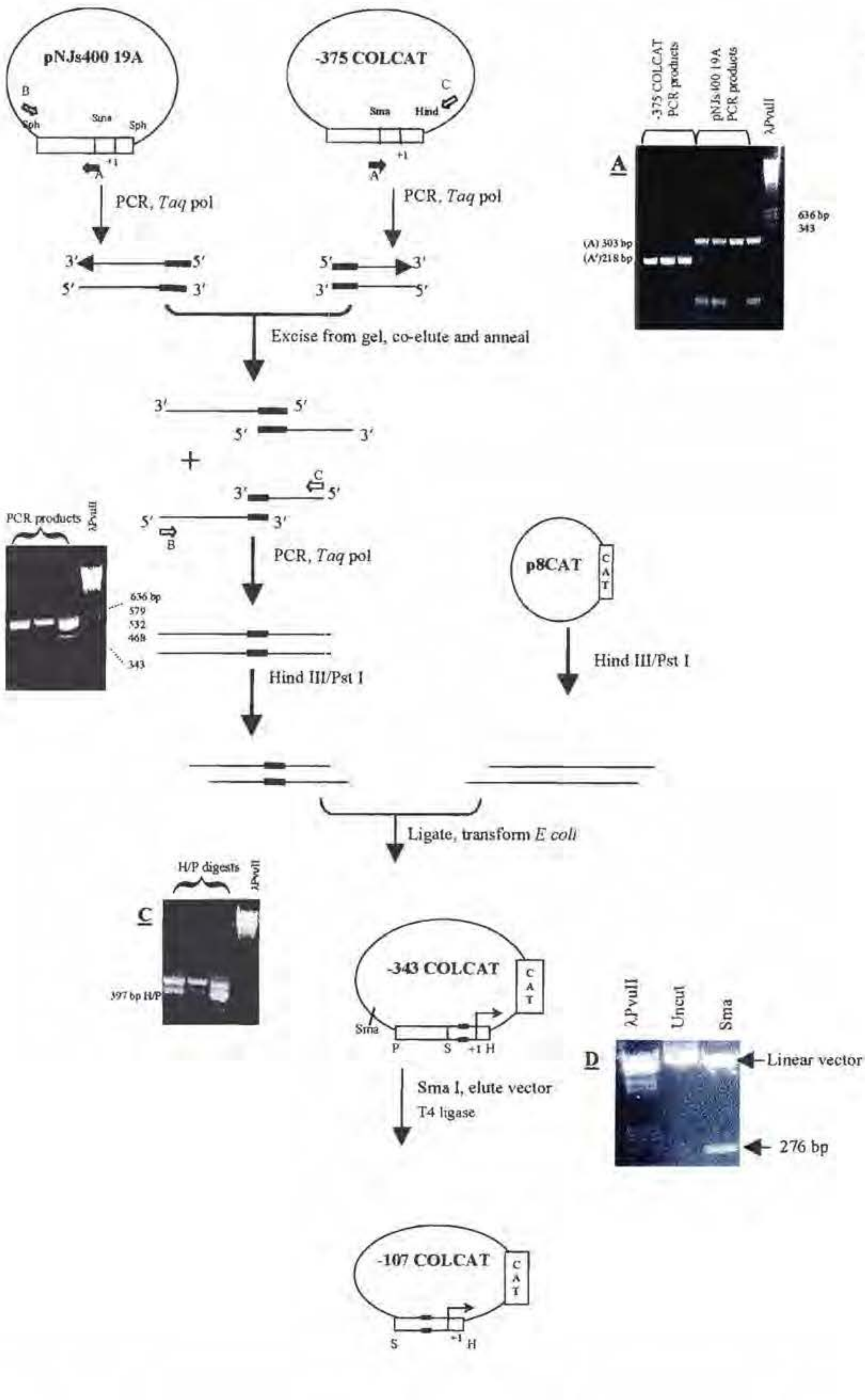


Figure 2.7: **Site-directed mutagenesis PCR of the α2(1) procollagen promoter.** Plasmids pNJ400s 19A and -375 COLCAT (see text for details) were amplified with primers A/B and A'/C (see Table 2.1and 2.2), respectively, to introduce point mutations in the α2(I) procollagen promoter. PCR products **(A)** were co-eluted from agarose gels (section 6.4.2), annealed and re-amplified with primer B and C to generate a 521 bp product **(B)**. The DNA was digested with HindIII and PstI **(C)** and cloned into p8CAT (see Appendix for the map of p8CAT) to generate the -343 COLCAT plasmid. This plasmid was digested with SmaI **(D)** and re-ligated to generate the -107 COLCAT plasmid. H; HindIII, P; PstI, S; SmaI.

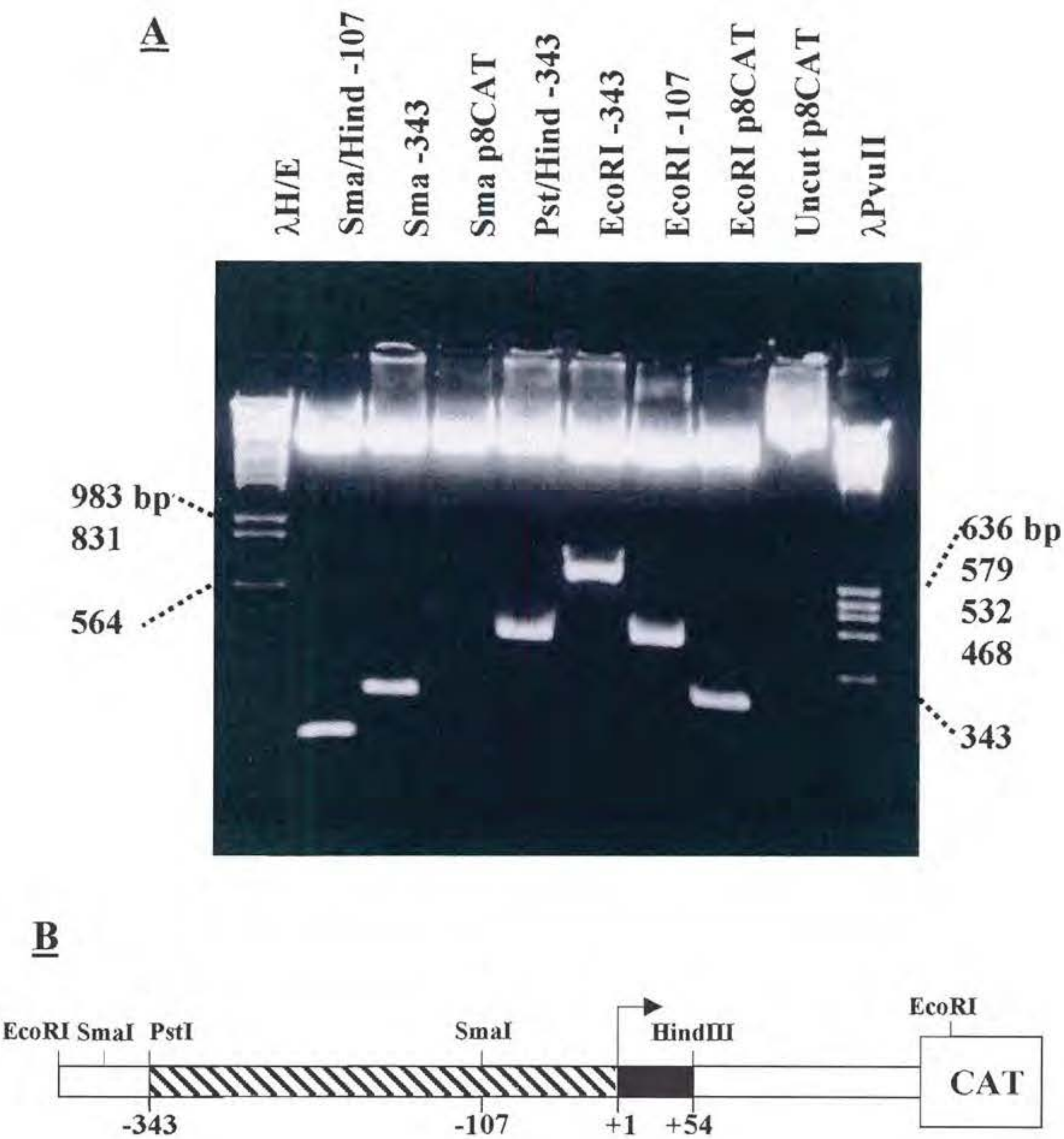


Figure 2.8: Restriction enzyme analysis of the α2(I) procollagen promoter-CAT constructs. Plasmid DNA was digested with the indicated restriction enzymes to release the promoter fragments cloned into p8CAT. Digestion with EcoRI releases the cloned promoter fragment plus 250 bp of p8CAT vector. **(A)** DNA fragments were resolved by electrophoresis on a 1.5 % agarose gel and stained with ethidium bromide. **(B)** Restriction map of the -343 COLCAT construct.

; p8CAT sequences, ; α2(I) procollagen promoter sequences,

; α2(I) procollagen exon sequences,

CAT; Chloroamphenicol acetyl transferase gene

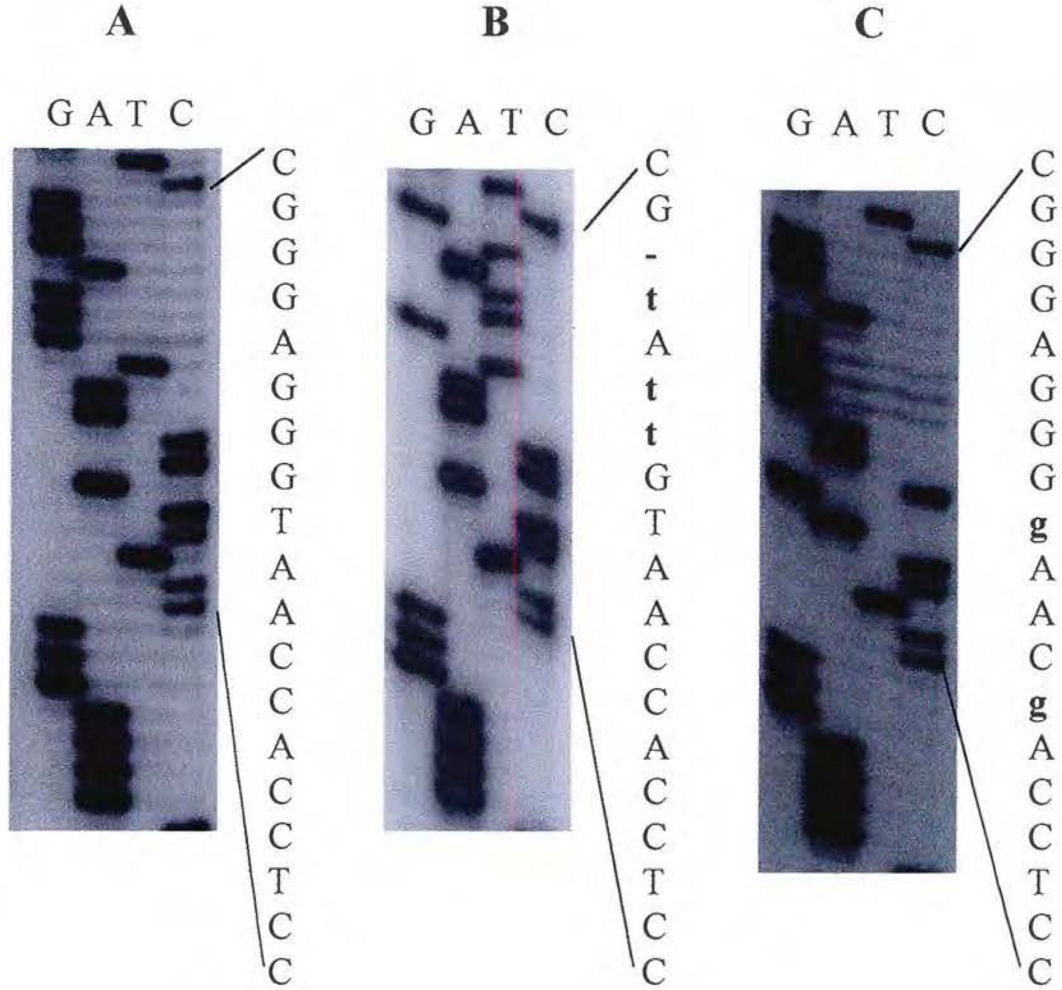


Figure 2.9: Analysis of mutations in the G/CBE region of the $\alpha 2(I)$ procollagen promoter. Mutations in the G/CBE element of the $\alpha 2(I)$ procollagen gene were introduced by site-directed mutagenesis as described in figure 2.5. The promoter was subsequently cloned into p8CAT reporter plasmid and sequenced by the Sanger dideoxy sequencing protocol as described in section 6.4.6 using the Col+1 primer (see Table 2.2). The lower case letters indicate the mutations. **(A)** Wild type, **(B)** Mut US, and **(C)** Mut CCAAT.

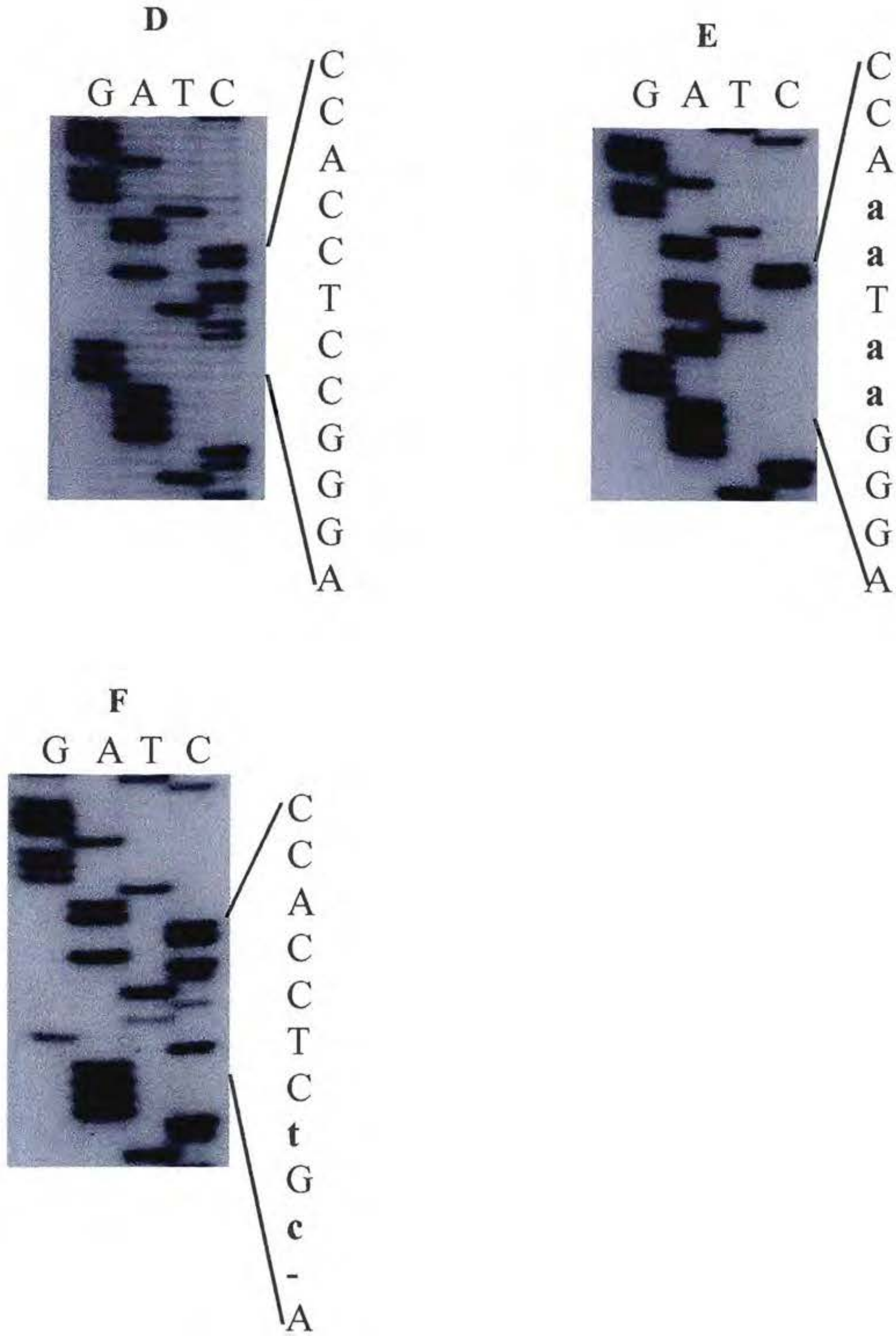


Figure 2.9 (continued): Analyses of mutations in the collagen modulating element (CME) of the $\alpha 2(I)$ procollagen promoter. Mutations in the CME region of the promoter of the $\alpha 2(I)$ procollagen gene were introduced by PCR as described in figure 2.5 and mutations confirmed by Sanger dideoxy method (section 6.4.6). **(D)** Wild type, **(E)** Mut CME1, **(F)** Mut ΔCME.

the two transformed WI-38 cell lines, namely, the α2(I) collagen producing cell line, CT-1 and non-α2(I) collagen producing cell line, SVWI-38 fibroblasts. Cells were transfected with 10-20 μg of plasmid DNA by the calcium-phosphate precipitation method as described in section 6.2. To control for transfection efficiency, cells were co-transfected with 2-5 μg of a cmvβgal plasmid, which contains the β-galactosidase gene under the control of the cytomegalovirus promoter. Cells were

Table 2.2: Primer sets used in site-directed PCR mutagenesis.

Oligonucleotide	Sequence (5′ → 3′)
Wild type A	5′ CCC AGC CCT CCC ATT GGT GGA GGC CCT 3′
Wild type A′	5′ AGG GCC TCC ACC AAT GGG AGG GCT GGG 3′
Mut US A	5′ CCC AGC AAT AAC ATT GGT GGA GGC CCT 3′
Mut US A′	5′ AGG GCC TCC ACC AAT GTT ATT GCT GGG 3′
Mut CCAAT A	5′ CCC AGC CCT CCC CTT GCT GGA GGC CCT 3′
Mut CCAAT A′	5′ AGG GCC TCC AGC AAG GGG AGG GCT GGG 3′
Mut CME1 A	5′ CCC AGC CCT CCC ATT GGT TTA TTC CCT 3′
Mut CME1 A′	5′ AGG GAA TAA ACC AAT GGG AGG GCT GGG 3′
Mut CME3 A	5′ CCC AGC CCT CCC ATT GGT GGA GAC GTT 3′
Mut CME3 A′	5′ AAC GTC TCC ACC AAT GGG AGG GCT GGG 3′
M13 Forward	5′ GTT TTC CCA GTC ACG ACG TTG TA 3′
p8CAT Forward	5′ GGA TAT ATC AAC GGT GGT ATA TCC 3′
COL+1	5′ AAA ATA ATA AAG CCC GGA TCT GCC C 3′

The base substitutions are underlined in all mutated oligonucleotides, G/CBE and CME elements are underlined in the wild type sequence; A and A′ are complementary oligonucleotides.

harvested and CAT and β-galactosidase activities were determined as described in sections 6.2.5 to 6.2.6. The aim was to establish the role of the CCAAT-binding factor in the two fibroblast cell lines that differ in their ability to express the α2(I) procollagen gene. Mutations in the inverted GGAGG element (Mut US) reduced promoter activity by 7-fold for the -343 construct and by 12-fold for the -107 construct in CT-1 fibroblasts (Figure 2.10 and 2.12). The mutation in the inverted CCAAT box (Mut CCAAT) reduced promoter activity by approximately 9-fold for the -343 and 7-fold for the -107 constructs (Figure 2.10 and 2.12). In SVWI-38 fibroblasts, on the other hand, the activity of the Mut US promoter construct was reduced by 3- and 8-fold for the -343 and -107 constructs, respectively, as shown in figure 2.11 and 2.13 while a 3-fold and 9-fold reduction in promoter activity was observed with the Mut CCAAT for both the -343 and -107 constructs. As expected, the collagen producing CT-1 fibroblasts generally had a much higher promoter activity than the SVWI-38 cell line following transfection with the wild type construct. Both the GGAGG and CCAAT box mutations had negative effects on promoter activity in both CT-1 and SVWI-38 fibroblasts. These results indicate that the integrity of the G/CBE element is important for the expression of the $\alpha2(I)$ procollagen gene, irrespective of the level of expression of the gene in a particular cell line.

2.2.4. DNA-protein interactions on mutated G/CBE oligonucleotides

The analysis of promoter activity in transient transfection assays assumes that the different transcription factors recognise their respective *cis*-elements, thus eliciting its effect in the transcription of the reporter gene, in this case the CAT gene. To confirm that the presence of mutations interfered with binding of transcription factors, *in vitro* DNA binding assays were carried out in the presence of a wild type probe and the different mutant oligonucleotides were used as competitors. The wild type and mutant DNA fragments (169 bp) were released from -107 COLCAT constructs by digestion with Sma I and Hind III and purified from polyacrylamide gels by the crush-soak method (section 6.4.2.3). Figure 2.14 illustrates the effects of the two upstream mutations on DNA-protein interactions. There was competition in the presence of the wild type construct, whilst the presence of the mutant oligonucleotides did not result in competition of complex I, and complex II was competed out because the sequence recognised by complex II binding proteins was not mutated in this oligonucleotide. Similar DNA-protein interactions and competition with mutant oligonucleotides were obtained for both CT-1 and SVWI-38 nuclear extracts (data not shown).

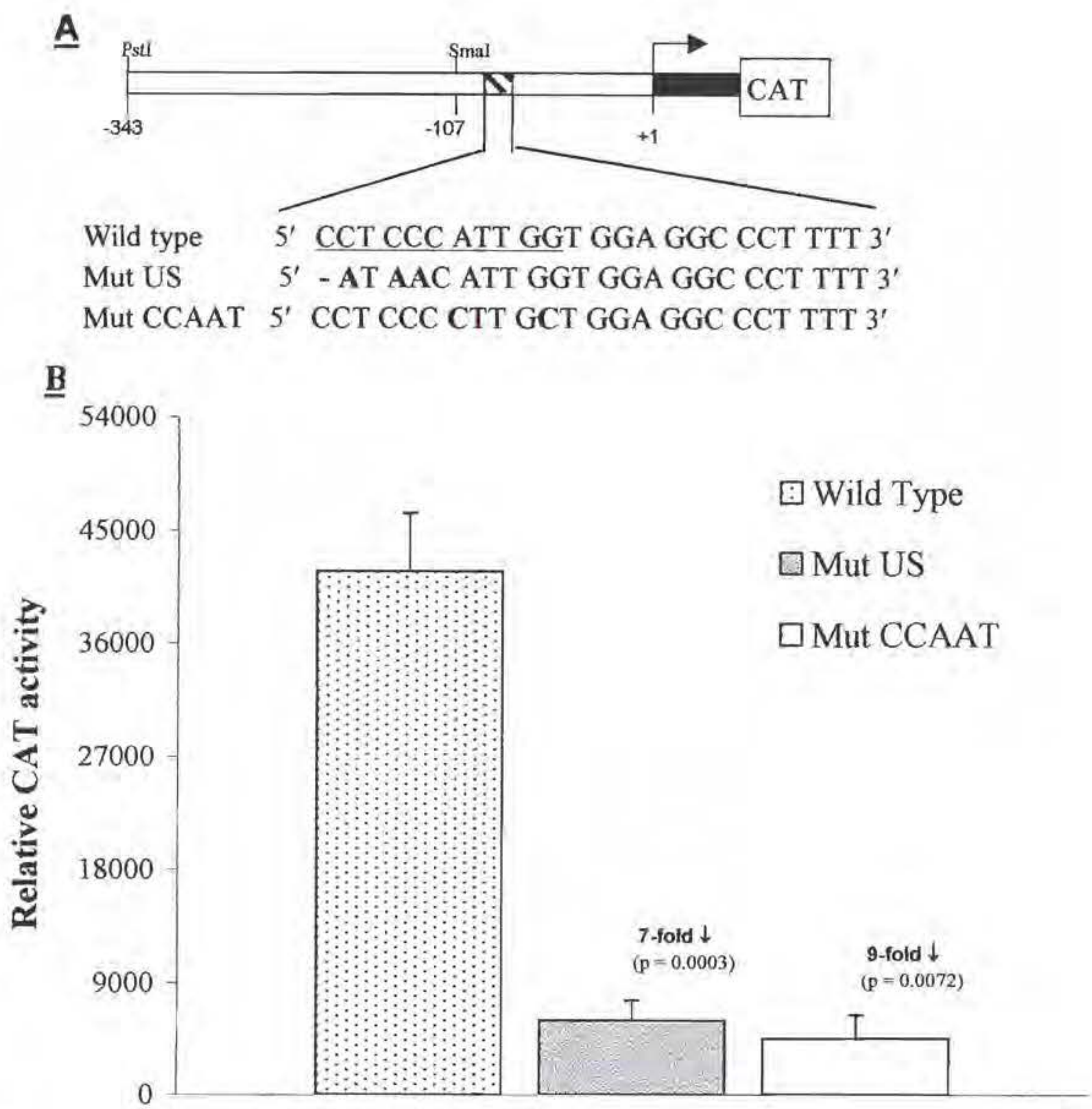


Figure 2.10: The effect of G/CBE mutations on the activity of the -343 COLCAT construct CT-1 fibroblasts. (A) Schematic representation of the -343 COLCAT constructs showing the region of the promoter where mutations were introduced; the G/CBE element is underlined; point mutations are shown in bold letters. **(B)** CT-1 fibroblasts were transfected with 20μg of wild type or mutant -343 COLCAT constructs by the calcium phosphate precipitation method (see section 6.2); 5μg of a cmvβgal plasmid was co-transfected to control for transfection efficiency. Cells were incubated in fresh medium for 48 hours, harvested and the post-nuclear supernatant (PNS) was obtained following 3 cycles of freeze-thawing in liquid N_2 and centrifugation at 13 000 rpm at 4 ° C. The supernatant was used in CAT and βgalactosidase assays as described in section 6.2.2. Graphical representation of the activity of the -343 COLCAT construct containing the indicated mutations. Promoter activity is expressed as the ratio of CAT activity to βgalactosidase activity. Results represent the mean ± SD of three separate experiments performed in triplicate; p values were calculated using the student *t* test.

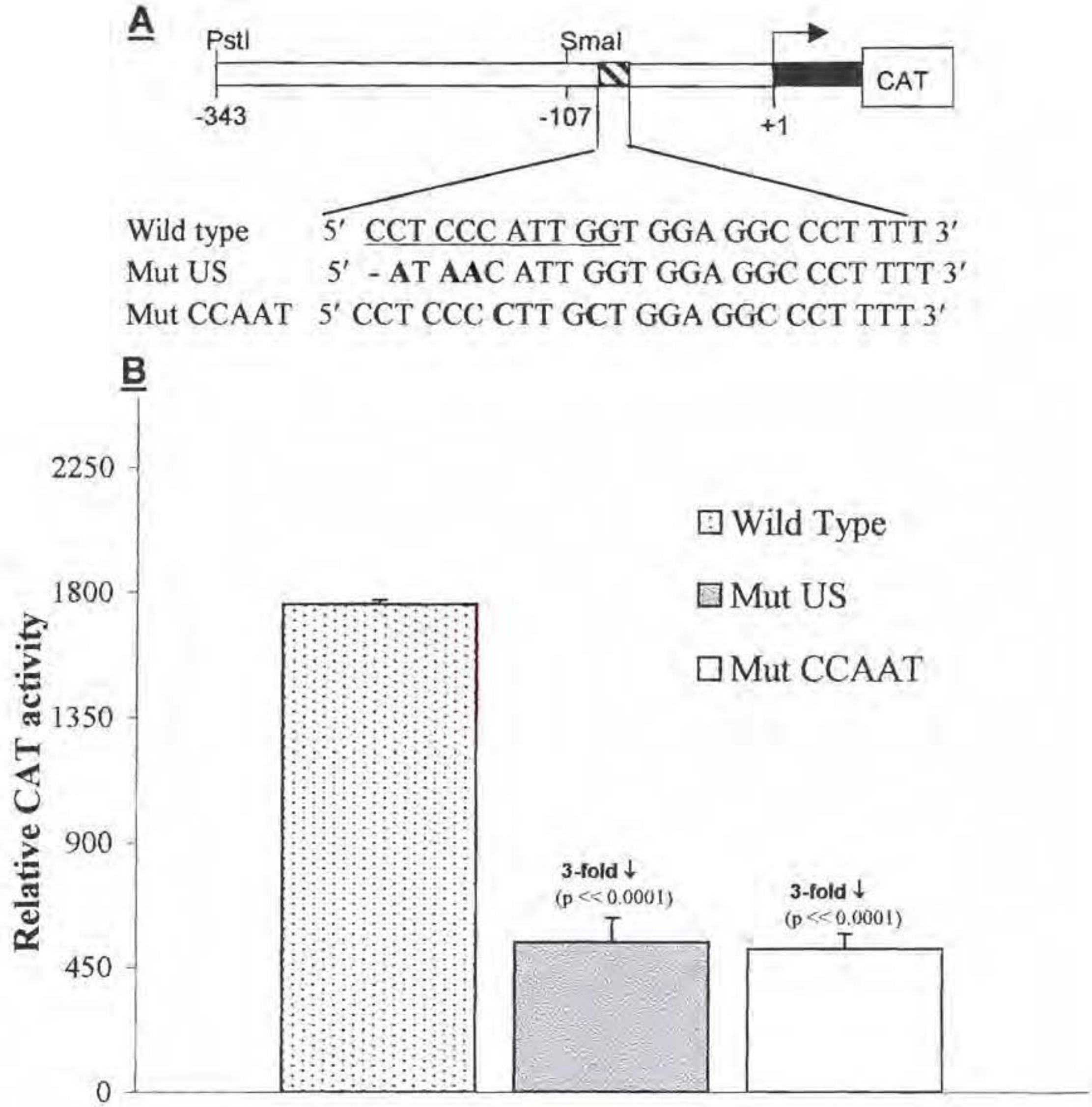


Figure 2.11: **The effect of G/CBE mutations on the activity of the -343 COLCAT in SVWI-38 fibroblasts.** (**A**) Schematic representation of the -343 COLCAT constructs used in transfection experiments. **(B)** SVWI-38 fibroblasts were transfected with -343 COLCAT constructs that contain mutations in the G/CBE region of the promoter as described in section 6.2. Cells were harvested and the CAT activity determined. The CAT activity is expressed relative to β-galactosidase activity (see section 6.2.2). Results represent the mean ± SD of three separate experiments done in triplicate, p values, calculated using the student *t* test, were much, much smaller than 0.0001.

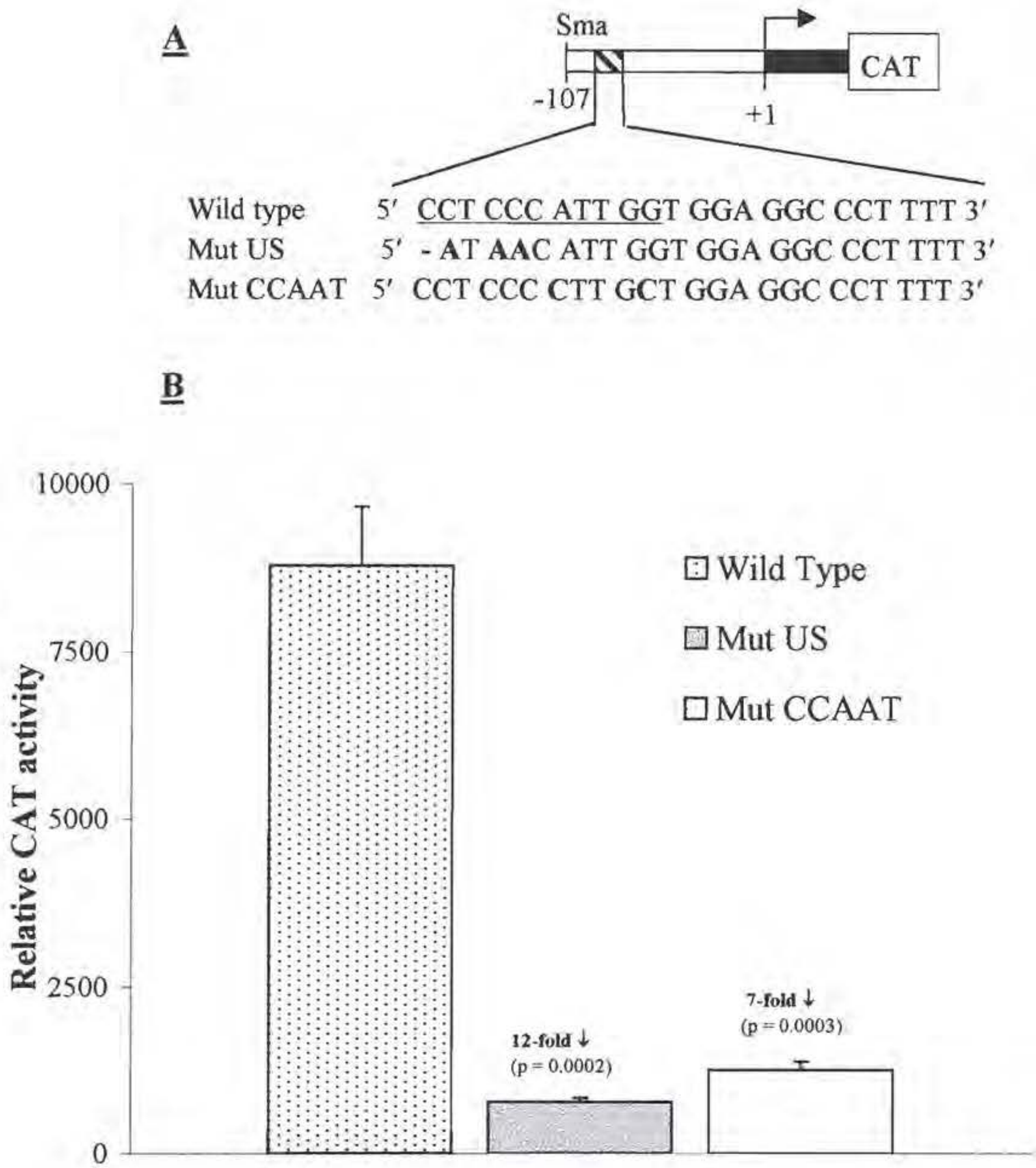


Figure 2.12: The effect of G/CBE mutations on the activity of the -107 COLCAT constructs in CT-1 fibroblasts. (A) A schematic outline of the -107 α2(I) procollagen promoter-CAT construct; the G/CBE region of the promoter is underlined; point mutations are indicated in bold letters. **(B)** Fibroblasts were transfected with 20μg of the -107 COLCAT constructs containing mutations in the G/CBE region of the promoter as described previously. Transfected cells were harvested and the post-nuclear supernatant was assayed for CAT and β-galactosidase activities as described in section 6.2. Results represent the mean ± SD of three separate experiments done in triplicate and p values were determined by the student *t* test.

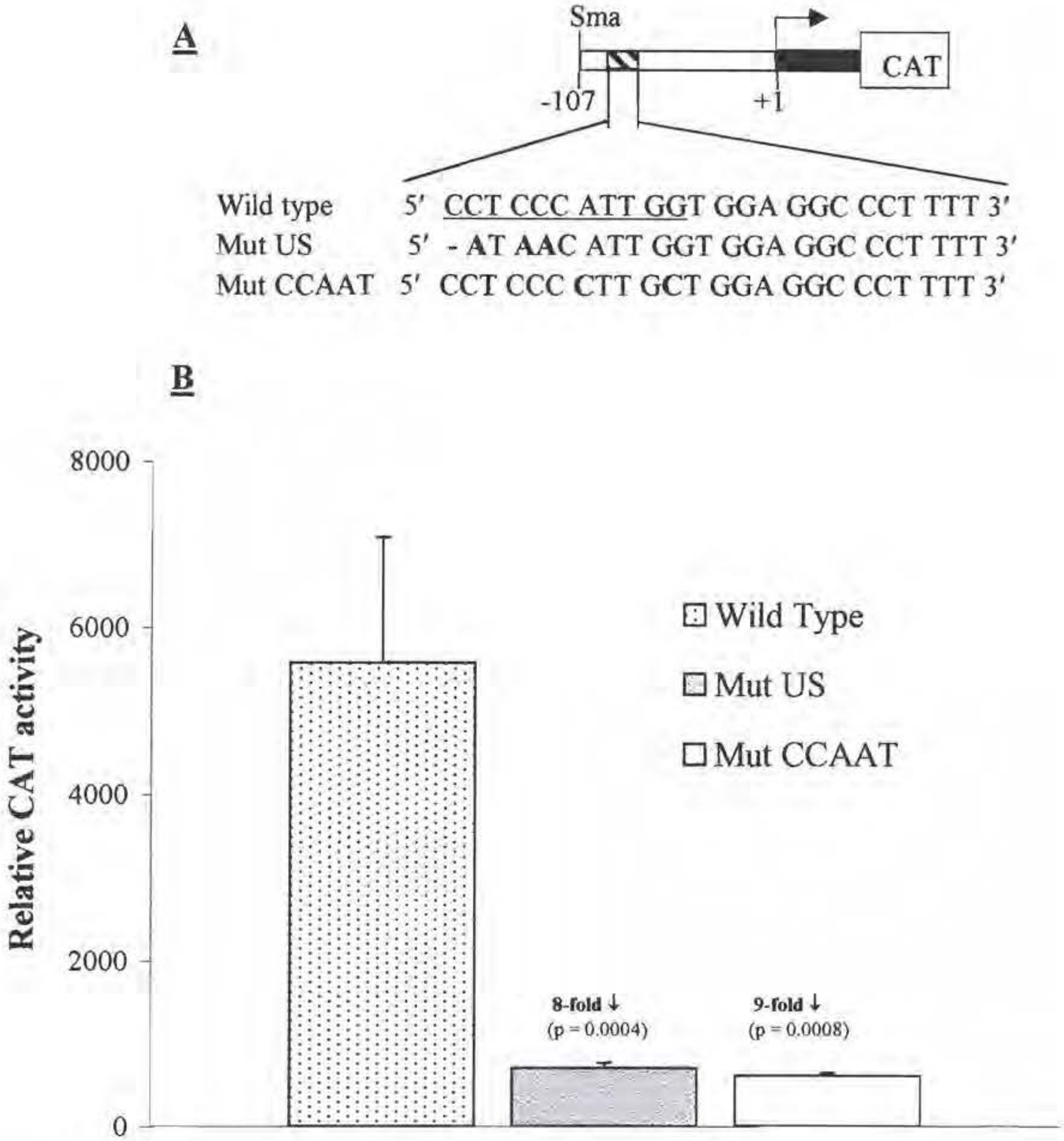


Figure 2.13: The effect of G/CBE mutations on the activity of the -107 COLCAT constructs in SVWI-38 fibroblasts. **(A)** A schematic representation of the -107 α2(I) procollagen promoter-CAT constructs; the G/CBE region is underlined and point mutations are indicated in bold letters. **(B)** SVWI-38 fibroblasts were transfected with -107 COLCAT constructs containing mutations in the G/CBE region of the promoter as described in figure 2.10. Enzyme assays were performed with the post-nuclear supernatant as described in section 6.2. Results represent the mean ± SD of three separate experiments carried out in triplicate and all p values were calculated using the student *t* test.

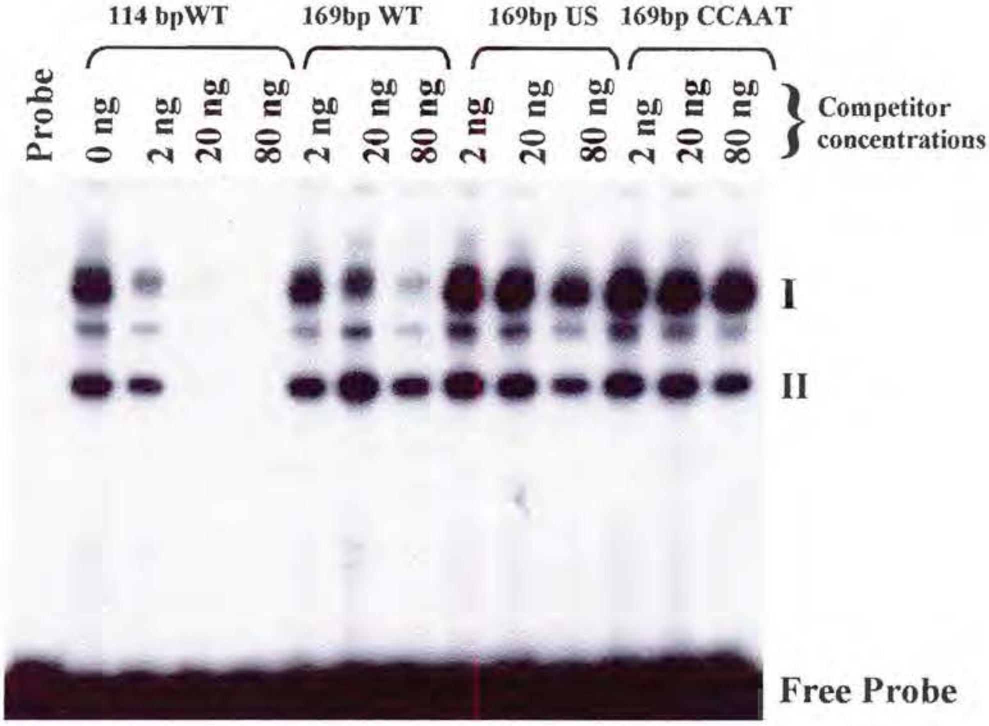


Figure 2.14: **Competition of DNA-protein complexes by mutated CCAAT box and the upstream GGAGG sequences.** CT-1 nuclear proteins were incubated with ^{32}P-labelled -107 to -60 α2(I) procollagen promoter (114 bp) and DNA-protein complexes were resolved by electrophoresis on a non-denaturing 5 % polyacrylamide gel. For competition EMSA, -107 COLCAT constructs were digested with SmaI and HindIII to release a 169 bp fragment (see figure 2.8) which was eluted from a polyacrylamide gel by the crush-soak method described in section 6.4.2.3. For competition EMSA reactions, nuclear proteins were incubated with poly dI.dC and the indicated concentrations of competitor DNA at room temperature for 10 minutes, a ^{32}P-labelled probe (114 bp) was added and the incubation was continued on ice for 30 min. DNA-protein complexes were separated by electrophoresis on a non-denaturing 5% polyacrylamide gel. Although both the 114 bp and 169 bp WT fragments contain wild type sequences, the 114 bp DNA fragment was a better competitor since it contains promoter sequences only between -107 and -60 whereas the 169 bp DNA fragment contains sequences between -107 and +54. Complex II proteins were sometimes competed at 80 ng of competitor DNA. WT; wild type, US; mutated upstream GGAGG box, CCAAT; mutated CCAAT box.

2.2.5. The CME binding protein(s) functions as a repressor of transcription

A second set of mutations was introduced into the CME of the α2(I) procollagen promoter. Two such mutations, Mut CME1 and Mut ΔCME3 (See Table 2.2), were cloned into p8CAT and transfected into the α2(I) procollagen expressing CT-1 and the non-expressing SVWI-38 fibroblasts. The promoter activity of the -343 Mut CME1 COLCAT construct increased by 1.6 fold over the wild type promoter construct in CT-1 fibroblasts while a 1.7 fold increase was observed with the ΔCME3 COLCAT construct (Figure 2.15). Transfection of the same plasmids into SVWI-38 fibroblasts also resulted in an increase in the activity of the promoter. The Mut CME1 COLCAT construct had a 1.9-fold increase in promoter activity whilst the ΔCME3 COLCAT construct had a 3-fold increase in activity (Figure 2.16). Surprisingly, the -107 COLCAT mutant constructs did not have any increased promoter activity, instead the activity decreased by 1.8- and 1.2-fold for Mut CME1 COLCAT and Mut ΔCME3 COLCAT constructs, respectively, in CT-1 fibroblasts (Figure 2.17). In SVWI-38 fibroblasts, the decrease was 1.7- and 1.2-fold for the Mut CME1 COLCAT and Mut ΔCME3 COLCAT constructs, respectively (Figure 2.18). No significant differences in promoter activity were observed between the -107 wild type and Mut ΔCME3 constructs. The observation from this section was the clear difference in the response of the two fibroblast cell lines to the presence of mutations in the CME element; although a significant increase was observed in both cell lines, it was higher in SVWI-38 fibroblasts. These results suggest that the CME is involved in DNA-protein interactions in SVWI-38 fibroblast that could possibly account for some of the down-regulation of the α2(I) procollagen gene. Hence, mutation of this element lead to a significant increase in the activity of the promoter since the inhibitory effect of these interactions was removed.

2.2.6. DNA-protein interactions on mutated CME oligonucleotides

In order to determine the effect of the mutations on *in vitro* DNA-protein interactions, electrophoretic mobility shift assay (EMSA) studies were performed as described in section 6.6.2. Binding reactions included the 169 bp Sma I/Hind III fragment,

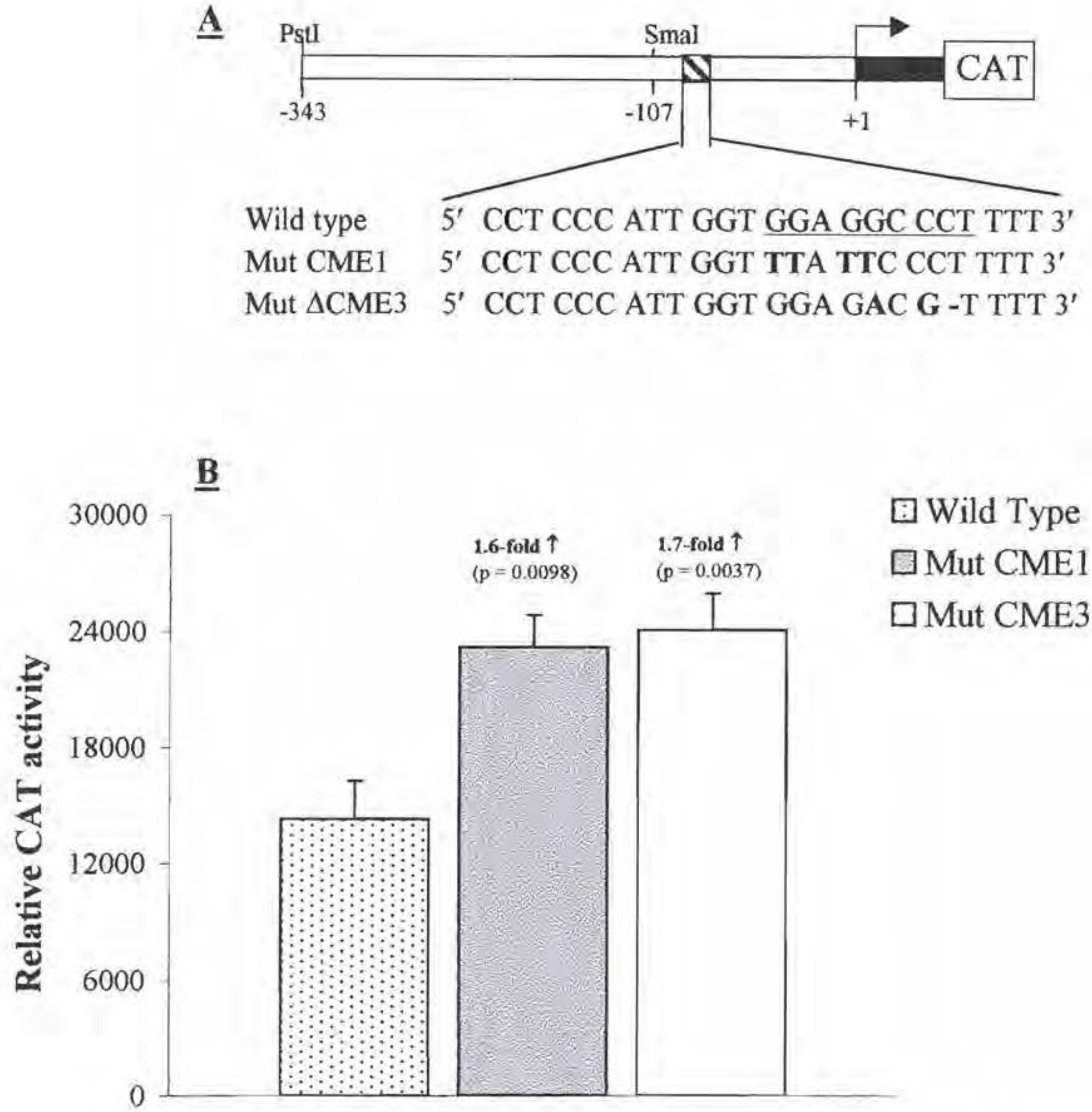


Figure 2.15: Transcriptional activity of the mutant CME -343 COLCAT constructs in CT-1 fibroblasts. CT-1 fibroblasts were transfected with the -343 COLCAT constructs using the calcium phosphate precipitation method. Cells were harvested and CAT and β-galactosidase activities measured as described in section 6.2, CAT activity is expressed relative to β-galactosidase activity. **(A)** A schematic representation of the -343 COLCAT constructs showing the CME region (underlined); point mutations are indicated in bold letters. **(B)** Relative CAT activity of CT-1 fibroblasts. Results represent the mean ± SD of five separate experiments performed in triplicate; p values were determined by the student *t* test.

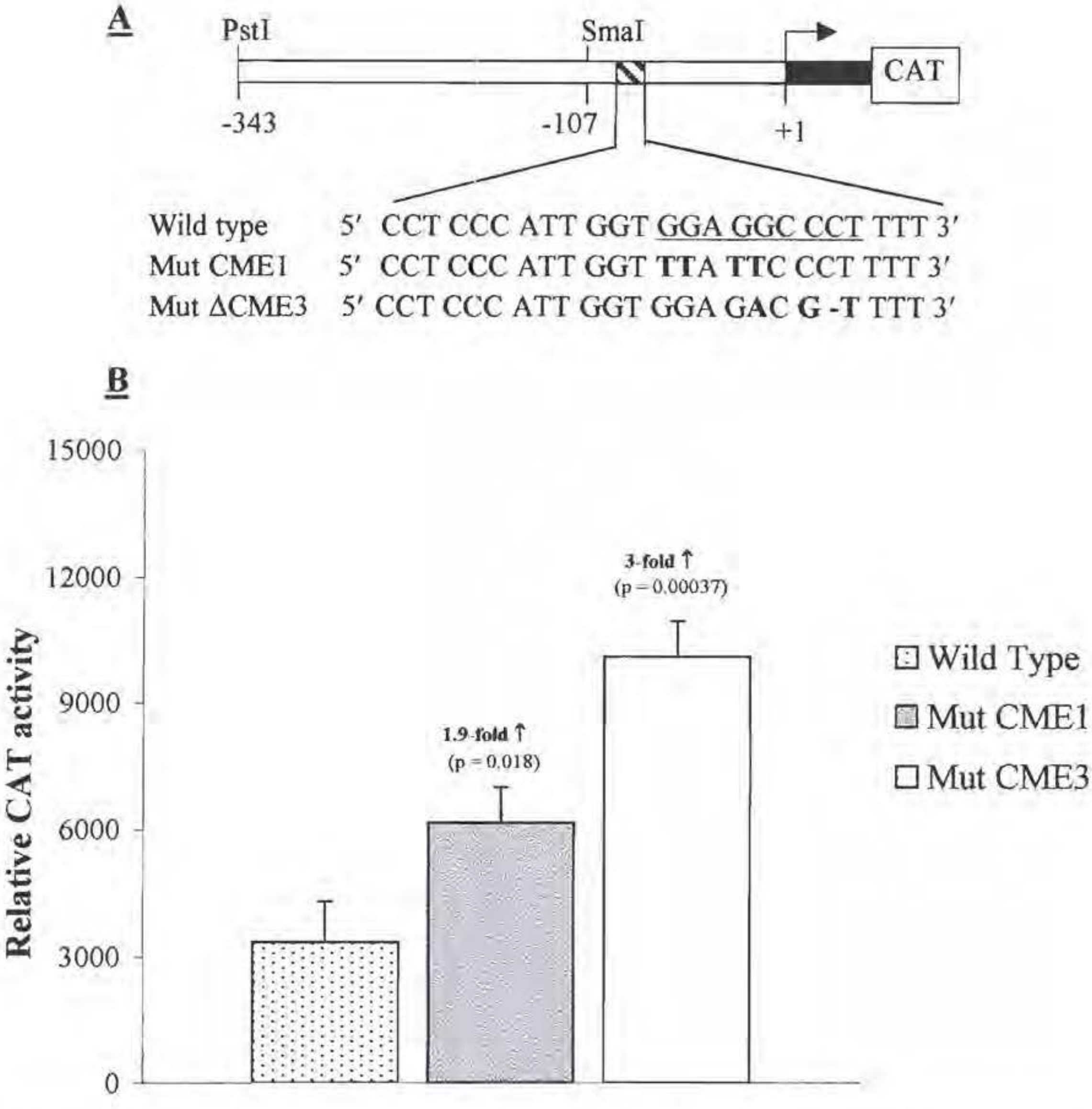


Figure 2.16: Transcriptional activity of the mutant -343 COLCAT construct in SVWI-38 fibroblasts. **(A)** A scheme showing the -343 $\alpha2(I)$ procollagen promoter-CAT construct; the CME is underlined and point mutations are indicated in bold letters. **(B)** The relative CAT activity of SVWI-38 fibroblasts transfected with the -343 COLCAT constructs containing mutations in the CME region. CME; collagen modulating element. Results represent the mean ± SD of five separate experiments done in triplicate and p values were calculated using the student *t* test.

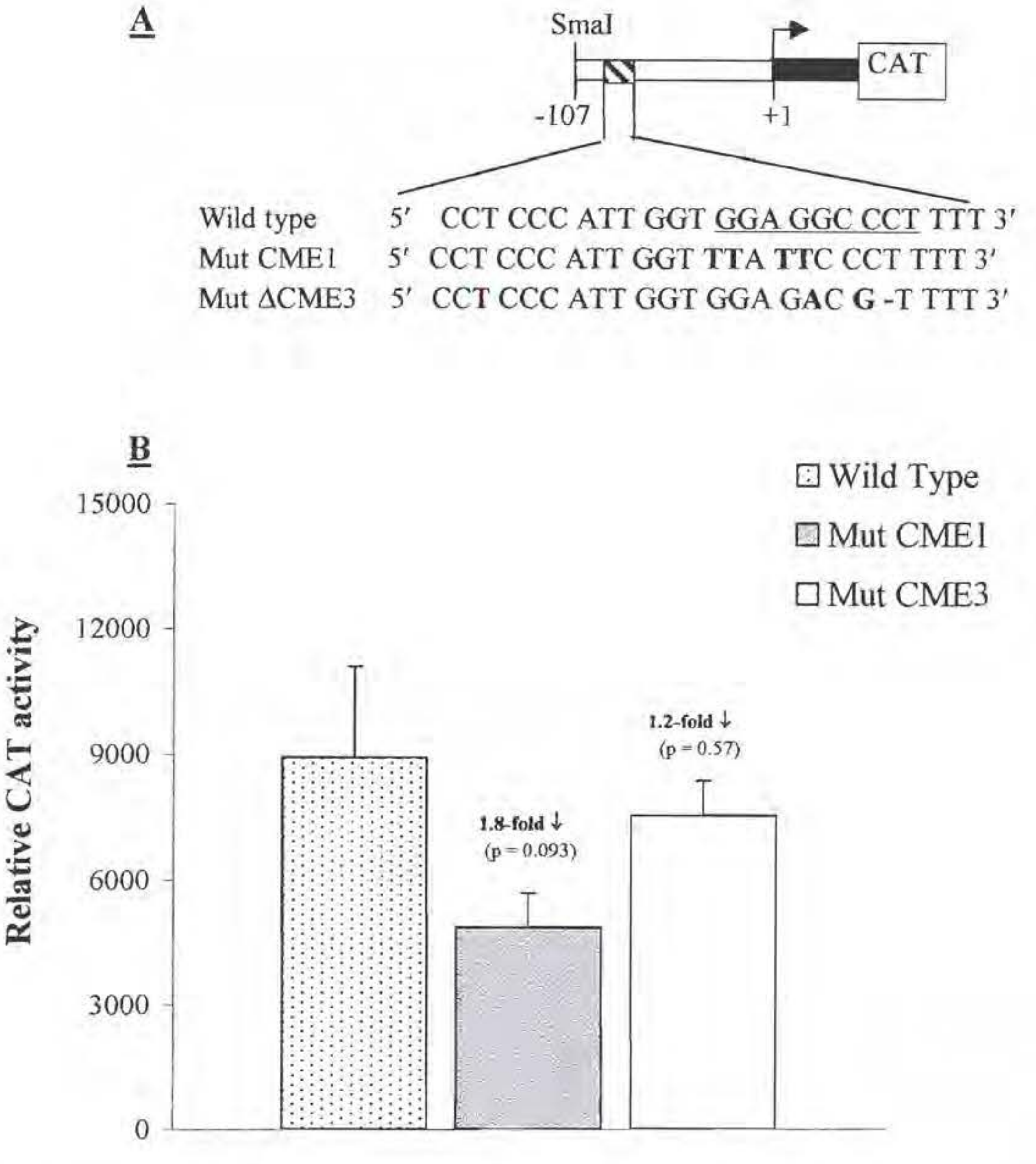


Figure 2.17: Transcriptional activity of the mutant CME -107 COLCAT constructs in CT-1 fibroblasts. (A) A schematic representation of the -107 COLCAT constructs; the CME region is underlined and point mutations are indicated in bold letters. **(B)** The relative CAT activity of CT-1 fibroblasts transfected with -107 COLCAT constructs with mutations in the CME region. Results represent the mean ± SD of five separate experiments done in triplicate; p values were determined by the student *t* test.

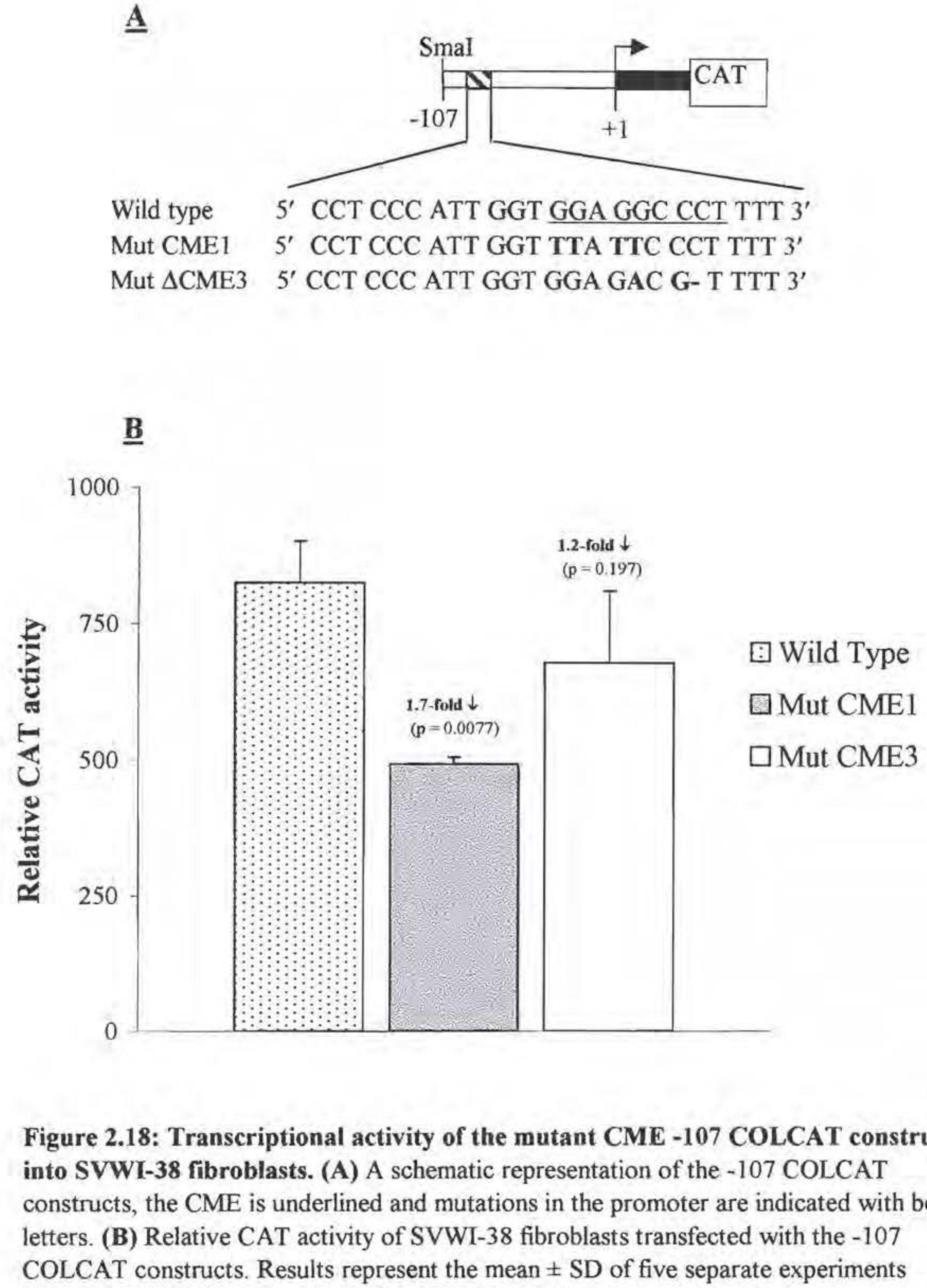


Figure 2.18: Transcriptional activity of the mutant CME -107 COLCAT construct into SVWI-38 fibroblasts. (A) A schematic representation of the -107 COLCAT constructs, the CME is underlined and mutations in the promoter are indicated with bold letters. **(B)** Relative CAT activity of SVWI-38 fibroblasts transfected with the -107 COLCAT constructs. Results represent the mean ± SD of five separate experiments performed in triplicate; p values were calculated using the student *t* test.

described in section 2.2.4, but in this case, the mutations were in the CME region. The wild type sequence resulted in competition of both complexes I and II, while the mutated oligonucleotides had lost the ability to compete complex II binding (Figure 2.19). The increase in promoter activity can therefore be correlated with failure of complex II proteins to interact with the mutated promoter constructs.

2.2.7. Comparison of the transcriptional activity of the human and mouse proximal α2(I) procollagen promoter.

Previous studies have demonstrated the presence of a three base pair mismatch between the CME regions of the mouse and human α2(I) procollagen promoters (59). This mismatch has been shown to abolish the formation of complex II when the mouse promoter is used as a probe in EMSA reactions in the presence of nuclear proteins from either mouse or human fibroblasts. However, complex II formation is observed with nuclear extracts from both human and mouse cells when the human promoter is used as probe (59). From section 2.2.5, it is evident that mutations of the CME increased promoter activity, especially in SVWI-38 fibroblasts. It was therefore decided to compare the transcriptional activity of the human and mouse α2(I) procollagen promoters, bearing in mind that the three base pair mismatch in the CME region is not the only region of mismatch in the proximal promoter of these genes. CT-1 and SVWI-38 fibroblasts were transfected with 20 μg of the -343 COLCAT human or mouse constructs and CAT activity determined as described in section 6.2.6. It is evident from figure 2.20 that the human promoter activity was 6-fold higher than the mouse promoter in CT-1 fibroblasts and a 9-fold difference was observed in SVWI-38 fibroblasts. The significance of these findings will be discussed in section 2.3.

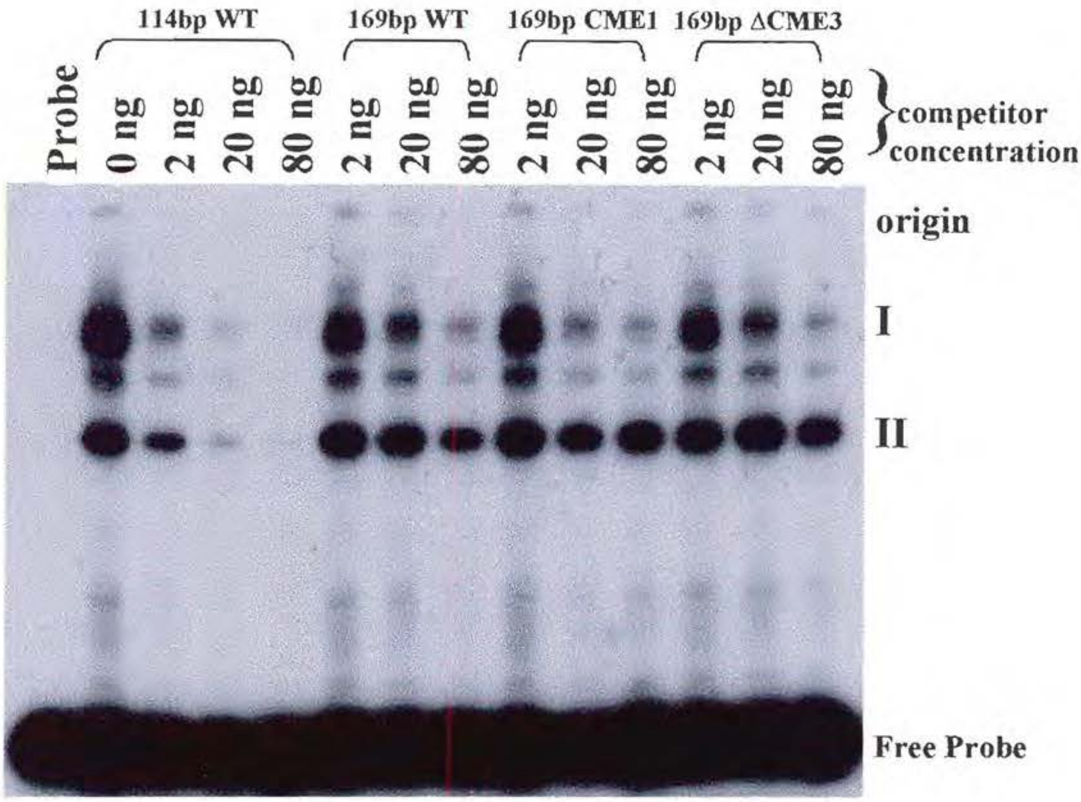


Figure 2.19: **Competition of DNA-protein complex formation by mutated CME oligonucleotides.** Nuclear proteins from CT-1 fibroblasts were incubated with a ^{32}P-labelled -107 to -60 α2(I) procollagen promoter fragment (114 bp) and DNA-protein complexes were resolved by electrophoresis on a non-denaturing 5 % polyacrylamide gel. The -107 COLCAT promoter constructs of mut CME1 and ΔCME3 were digested with SmaI and HindIII to release a 169 bp fragment, DNA was eluted as described in section 6.4.2.3 and used as competitors in the EMSA reaction. WT; wild type, CME, collagen modulating element (see Table 2.2 and Figure 2.7 for the sequence of the mutations)

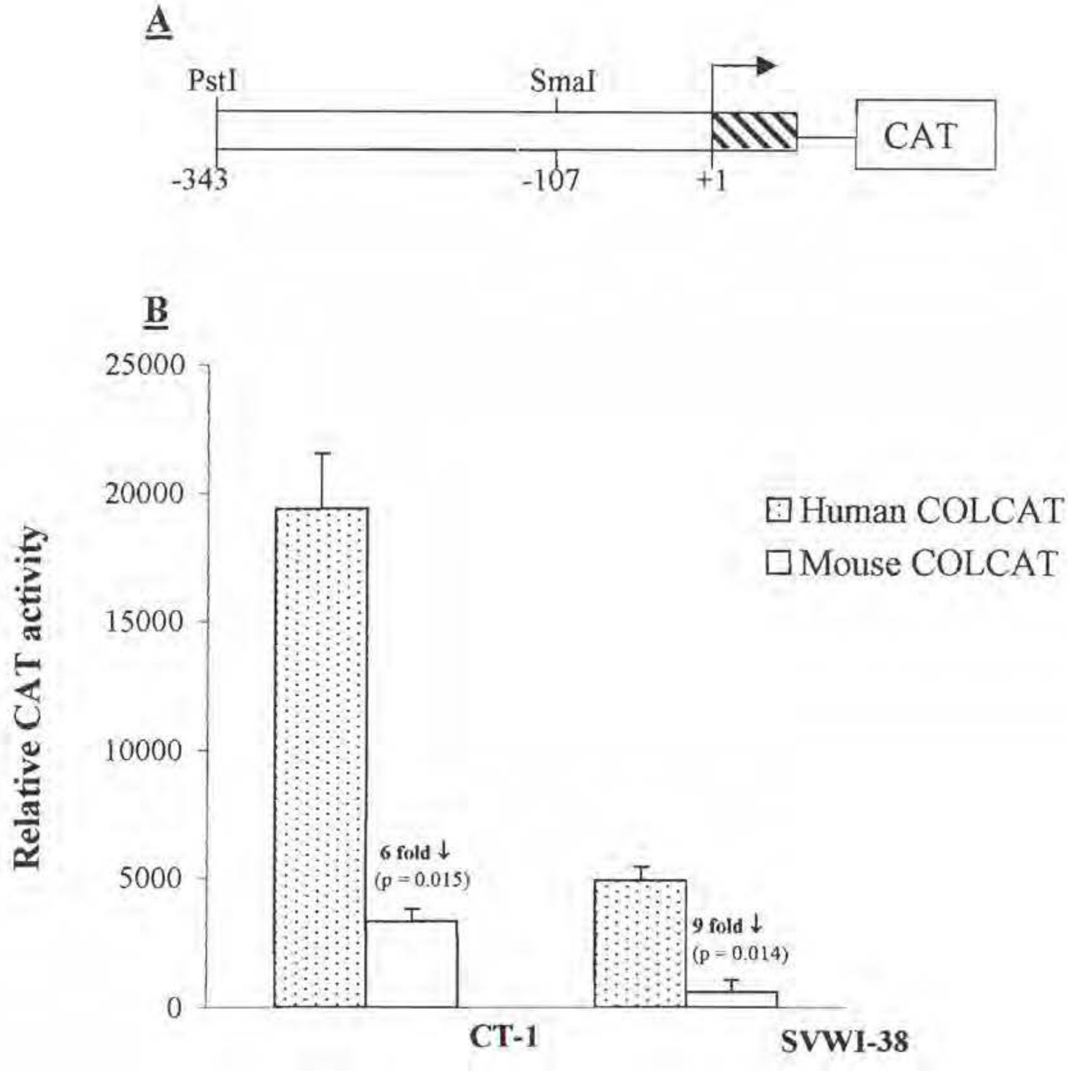


Figure 2.20: Comparison of the transcriptional activity of the human and mouse -343 $\alpha 2(I)$ procollagen promoters in CT-1 and SVWI-38 fibroblasts. (A) A schematic representation of the -343 $\alpha 2(I)$ procollagen promoter-CAT constructs. Cells were grown to 40 % confluency and transfected with 20 μg of the -343 human or mouse $\alpha 2(I)$ procollagen promoter-CAT constructs using the calcium phosphate precipitation method as described in section 6.2. Cells were harvested and **(B)** the relative CAT activity of -343 COLCAT constructs transfected into CT-1 and SVWI-38 fibroblasts were determined as described in section 6.2. The above results represent the mean ± SD of three separate experiments performed in triplicate; the student *t* test was used to calculate p values.

2.3 DISCUSSION

Type I collagen is synthesised by a limited number of cell types in intact animals and embryos, *inter alia* osteoblasts, odontoblasts, fibroblasts, mesenchymal cells and smooth muscle cells (54). A number of reports have described different elements both upstream and downstream of the transcription initiation site of the $\alpha1(I)$ and $\alpha2(I)$ procollagen genes that modulate their expression. Procollagen genes which contain transcriptional regulatory regions in the first intron include: $\alpha1(I)$, $\alpha2(I)$, $\alpha1(II)$ and $\alpha1(IV)$ (62). The proximal promoter region harbours a number of transcription factor binding sites of which the CCAAT binding protein (CBF), also referred to as NF-Y and CP1, is the best characterised, although most of the work was carried out with the mouse $\alpha2(I)$ procollagen promoter (53). CBF recognises CCAAT motifs in most promoters of mammalian genes transcribed by RNA polymerase II (53).

In this study, the regulation of the human $\alpha2(I)$ procollagen gene by the CCAAT-binding protein, referred to as CBF, was investigated. Furthermore, this study sought to characterise the role of this protein in procollagen and non-procollagen expressing fibroblasts. Mutations were introduced in the promoter in the previously described G/CBE element (13,51) and the effects of these mutations analysed by transient transfection assays using a CAT reporter gene. From the results section it is clear that mutations in this element drastically decreased promoter activity regardless of the ability of the cells to express $\alpha2(I)$ procollagen. This feature of CBF to regulate transcription of most genes can be explained by the observation by Coustry *et al* (187) that CBF complexes with dTAF110, a TATA box binding protein-associated factor. TAFs are components of the general transcription machinery and they play important roles in transcription activation (188). This suggests that CBF is required for the formation of the pre-initiation complex by allowing formation of further protein-protein interactions while anchored via TAF110 to the TATA binding protein (TBP) with a consequent increase in transcriptional activity (187). The protein-protein interaction between the activator and TAF might stabilise the binding of

TFIID to the TATA box, or it might induce a conformational change in the TFIID complex that promotes the assembly of the initiation factors into an active complex. To support the suggestion that CBF might be involved in protein-protein interactions with the basal transcription factors, transfections carried out with the basal promoter element between -107 and +54. The mutated G/CBE element had significantly reduced promoter activity in both CT-1 and SVWI-38 cell lines. These results suggest that transcriptional activation of the α2(I) procollagen gene by CBF can take place in the absence of upstream elements.

Previous analysis of the CBF binding sites have indicated that mutations of each nucleotide of the CCAAT sequence in the promoter of the mouse α2(I) procollagen gene decreases promoter activity and abolishes CBF binding (53). A similar observation was made by Ihn *et al* (47) whereby mutations in the CCAAT box not only abolished CBF binding but also resulted in a significant reduction in the activity of the proximal promoter of the human α2(I) procollagen gene. In a related study, Chen *et al* (189) demonstrated that mutation of the two CCAAT boxes present in the human cdc2 promoter interfered with basal promoter activity and the promoter failed to be transactivated by the SV40 large T antigen. In addition 5′-deletion analysis and substitution mutagenesis of the fatty acid synthase promoter revealed that an essential element required for cAMP antagonism of the insulin effect was an inverted CCAAT box located between nucleotides -99 and -92 in the fatty acid synthase gene (190). An extensive review on the role of the CCAAT binding protein in transcriptional regulation is presented in a report by Maity and de Crombrugghe (53). Interestingly not all promoter elements containing the CCAAT sequence bind CBF, suggesting that specific nucleotide sequences other than CCAAT are necessary for CBF binding (53). High affinity CBF binding sites have been shown to contain [T/C][A/G][A/G] as the 5′ flanking sequence and CA as the 3′ flanking sequence of CCAAT (53). Our studies have shown that whilst the integrity of the 5′ flanking sequences are essential for CBF binding in the α2(I) procollagen gene, the 3′ flanking sequences are not required. Mutations in the upstream inverted GGAGG box of the human α2(I) procollagen promoter effectively reduced the activity of the promoter to the same extent as mutations in the CBF binding sequence. In order to determine whether the effect of mutations in the G/CBE element *in vivo* can be correlated with *in vitro* DNA-protein complex formation, complex formation was assayed by EMSA. It is evident that the decrease in the expression of the CAT gene in these mutated constructs can be ascribed to decreased DNA binding. Thus, it can be concluded that the CCAAT box and the adjacent 5′ flanking inverted GGAGG sequences are important in the regulation of the expression of the $\alpha2(I)$ procollagen gene and this region of the promoter binds a factor with a positive effect on transcription of the gene. The significance of these results can be extended to the finding that the regulation of the expression of the $\alpha2(I)$ procollagen gene by the G/CBE binding proteins takes place regardless of the ability of a particular cell line to express the gene.

Comparison of the mouse and human $\alpha2(I)$ procollagen promoters revealed that there is about 86 % sequence similarity between nucleotides -324 and -224 with homology decreasing in both the 5′ and 3′ directions. For example, NF1 binds to the mouse sequence between -290 and -304, but the equivalent region in the human gene does not contain this element (46). The mouse $\alpha2(I)$ procollagen promoter region between nucleotides -180 and -136 has been shown to bind BFCOL (48). Although mutations in the binding site of this protein results in an increase in promoter activity, the authors argued that BFCOL is not a repressor but that these mutations allow binding of more potent transcriptional activators with a net increase in promoter activity. In contrast, the human $\alpha2(I)$ procollagen promoter has been shown to possess a repressor binding site between -173 and -155 (47). Based on the sequence homology between the human and mouse $\alpha2(I)$ procollagen promoter in this region, it is very tempting to suggest that the protein(s) that binds to this region is similar to BFCOL. A small 3-base difference in the sequence between the $\alpha2(I)$ procollagen promoters of mouse and human occur in the region between -80 and -67 (13,51). This region has been shown previously to harbour a potential repressor binding site in the human $\alpha2(I)$ procollagen promoter (59). To further characterise this segment of the promoter, functional studies using site-directed mutagenesis together with transient transfection assays with a CAT reporter gene were performed. Whilst the activity of the promoter was significantly increased in the two cell lines in the presence of the mutations in the CME element, the increase was higher in the SVWI-38 fibroblasts. Previous transfection results with a wild type promoter CAT construct have demonstrated that the activity of the promoter is markedly decreased in SVWI-38 compared to that of CT-1 and WI-38 fibroblasts (13). This observation suggested that there is an active repressor in SVWI-38 fibroblasts that could possibly account for this reduction in promoter activity. Interestingly, there were no differences in the DNA binding profiles for nuclear proteins extracted from the latter two cell lines as was previously reported (13,51,59). Mutations that increase promoter activity compared to the wild type promoter can be accounted for by the possibility that: (a) the protein that binds the mutant sequence is a *bona fide* transcriptional repressor, or (b) the DNA binding protein is a weak activator and the increase in promoter activity of the mutant promoter is due to binding of a stronger activator to the mutant promoter (54). Although mutations in the BFCOL binding site yielded the same phenomenon and the authors suggested that this site could overlap binding sites for potent activators such that in the presence of mutations, these activators bind more efficiently to the mutated region as compared to BFCOL (48). If the CME-binding protein was a less potent activator, inhibition of its binding by the presence of mutations in its binding site, accompanied by improved binding of CBF should increase promoter activity significantly in both cell lines, or more so in CT-1 fibroblasts. The CCTCC motif in G/CBE and the GGAGG motif in the CME (Figure 2.21) could possibly engage in the formation of a secondary structure, for instance dyad symmetry (51); mutations in the CME could alter this conformation and therefore the promoter will acquire a much more relaxed conformation which will then allow improved interactions at the G/CBE element with a consequent increase in promoter activity. The secondary structure can be disrupted by introducing a spacer sequence between the G/CBE and CME binding motifs that would allow analysis of independent contribution of these motifs in activation of transcription. In this study, however, the increase in the activity of the promoter construct containing mutations in the CME was greater in SVWI-38 fibroblasts compared to CT-1 fibroblasts. The activity of the promoter construct with CME mutations transfected into SVWI-38 fibroblasts increased to levels similar to the activity of the wild type promoter construct transfected into CT-1 fibroblasts. Therefore, the possibility that mutations in the CME resulted in increased binding of potent activators to the mutated sequences does not appear to be the case in this study. The more favourable argument is that the CME binding protein acts as a repressor. It should be emphasised that the increase in promoter activity was only evident with the proximal promoter construct suggesting a possible involvement of upstream elements. Transfection of the -107 promoter construct did not yield the same increase, instead the activity of the promoter was lower than for the wild type construct. Results obtained with the -107 promoter constructs that contained CME mutations are contrary to results obtained when cells were transfected with the -107 promoter constructs that contained mutations in the G/CBE; it would appear that the G/BE binding protein is required in the activation of the α2(I) procollagen gene via interaction with the general transcription factors whilst the CME binding protein requires upstream elements/factors to regulate gene expression.

5′-**CCT C**CA TTG GT**G GAG G**CC CTT TT-3′
3′-**GGA GG**T AAC CAC **CTC C**GG GAA AA-5′

Figure 2.21: Schematic representation of the G/CBE-CME region in the promoter of the human α2(I) procollagen gene. Regions represented in bold could engage in formation of a secondary structure.

Ihn and Trojanowska (40) have demonstrated that although Sp1/Sp3 are transcriptional activators of the human α2(I) procollagen gene, they also bind to a repressor site between -164 and -159 region, which is characterised by a TCCCCC motif. It is therefore possible that mutations in the basal -107 promoter construct do not yield the desired effect as evidenced with the -343 construct since the basal promoter construct lacks regulatory sequences such as the upstream TCCCCC motif which might be recognised by proteins that act in synergy with the CME binding function as being important in determining the mechanism of repression rather than the repressor complex only. Recently identified silencers with a motif homologous to the CME were reported in the human platelet derived growth factor A-chain (+1605 to +1630: CG**GGGAGG**GGGAGTGG**GGGACG**CC and -1418 to -1388: **GGGGAGG**GGG), and the rat cytochrome P-450 CYP1A1 (-873 to -851: **CCTCC**ACTCT) genes, however, the identity of the factor(s) that recognise these elements, however, are still unknown (188).

It has been shown that the *in vitro* activity of the mouse α2(I) procollagen enhancer is not limited to fibroblasts, although it may function only in cell types that are capable of synthesising type I procollagen (62). Some of the cell lines in which the enhancer is active include NIH/3T3 and 3T3-L1 mouse fibroblasts, HeLa human cervical carcinoma cells, and CV-1 African Green Monkey kidney cells (62). In addition a number of non-fibroblastic cell types such as kidney epithelial cells, smooth muscle cells and macrophages can synthesise type I collagen (191-193). In view of this, it is interesting to note that the CME binding proteins are ubiquitous, they have been shown to be present in a wide range of human cell lines, namely, U937 monocytic cells, HT1080 fibrosarcoma cells, HepG2 hepatocellular carcinoma cells, L-77 lymphoblasts, MDA-MB-231 and ZR-72-2 breast cancer epithelial cells, normal and transformed fibroblasts (59). In addition, these proteins have been shown to be present in extracts from rodent cell lines such as p388D_1 mouse macrophage cells, 3T3 and 3T6 rodent fibroblasts (59). Although transfection studies showed that the CME binding proteins are involved in the regulation of the α2(I) procollagen gene expression, they do not, however, explain the ubiquity of these proteins.

In a study by Chen *et al* (52), it was demonstrated that comparison of the *in vitro* DNase 1 footprints for each of the procollagen promoters from cell lines which either expressed or did not express type I procollagen revealed a similar pattern of protection, suggesting that the different binding proteins were present in all cell types studied. *In vivo* footprints over the same promoter segments that produced *in vitro* footprints were, however, cell specific, occurring only in fibroblasts and not in other cell types that did not express type I procollagen (52). These results suggested that although all cell types tested contained nuclear proteins that can bind to the proximal promoter of the type I procollagen genes *in vitro*, it is only in fibroblasts that these proteins actually bind to their cognate sites *in vivo*. Liver nuclear extracts from Phenobarbital treated and untreated rats produced the same DNA-protein interactions at the Phenobarbital response element identified in the rat cytochrome2B1/2 and mouse cytochrome2b10 genes (194). There were no differences in the protected region in *in vitro* footprints, but analyses of the native chromatin *in vivo* obtained from treated and untreated rats, showed dramatic differences in the protection pattern. In the major histocompatibility class II genes, binding sites for ubiquitous proteins have also been shown to be occupied only in cells that express the genes (195). These reports suggest that although a specific factor is ubiquitous, it may interact only with its specific binding site in cells where it has function.

To further characterise the role of CME binding proteins, CT-1 and SVWI-38 fibroblasts were transfected with human and mouse wild type proximal promoter constructs. The rationale being that, if mutations in the CME region increase promoter activity, then transfection of the mouse promoter should generate the same effect since the mouse promoter contain a three base pair mismatch in the CME region. Interestingly the mouse promoter was less active in human cells. These results further support the notion that the CME region alone is not sufficient for repression of the $\alpha2(I)$ procollagen gene. Alternatively, these observations could simply reflect differences in regulatory elements present in the human and mouse promoters, where different DNA-protein complexes are formed resulting in variations in the activities of these promoters.

To summarise, this chapter demonstrated that only a single factor, complex II, associates with the CME region of the $\alpha2(I)$ procollagen promoter in *in vitro* DNA binding assays whereas the previously characterised complex III is a specific proteolytic fragment of complex II. Although complex II proteins are present in $\alpha2(I)$ collagen producing and non-producing cells, namely CT-1 and SVWI-38 fibroblasts, binding assays whereas the previously characterised complex III is a specific proteolytic fragment of complex II. Although complex II proteins are present in α2(I) collagen producing and non-producing cells, namely CT-1 and SVWI-38 fibroblasts, respectively, it appears that the regulation of expression of the α2(I) procollagen gene by the CME is different in the two cell lines. A significant increase in promoter activity was observed in both CT-1 and SVWI-38 fibroblasts following transfection with promoter constructs that contained mutations in the CME but the increase was greater for SVWI-38 fibroblasts; thus leading to the conclusion that the CME binding protein acts as repressor of transcription of the α2(I) procollagen gene, and perhaps this effect is cell-specific as suggested by Leaner (59). The existence of a transcription factor that represses α2(I) procollagen gene expression in fibroblasts suggests a high degree of complexity in the regulation of this gene, and obviously this regulation will impact on synthesis of overall type I collagen. Further studies will seek to understand why fibroblasts need an inhibitor of transcription of procollagen genes and perhaps also to establish the role of the repressor during development, hormonal regulation as well as transformation and metastasis. It would be interesting to investigate the role of the specific proteolytic product (complex III) in the regulation of expression of the α2(I) procollagen gene. The presence of an intact DNA-binding proteolytic fragment could be an indication of the presence of a functionally autonomous DNA-binding domain. Although the proteolytic fragment might have no bearing on regulation of the gene, it might be an essential tool for mapping amino acid residues engaged in direct contact with DNA. The G/CBE binding protein, on the other hand, is required for gene activation in both cell lines. Members of the CCAAT binding factor family of proteins have been shown in a number of studies to activate transcription of a variety of genes, thus it is not surprising that the G/CBE binding protein also activates transcription of the human α2(I) procollagen gene in transformed cells. In addition, these results agree with observations that the CCAAT element is involved in the activation of the mouse α2(I) procollagen gene. It can be concluded that the G/CBE binding protein activates the expression of the α2(I) procollagen gene whilst the CME binding protein is involved in interactions that result in repression of the gene.

CHAPTER 3

PURIFICATION AND CHARACTERISATION OF THE PROCOLLAGEN MODULATING ELEMENT (CME) BINDING PROTEIN(S)

3.1 INTRODUCTION

A number of studies have reported on the presence of *cis*-regulatory sequences in the promoter and first intron of the human α2(I) procollagen gene (see previous chapters). Most of these studies focused on the identification and characterisation of the regulatory sequences by demonstrating their requirement in basal or induced promoter activity. An understanding of the molecular mechanisms by which they control gene expression requires the identification and characterisation of these transcription factors. Progress in this field, however, has been retarded mainly by the fact that transcription factors are present in extremely low quantities in cells and some of these factors might require induction and/or have high turnover rates.

Different approaches have been employed to purify transcription factors. Two approaches that have generally enjoyed success are (a) molecular cloning of transcription factors by screening cDNA expression libraries and subsequent expression of the protein in bacteria or yeast cells, and (b) conventional biochemical purification from tissues or cultured cells. The main advantage of the first method is that in addition to the isolation of the cDNA encoding the transcription factor, the method requires only the use of the consensus binding sequence to "pull-out" the correct clone. This approach, however, has been shown to generate many false positives. Some of the transcription factors isolated using this approach include oct-1/2, CREB and BFCOL (48,196-198). A number of researchers have reported the successful use of conventional biochemistry in the purification of transcription factors, such as OTF-1/2, TFIIA, CBF, Sp1 and BBF (199-205). The purified transcription factor was then used to raise antibodies or to generate a partial amino acid sequence of the protein. Recently, transcription factors have been analysed by the use of **m**atrix-**a**ssisted **l**aser **d**esorption/ionisation at the **t**ime of **f**light **m**ass **s**pectrophotometer (MALDITOF-ms), as will be described later in this chapter.

This chapter will describe biochemical purification of complex II proteins by ion-exchange chromatography, sequence-specific DNA affinity chromatography, sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and attempts to identify and characterise the complex II proteins.

3.2 RESULTS

3.2.1 Purification of the CME binding protein

As mentioned in Chapter 2, complex II binding proteins are fairly ubiquitous, in that they are expressed in many different cell lines, including cells such as haemopoietic cells, epithelial cells and fibroblasts (Figure 3.1). The objective of the previous chapter was to demonstrate that complex II proteins play a functional role in the regulation of $\alpha2(I)$ procollagen gene expression. For the purpose of purification, the cell line of choice was the U937 monocytic cell line because they grow rapidly in suspension cultures, thus making it cheaper and faster to work with these cells than with fibroblasts. Cells were grown in RPMI-1640 supplemented with 10 % foetal calf serum and antibiotics in a humidified atmosphere in $150cm^2$ tissue culture flasks as described in section 6.1.2. Confluent cells were harvested by centrifugation and nuclear proteins were prepared using a modified version of the Dignam protocol (Section 6.6.1). Nuclear proteins were assayed for DNA binding activity by EMSA (see section 6.6.2) prior to purification and the activity was again monitored at each purification step.

3.2.1.1 Ion-exchange chromatography

Complex I and II DNA-binding activities can be separated from each other by ion-exchange chromatography on a heparin-agarose matrix (57). Elution of these proteins was effected with a stepwise KCl gradient, complex I and II proteins eluted at 0.2M and 0.4M KCl, respectively. Approximately 40 mg of the crude nuclear extract was loaded onto a heparin-agarose ion-exchange column (15 cm x 0.5 cm), and eluted with increasing concentrations of KCl. Figure 3.2 illustrates a typical elution profile of transcription factors that bind to the -107 to -60 $\alpha2(I)$ procollagen promoter; each fraction was assayed for DNA binding activity by the electrophoretic mobility shift assay (EMSA described in section 6.6.2). The flowthrough fraction occasionally contained complex I and II proteins depending on the starting protein concentration of the nuclear extracts. It is evident that chromatography by ion-exchange is an important step in the purification of complex II proteins since its activity could be

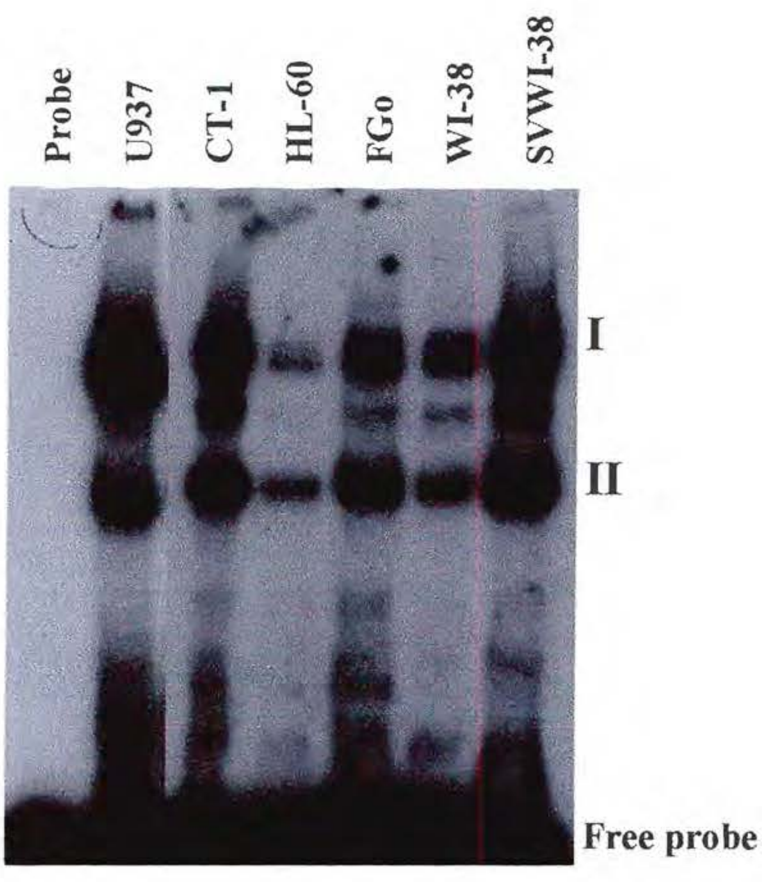


Figure 3.1: Electrophoretic mobility shift assay of nuclear proteins extracted from different cell lines. Nuclear proteins were extracted from different cell lines following a modified method of Dignam *et al* (245) as described in section 6.6.1. Crude nuclear protein extract (4μg) was incubated with 10^4 cpm of a ^{32}P-dCTP labelled -107 to -60 $\alpha 2$(I) procollagen promoter fragment;2μg of polydI.dC was included in the reaction to minimise non-specific binding. DNA-protein complexes were resolved by electrophoresis on a non-denaturing 5% polyacrylamide gel, the gel was dried and exposed to X-ray film for 16 hours at -70 °C. **I**; complex I, **II**; complex II.

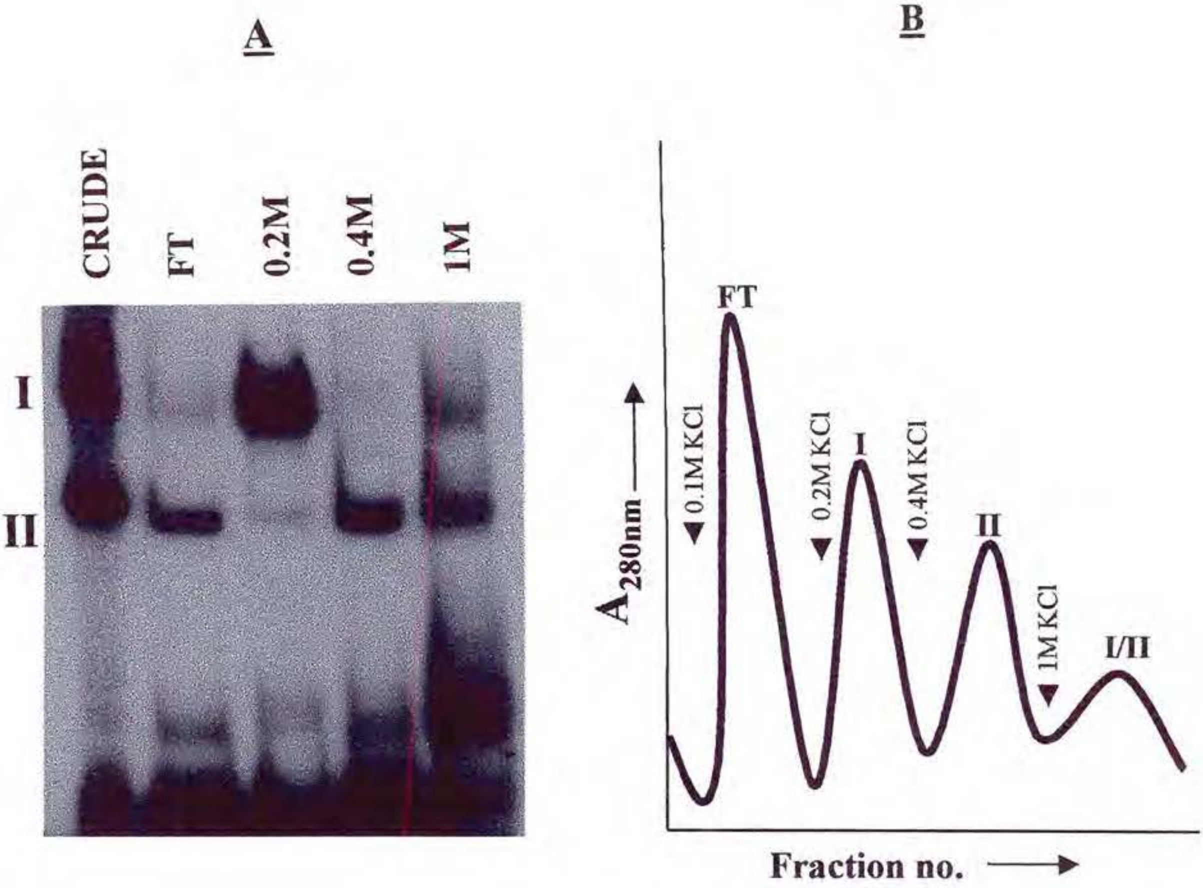


Figure 3.2: Heparin-agarose ion-exchange chromatography of U937 nuclear proteins. Nuclear proteins were extracted from U937 monocytic cells as described in section 6.6.1 following a modified method of Dignam *et al* (245). 40 mg of the crude nuclear extract was loaded onto a 15 cm x 0.5 cm heparin agarose column. Proteins were eluted with CB buffer containing 0.1M, 0.2M, 0.4M or 1M KCl (see section 6.7). Each fraction was assayed for DNA-binding activity and DNA-protein complexes were resolved by electrophoresis on a non-denaturing 5% polyacrylamide gel. **(A)** Electrophoretic mobility shift assay (EMSA) of heparin agarose fractions; **(B)** A typical elution profile of nuclear proteins from the heparin agarose column monitored with a UV detector at λ_{280nm}. Each peak represents a specific fraction eluted at thenindicated [KCl] in the buffer. **FT**~ flowthrough fraction, **I**~ complex I and **II**~ complex II proteins.

separated from that of complex I. These fractions, however, are not homogeneous since many non-DNA binding proteins are still present.

3.2.1.2 DNA affinity chromatography

The next step in the purification process involved preparation of a sequence specific DNA affinity matrix (section 6.6.3.2.1). This was accomplished by following the procedure of Kerrigan and Kadonaga (206). Briefly, the method entails annealing the complementary oligonucleotides recognised by the protein of interest, phosphorylation and ligation of the oligonucleotides to generate concatamers (see Figure 3.3). Ideally, oligonucleotide multimers of 10-mers or larger should be coupled to the matrix to increase specific binding. Whilst smaller multimers of less than 10-mers can be used, they tend to give very poor yields of the protein factor of interest due to steric hindrance that arises when a short oligonucleotide is used against the Sepharose support. Multimers of complementary CME oligonucleotides were prepared and analysed by electrophoresis on a 1.5 % (w/v) agarose gel (Figure 3.4). This oligonucleotide mixture was coupled to cyanogen-bromide activated-Sepharose 4B. A 5ml column (7 cm x 0.5 cm) was then packed with the matrix and equilibrated with a 1x DNA binding assay buffer (see section 6.7). The 0.4M KCl fractions from heparin-agarose were pooled and pre-incubated for 30 minutes on ice with polydI.dC to reduce non-specific binding to the matrix, diluted to about 0.1M KCl and loaded onto the DNA affinity column. Proteins were eluted with the EMSA binding buffer in the presence of increasing concentrations of KCl. A typical elution profile from the affinity matrix is illustrated in figure 3.5. The highest complex II DNA-binding activity was obtained in fractions eluting between 0.5M and 0.7M KCl, hence in subsequent purifications, elution of complex II proteins was effected in 0.7M KCl (Figure 3.6). Although the affinity matrix bound sufficient amounts of complex II proteins, it appeared that it reached saturation very rapidly, with at least 50 % of the remaining proteins present in the flowthrough fraction. In addition to DNA binding assays, purification was monitored by SDS-PAGE followed by silver staining of the gels (see section 6.6.5). A peculiar observation from these SDS gels (Figure 3.7) was the fact that the presence of β-mercaptoethanol (β-MSH) or dithiothreitol (DTT) in

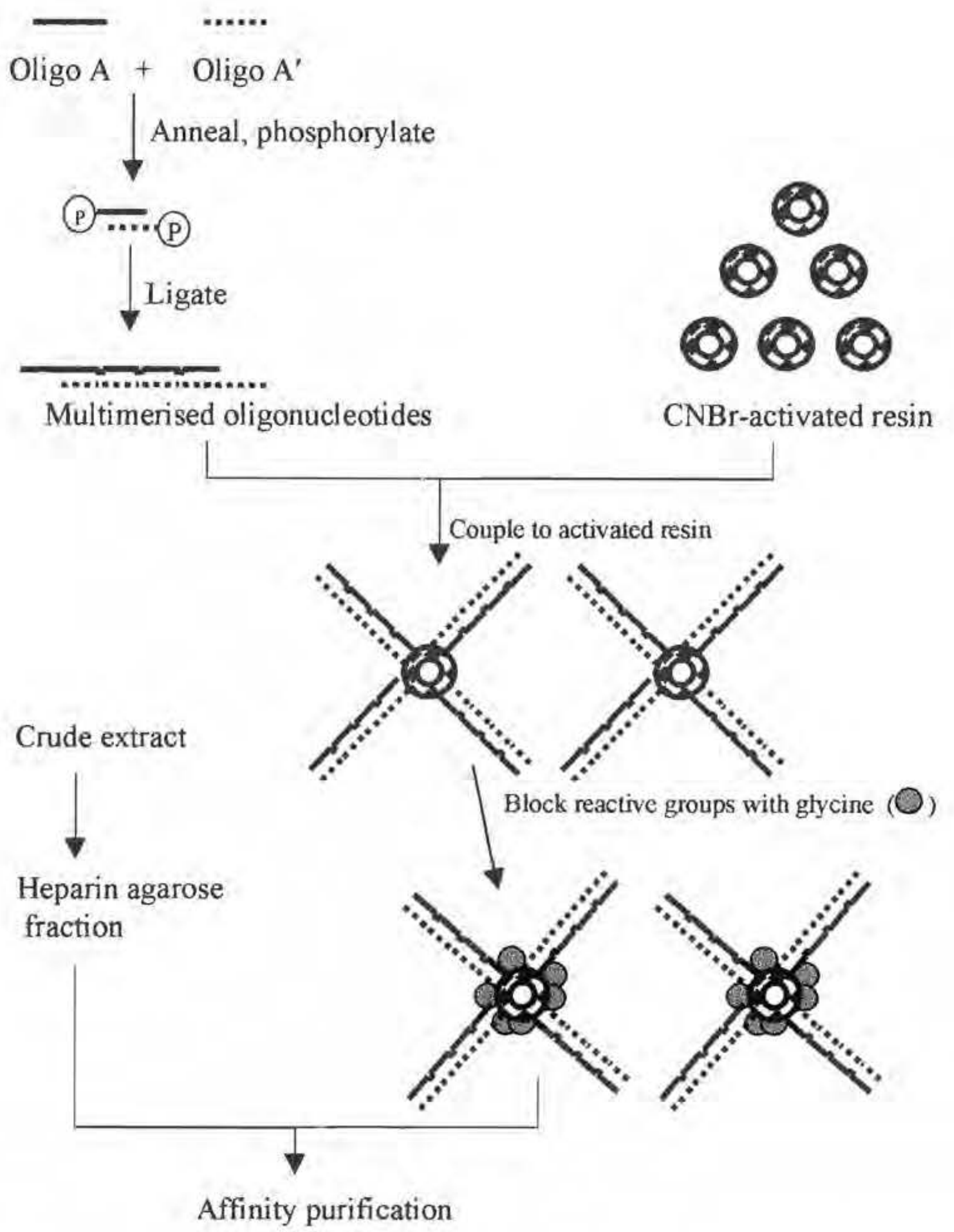


Figure 3.3: **Affinity purification of DNA-binding proteins.** The scheme was adapted from the method of Kerrigan and Kadonaga (206). Complementary CME oligonucleotides (A and A′) were annealed, phosphorylated and ligated to generate multimers of no less than 10-mers. The multimeric oligonucleotides were coupled to cyanogen bromide(CNBr) activated Sepharose 4B and reactive groups were blocked by successive washes with 0.2M glycine (section 6.6.3.2.1). The matrix was then packed into suitable columns and used to purify proteins pooled from the heparin agarose purification step. See a detailed description of the method in section 6.6.3.2.

A; 5′ A ATT CGG AGG CCC TTT TGG AGG GGA GGC CCT TTT GGA GG 3′,
A′; 5′ A ATT CCT CCA AAA GGG CCT CCC CTC CAA AAG GGC CTC CG 3′;
The oligonucleotides were designed such that they had 5′ AATT 3′ sticky ends for ease in the ligation step.

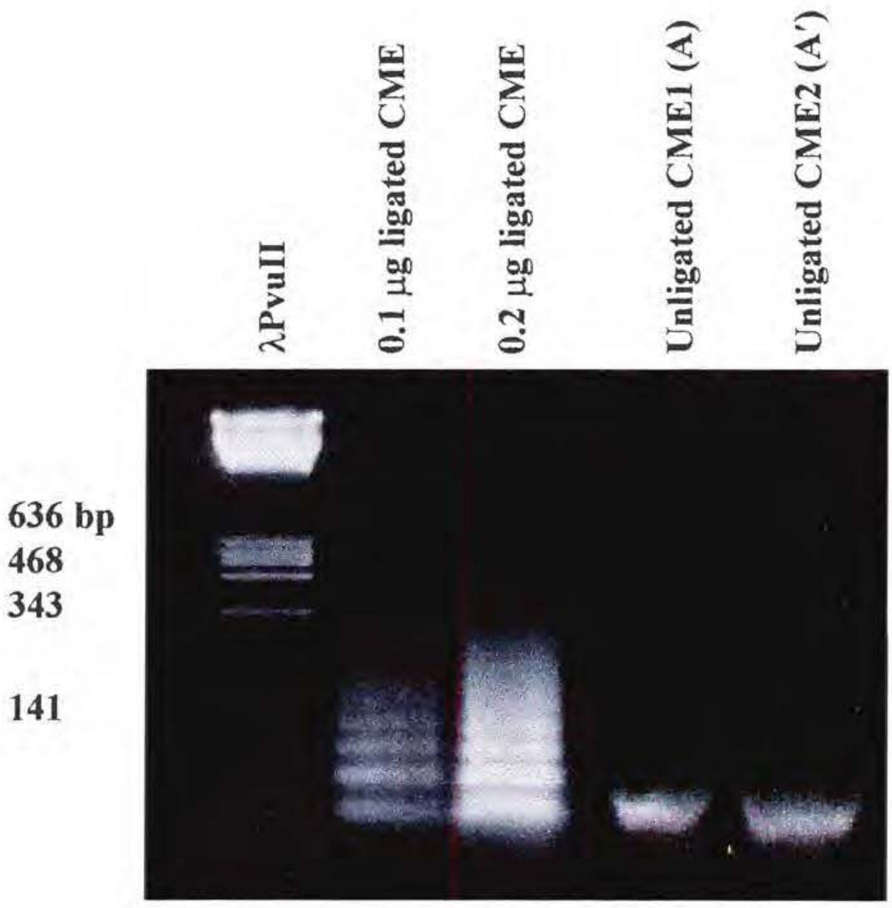


Figure 3.4: Agarose gel electrophoresis of multimers of CME oligonucleotides. Complementary CME oligonucleotides (A and A') were annealed, phosphorylated and ligated to generate multimers as outlined in figure 3.3; successful multimerisation was confirmed by electrophoresis on a 1.5 % agarose gel and the gel was stained with ethidium bromide. The sequences for **A** and **A'** are as described on figure 3.3.

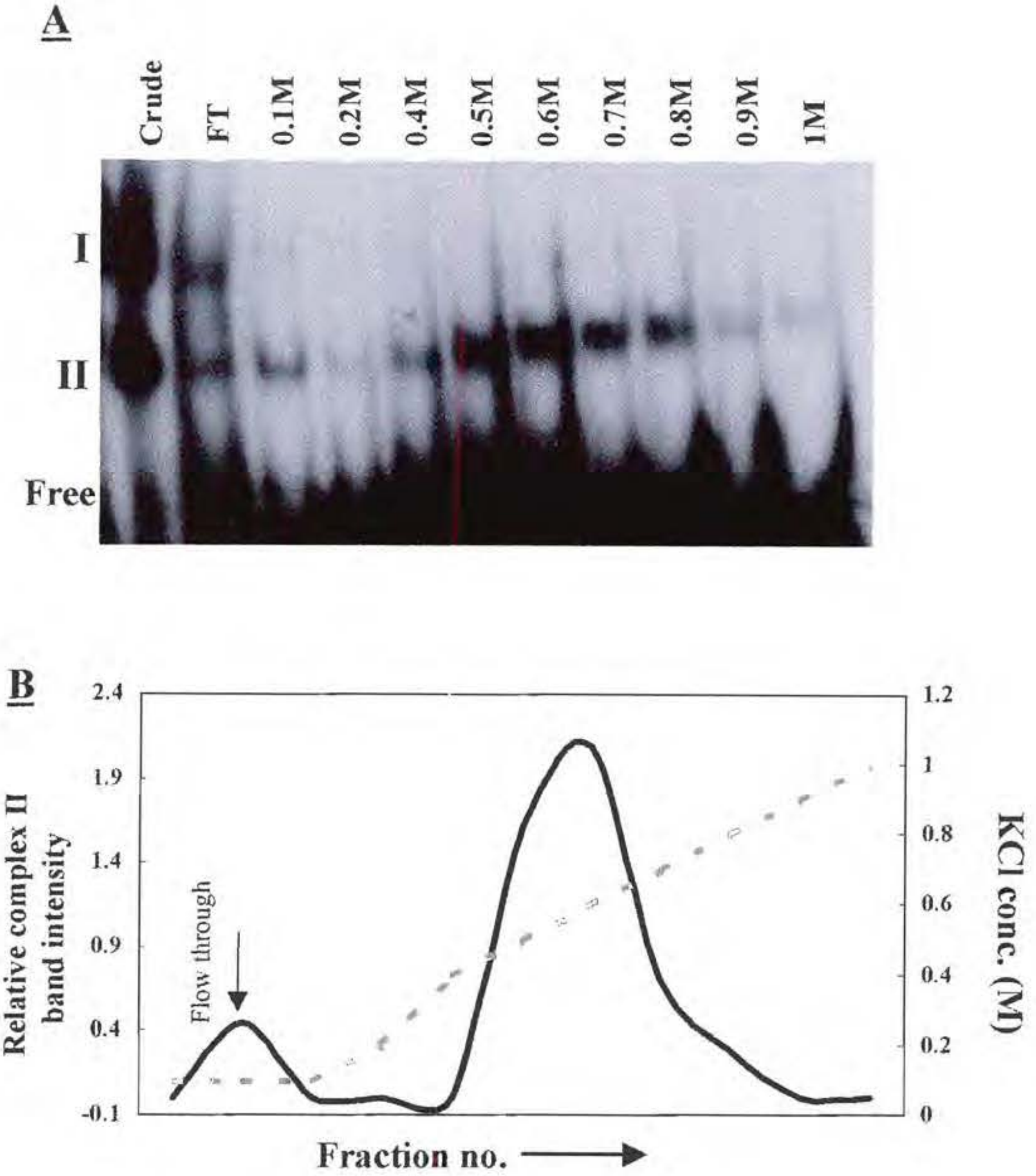


Figure 3.5: DNA affinity purification of Complex II proteins. The heparin agarose fraction containing complex II DNA-binding activity (Figure 3.2) was pre-incubated with 0.5 μg polydI.dC for 30 min on ice, diluted to 0.1 M KCl with a salt-free EMSA incubation buffer (section 6.7) and loaded onto the DNA affinity column (7 cm x 0.5 cm). Proteins were eluted with a [KCl] gradient from 0 to 1M with EMSA buffer. Fractions were analysed for DNA binding activity by EMSA as described in section 6.6.2. (**A**) Non-denaturing polyacrylamide gel electrophoresis of DNA affinity fractions analysed by EMSA. DNA affinity fractions were incubated with a ^{32}P-labelled -107 to -60 fragment of the α2(I) procollagen promoter. **B**) The relative band intensity of the autoradiogram was plotted against fraction number, the 0.5M and 0.6M KCl fractions contained the highest complex II DNA-binding activity. The broken line represents the KCl concentration gradient. **FT**~ flowthrough.

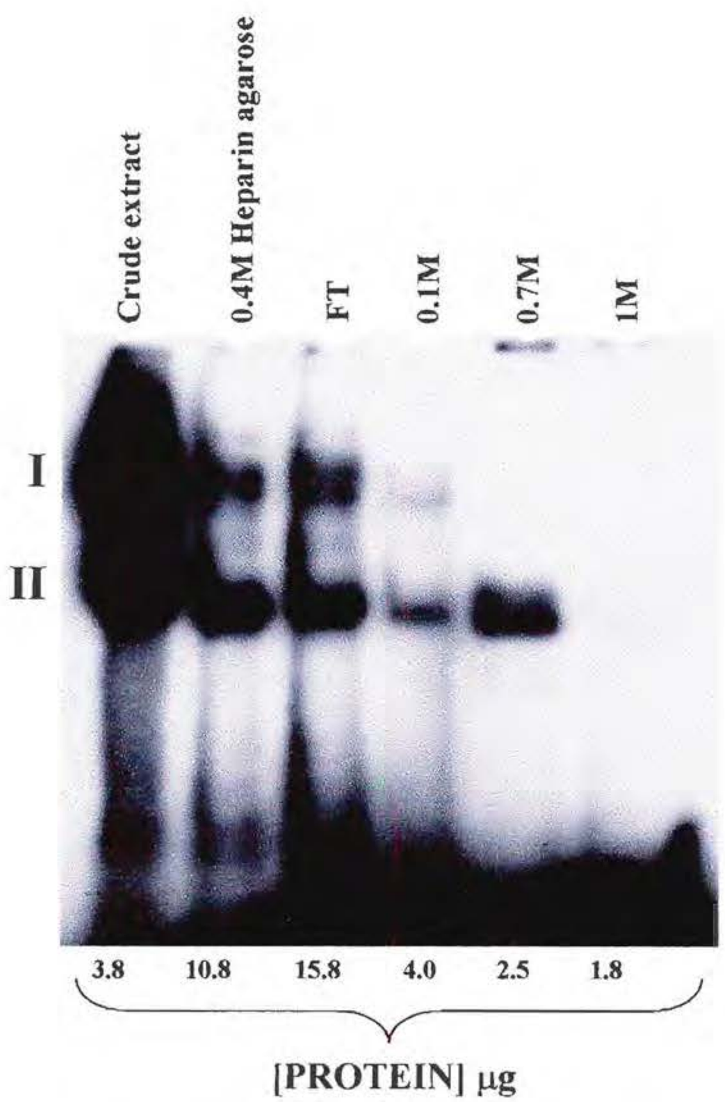


Figure 3.6: DNA-binding activity of proteins eluted from the DNA affinity column. The 0.4M heparin agarose fraction was incubated with polydI.dC for 30 min on ice and loaded onto the DNA affinity column (7cm x 0.5cm). The column was washed with 0.1M KCl in EMSA buffer and complex II proteins were eluted with buffer containing 0.7 M KCl. DNA affinity fractions were incubated with the -107 to -60 $\alpha 2(I)$ procollagen promoter fragment and DNA-protein complexes resolved by electrophoresis on a non-denaturing 5% polyacrylamide gel. The gel was exposed to X-ray film for 16 hrs at -70°C. The equivalent protein concentrations in each EMSA reaction are indicated at the bottom of each lane.

the SDS-sample buffer resulted in non-specific bands between ~50 and ~67 kDa, effectively making it very difficult to ascertain whether purification had been achieved. The exclusion of either DTT or β-MSH in the sample buffer, however, cleared the gel of any non-specific bands (see Figure 3.9). Although this was an affinity matrix, there were several contaminating bands in the complex II fraction but the complex was considerably homogeneous in comparison with the heparin-agarose fraction. A second affinity purification was performed in which case all complex II containing fractions from the first affinity purification were pooled and loaded onto a second affinity matrix column. Figure 3.8 illustrates the elution of complex II activity from the second affinity purification step; the accompanying silver stained SDS-polyacrylamide gel under reducing and non-reducing conditions is depicted in figure 3.9A and B. Two major polypeptides of ~50 kDa and ~67 kDa are present in the fraction containing complex II activity. Different volumes of the affinity fraction were loaded to ensure that the two major bands were indeed an indication of the extent of purification. Although a third affinity purification step was performed, too much of the sample was lost and this step was omitted from the purification procedure.

A summary of the purification results is presented on Table 3.1. Although it is difficult to calculate the activity of the transcription factors, it is generally assumed that 4ng of probe is required to bind ~1ng of complex II protein under standard EMSA conditions. Therefore, the concentration of the probe used in the DNA binding reaction is in excess of a specific transcription factor. To correlate radioactivity to units for the activity of each transcription factor, the amount of radioactivity in each band was determined by scanning the gel with a Packard Instant Imager. From Table 3.1, it is evident that the second affinity purification step improved the purification by 1186-fold but the yield was less than 10 %. The high specific activity support the SDS-PAGE results in that the last purification step showed very little contaminating protein bands, although they might be present in small quantities that are beyond the detection limit of the silver nitrate staining protocol.

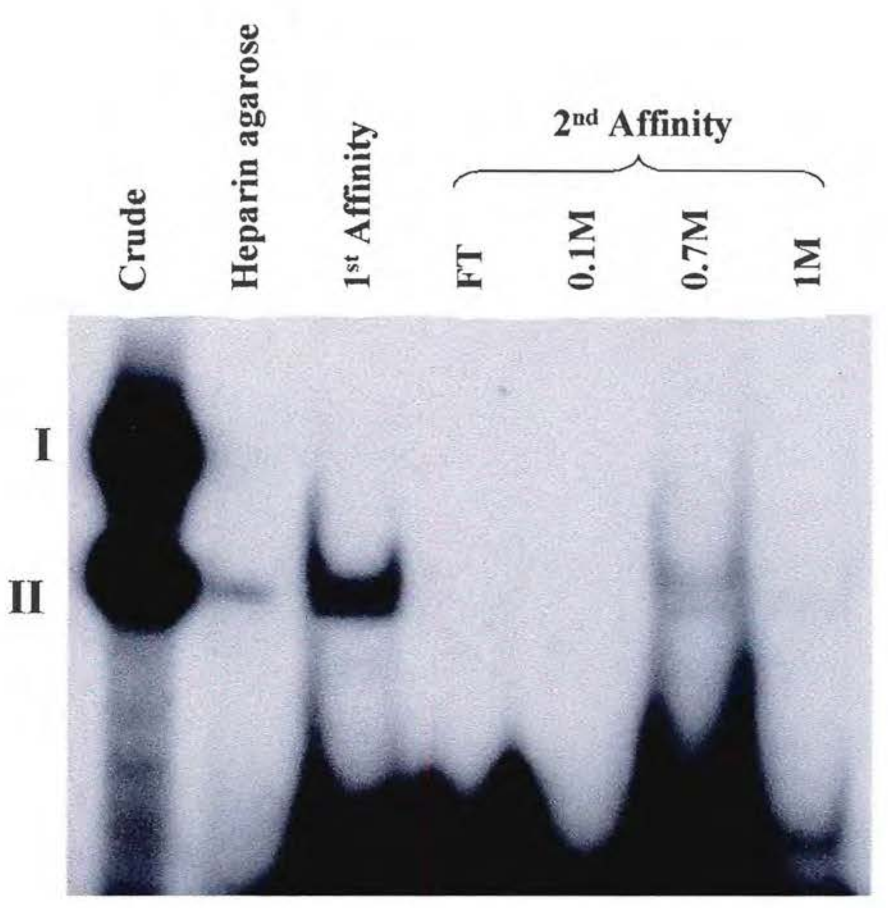
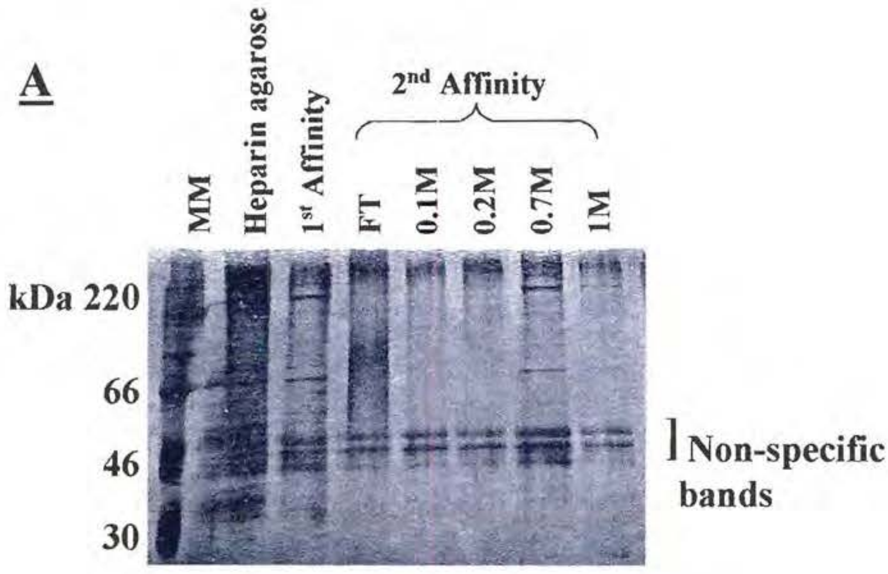


Figure 3.8: **EMSA analysis of the second DNA affinity purified complex II proteins.** The 0.7M KCl first affinity fraction was pre-incubated with polydI.dC on ice for 30 min, diluted to 0.1M KCl with the salt-free elution buffer and loaded onto second DNA affinity column (2 cm x 0.5 cm). Complex II proteins were eluted with a 0.7M KCl buffer and analysed for DNA-binding activity by EMSA as described previously. The autoradiograph shows a loss of complex II activity in the second affinity purification step.



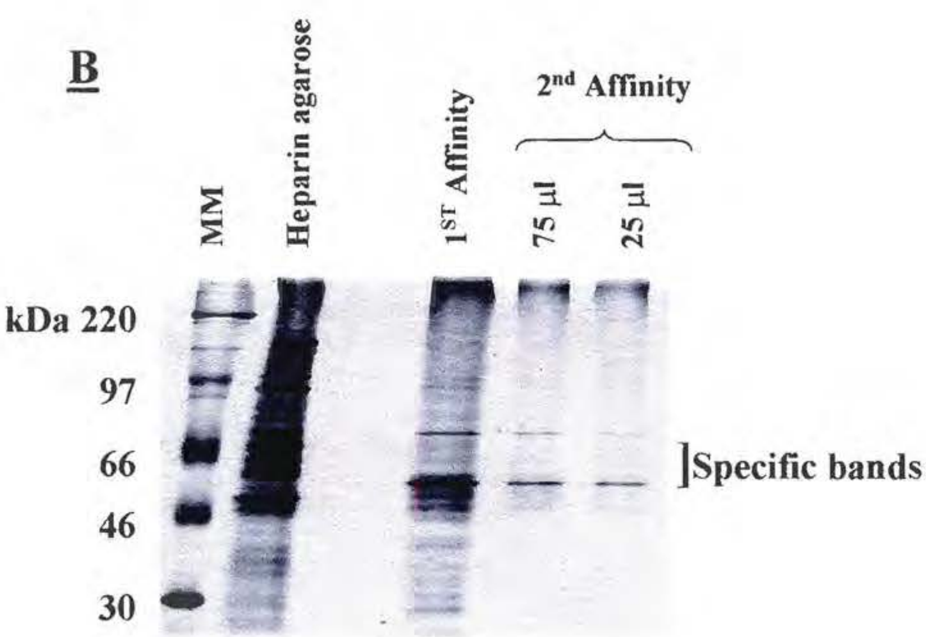


Figure 3.9: SDS-PAGE analysis of affinity fractions of complex II proteins. Proteins were precipitated with ice-cold acetone, the pellet washed with 70% (v/v) ethanol, dried and resuspended in Laemmli buffer (see section 6.7) in the presence **(A)** or absence **(B)** of β-mercaptoethanol. Proteins were resolved by electrophoresis on SDS-12 % polyacrylamide gels and the protein bands visualised by silver staining as described in section 6.6.5.
In **(B)** only the 0.7M KCl fractions were loaded on the gel; 75 µl and 25 µl of the 2nd affinity fraction were loaded to ensure that the bands observed were not non-specific artifacts.

Table 3.1: Summary of purification results

Purification step	[Protein] μg/ml	Activity (units) X 10^3	S.A[1] (units/μg) X 10^3	Yield[2] (%)	Purification factor (fold)
Crude	4865	147.37	0.0303	100	-
HepAgarose	12.15	94.52	7.779	64	257
1^{st} Affinity	1.50	21.59	14.393	15	475
2^{nd} Affinity	0.28	10.06	35.93	7	1186

The different fractions were incubated in the presence of a ^{32}P-labelled probe (section 6.4.9.1) and DNA-protein complexes resolved by electrophoresis on a non-denaturing 5 % polyacrylamide gel (section 6.6.2). The activity of the DNA binding protein was estimated from the radioactivity in each band on the EMSA gel using a Packard Instant Imager, they are expressed as relative units not absolute.

1~ specific activity was calculated by dividing the activity by the protein concentration;

2~ the yield was calculated as a percentage of the ratio of the activity of the preceding step.

3.2.2 Analysis and characterisation of CME binding proteins

Transcription factors are present in the cell at very low quantities of about 0.01 % or less of total cellular proteins, hence purification of transcription factors will remain one of the challenges in the study of gene regulation. To characterise the purified proteins, a number of strategies were employed in this study, these included the generation of antibodies, partial N-terminal sequence analysis, and matrix-assisted laser desorption/ionisation at the time of flight mass spectrophotometer (MALDITOF-ms) analysis.

3.2.2.1 Generation of antibodies against complex II protein

To raise antibodies, the single affinity purified fractions were used since this procedure requires the use of relatively high concentrations of protein. 10 μg of protein was mixed with complete Freund's adjuvant and the mixture injected subcutaneously into either three weeks old Balb/C mice or 6 months old rabbits. Animals were boosted with protein plus incomplete Freund's adjuvant every two weeks. Animals were given the last boost one week prior to a test bleed where 200 μl of blood was collected from the tail vein of mice or 2 ml from the hind leg of each rabbit. The blood sample was allowed to clot at 37°C for 16 hours and the serum was collected by centrifugation at 13 000 rpm at 4 °C for 10 minutes. 40 μg of crude nuclear proteins were resolved by electrophoresis on SDS-12 % polyacrylamide gels and immobilised onto nitro-cellulose membranes as described in section 6.6.7.2. The membrane was blocked with 5 % (w/v) fat-free milk in 1x Tris-buffered saline supplemented with 0.1 % (v/v) Tween 20 (TBST) for 2 hours or 16 hours. This incubation was important in order to reduce non-specific binding as shown in figure 3.10. The membrane was cut into strips and incubated with sera (200x dilution) from immunised or non-immunised mice or rabbits for 2 hours at room temperature, washed and incubated with a secondary antimouse IgG- or antirabbit IgG-horseradish peroxidase (HRP) conjugate. Bands were visualised by staining the membrane in the presence of hydrogen peroxide and 4-chloro-1-naphthol as a substrate. Only sera from rabbits contained antibodies to some of the proteins (Figure 3.10) with three major bands showing intense staining at ~38, 50 and 70 kDa. It was necessary to test if any of these antibodies were specific to complex II proteins since animals were immunised with the first affinity fraction that was not totally homogeneous. Nuclear proteins were pre-incubated with sera in the presence of polydI.dC, the -107 α2(I) procollagen probe was added to the reaction, and DNA-protein complexes were separated by electrophoresis on a non-denaturing 5 % polyacrylamide gel. From figure 3.11, it is evident that none of the sera contained antibodies against complex II proteins; the presence of a specific antibody should either inhibit complex formation by blocking the DNA-binding domain or supershift an existing complex due to the increased size associated with the presence of the antibody-protein-DNA complex.

3.2.2.2 Partial N-terminal sequencing of complex II proteins

The second affinity fractions that contained complex II activity (Figure 3.8) were pooled and ~40-60 μg protein separated by electrophoresis on a SDS-12 % polyacrylamide gel, transferred to a polyvinylidine difluoride (PVDF) membrane and stained with Coomassie blue (see section 6.6.7.1). Protein bands were excised and analysed on an automated protein sequencer. Attempts at sequencing generated only a 12 amino acid sequence for the 50 kDa protein: -

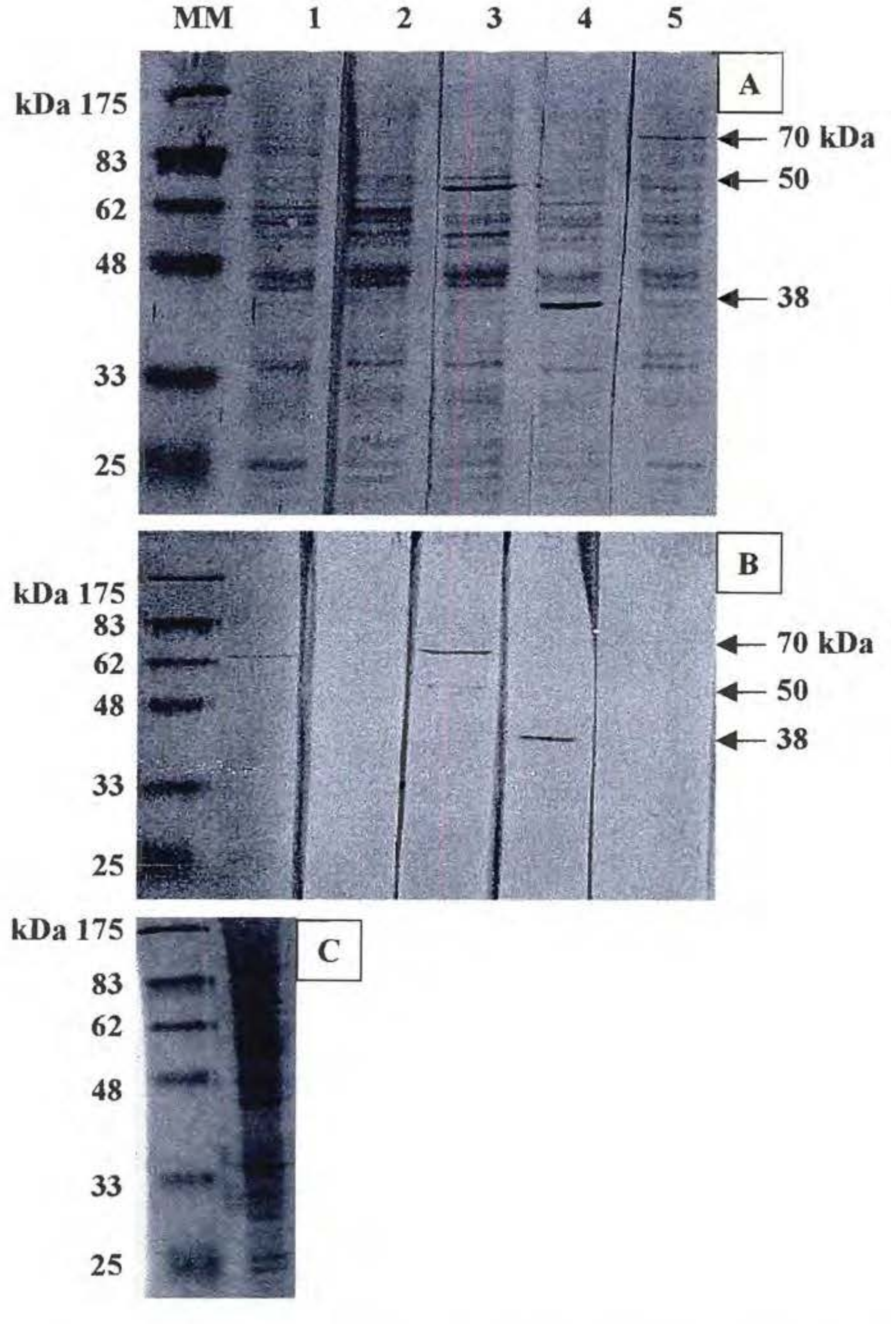


Figure 3.10: Western blot analysis of polyclonal antibodies raised in rabbits. 50μg of nuclear proteins were separated by electrophoresis on SDS-12 % polyacrylamide gels and immobilised onto nitro-cellulose membranes. The membranes were blocked with 5 % (w/v) fat-free milk in Tris-buffered saline containing 0.01% (v/v) Tween 20 (TBST) for 2 hours (**A**) or 16 hours (**B**), cut into strips and probed with polyclonal sera (200x dilution) from different immunised (**lanes 1-4**) and non-immunised (**lane 5**) rabbits. Specific protein bands were visualised by incubating with a secondary anti-rabbit IgG-horseradish peroxidase (HRP) conjugate for 45 minutes followed by staining with 4-chloro-1-naphthol dissolved in methanol and H_2O_2 as substrates (see section 6.6.7.2) **C**) A gel stained with Coomassie blue to show protein bands in the samples.

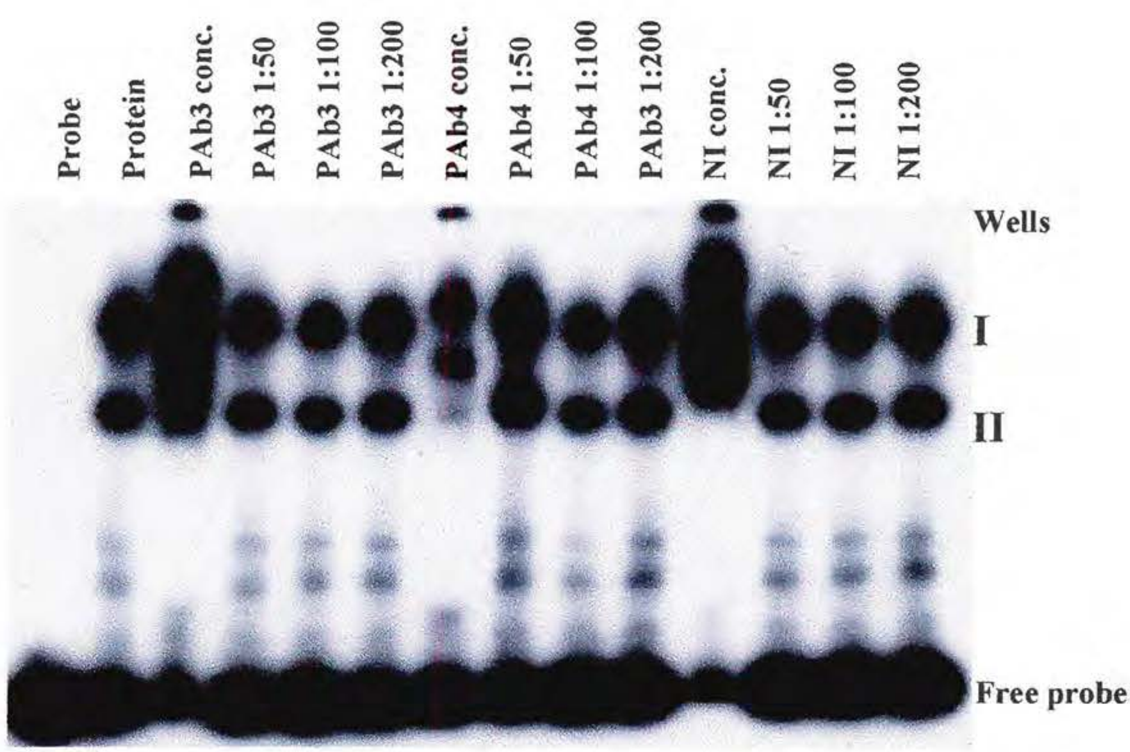


Figure 3.11: DNA-protein interaction on the -107 α2(1) procollagen promoter in the presence of rabbit polyclonal antibodies. Nuclear proteins were incubated with sera, at different dilutions, from immunised and non-immunised rabbits for 10 minutes at room temperature, a ^{32}P-labelled -107 to -60 α2(I) procollagen promoter fragment was added followed by a further incubation for 20 min on ice. DNA-protein complexes were resolved on a non-denaturing 5 % polyacrylamide gel. The appearance of supershifted bands was observed in all concentrated PAb samples, including the non-immunised serum sample, thus implying overloading of protein. **PAb**~ polyclonal sera, **NI**~ non-immunised sera.

Q[EGM]-EVG-EE[G]-E[G]-G; where the [ ] refer to either/or, while (-) indicate no signal.

These results were irreproducible and on a second sequencing attempt no signal was obtained even after using more protein. The 67 kDa protein did not generate any sequence, in fact no signal was obtained, suggesting that this polypeptide was N-terminally blocked.

3.2.2.3 MALDITOF-ms analysis of complex II proteins

Mass spectroscopy has become a powerful tool for the analysis of proteins, but more importantly, it has revolutionised the field of molecular biology and protein chemistry (207). MALDI is usually combined with a time-of-flight (TOF) mass spectrometry; ions are accelerated to a kinetic energy of approximately 25keV and subsequently allowed to fly through a fixed distance field-free region of ~1metre before they are recorded by the electric signal generated upon impact at the detector. The TOF in the field-free region is related to the mass:charge (*m/z*) ratio of a given ion. The use of mass spectroscopy coupled to MALDITOF permits identification and characterisation of proteins present at very low concentrations in cells, specifically transcription factors. Second affinity purified complex II proteins were separated by electrophoresis on a SDS-12 % polyacrylamide gel and the bands of interest (~50 and 67 kDa polypeptides) were excised and treated for *in situ* trypsin digestion as described in section 6.6.8. Proteins were digested with trypsin for 16 hours at 37°C and peptides eluted by alternate washes with 100 % (v/v) acetonitrile and 10 % (v/v) formic acid. The supernatant was dried and resuspended in a 0.1 % (v/v) triflouroacetate:acetonitrile (2:1) solution (TFA:ACN) and analysed by MALDITOF-ms using sinapinic acid as a matrix. Figures 3.12 and 3.13 illustrate a typical MALDITOF-ms spectra of peptide fragments present in the TFA: ACN-sinapinic acid mixture. The peptide masses were challenged against existing databank sequences using SwissProt.r35 on the http://prospector.ucsf.edu site. With a peptide mass tolerance of 1.500 Da, a 100 % match with part of a mouse zinc finger protein ZFP-27 (MKR4 protein) was obtained (Table 3.2). These preliminary results suggest

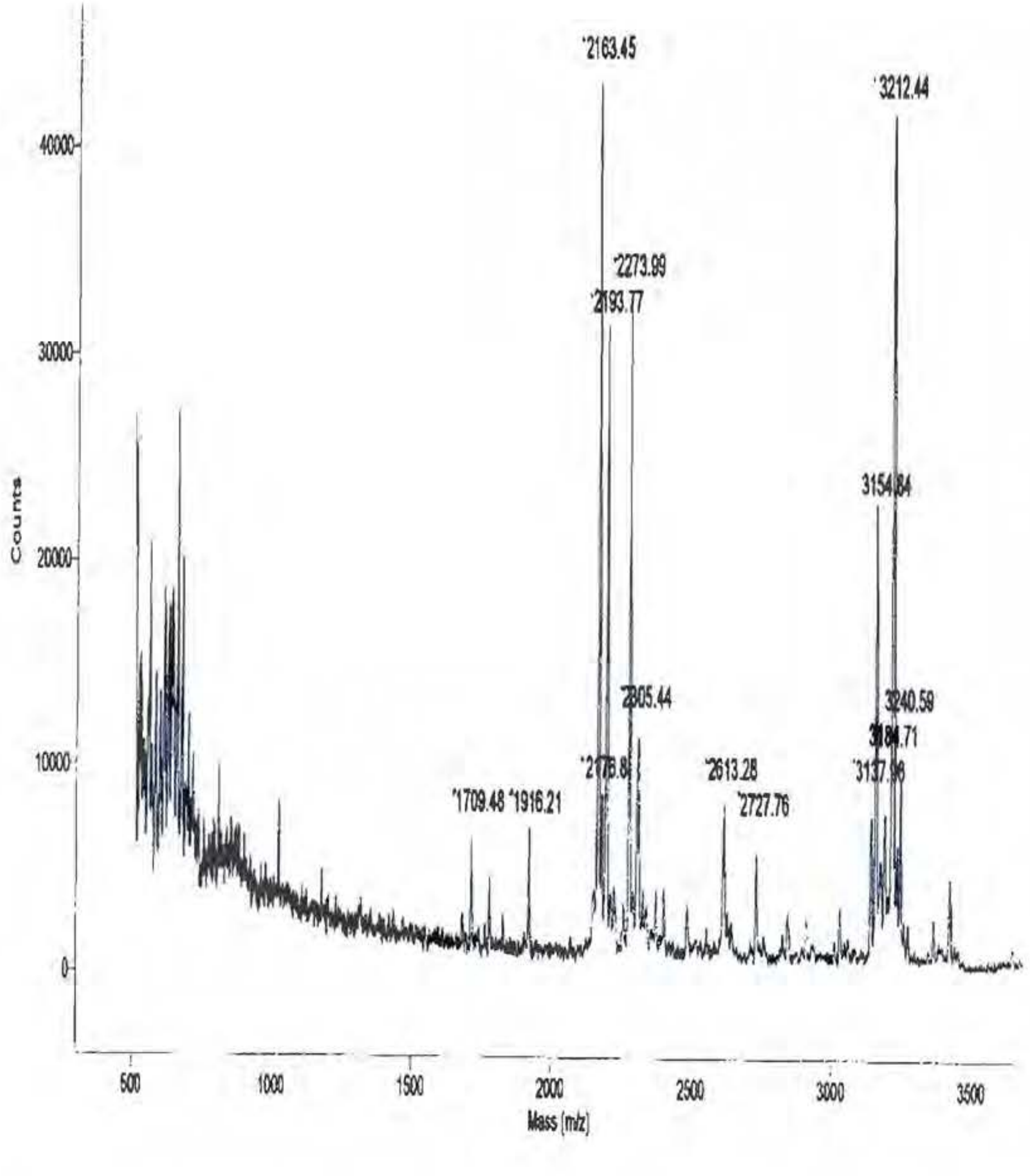


Figure 3.12: Mass spectrum for the 50 kDa polypeptide. Proteins purified by DNA affinity chromatography were separated by electrophoresis on a SDS-12% polyacrylamide gel, stained with Coomassie blue, the 50 kDa polypeptide band was excised, and processed for in-gel trypsin digestion as described in section 6.6.8. Peptide fragments were analysed by MALDITOF-ms (see section 6.6.9 for details of conditions). Numbers on the spectra indicate the mass:charge (m/z) ratios for each peptide.

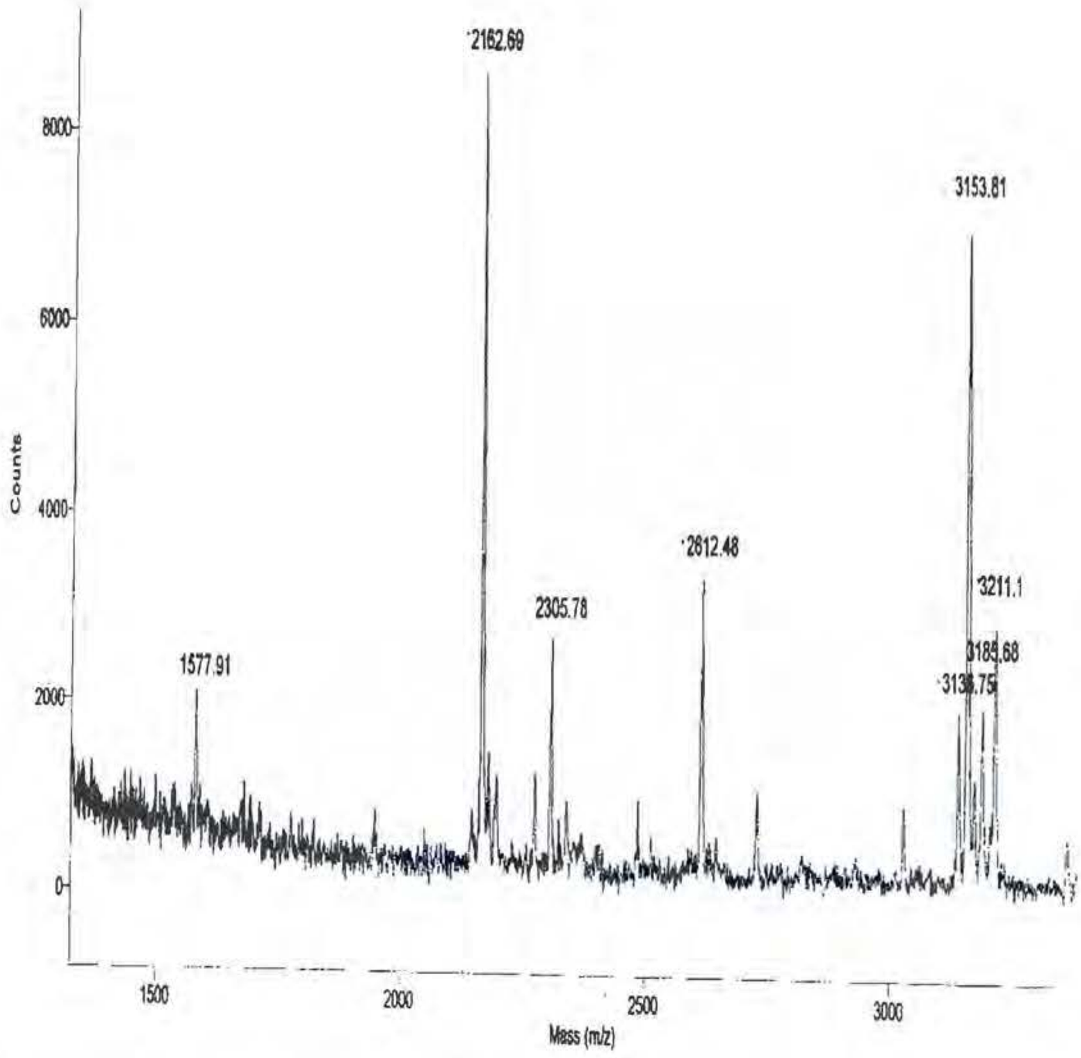


Figure 3.13: Mass spectrum for the 67 kDa polypeptide.

Proteins were separated by electrophoresis on a SDS-12% polyacrylamide gel as described in figure 3.11A, the 67 kDa polypeptide band was excised, and processed for in-gel trypsin digestion. Peptides fragments were analysed by MALDITOF-ms.

that polypeptides that constitute the complex II protein contain a zinc finger motif(s). Although a 100 % match was obtained, it does not follow that the complex II protein is ZFP-27, since the matched peptides covered only 23 % of the protein, furthermore, the molecular weight of ZPF-27 is only 47.341 kDa. If it is assumed that both the 67 and 50 kDa polypeptides associate to form complex II proteins, therefore the apparent molecular weight of this protein would be approximately 117 kDa. These results, although preliminary, provide evidence for classification of complex II proteins in the same class as Sp1, a zinc-finger transcription factor as well as a GC-rich binding protein. Further manipulations include the use of other proteases such as endoproteinase Lys-C to confirm the MALDITOF-ms profile obtained with trypsin and ultimately sequencing of some of the peptides remains crucial. The significance of the primary structure is crucial for designing primers to screen a cDNA library to search for the gene coding for this protein and of course at this stage it is possible to raise antibodies against some of the peptides to further characterise this protein.

Table 3.2: MS-Fit search results using SwissProt.r35

Data submitted	MH+ matched	Delta Da	Start A.A	End A.A	Peptide sequence
1577.9100	1576.8681	1.0419	58	70	(K)AFTQKSTLRMHQK(I)
3137.9600	3136.6821	1.2779	338	365	(K)AFNQKSILIVHQKIHTGEKPQVCAECGRA(A)
3184.7100	3185.6299	-0.9199	282	309	(K)AFTQRSEPVTHQRIHTGEKPYGCRPCGK(A)
3240.5900	3240.6597	-0.0697	142	169	(K)AFTYRSELHHQRTHTGEKPYQCGDCGK(A)

The molecular weight of ZFP-27 is 47341.8 Da, and the Acc. # P10077. The matched peptides cover 23 % of the protein.

3.3 DISCUSSION

The work described in this chapter demonstrates the difficulties experienced with purification of transcription factors. Successful purification depends on a high concentration of the starting material, the type, i.e. basal or inducer-specific, and stability of the transcription factor. Nuclear proteins were extracted from U937 monocytic cells, the first purification step involved ion-exchange chromatography using a heparin-agarose matrix. Complex II proteins eluted at high salt concentration, suggesting that these proteins form much stronger electrostatic interactions with the heparin-agarose matrix; thus implying that these proteins might have several exposed positively charged side groups which could be important for DNA-binding activity. The important aspect of the ion-exchange step, however, was that it allowed separation of complex I and II activities.

Purification of Sp1 by sequence specific DNA affinity chromatography was recommended for purification of other DNA binding proteins since this step enhanced the extent of purification by more than 60-100 fold in just one step (208). The next step in the purification process involved preparation of a sequence-specific DNA affinity matrix for further purification of complex II proteins. An interesting observation was that, although the matrix was specific for complex II proteins, other non-specific proteins were retained. This could arise from the non-ionic interactions between proteins and the Sepharose support through hydrophobic bonds. With repeated affinity purification, however, it was possible to improve the purity of the protein. In some of the published work on purification, authors used either hydroxylapetite, DNA-agarose and mutated or non-specific oligonucleotides to concentrate DNA-binding proteins prior to purification of a specific transcription factor on a sequence-specific DNA affinity matrix (199-201).

The use of conventional chromatography in the purification of transcription factors might appear tedious and old fashioned, however, it is still one of the common approaches in modern molecular biology. In a recent report, Wang and Kudlow (209) purified TEF1, a transcription factor that controls the human TGFα promoter, by the sequence specific DNA affinity chromatography from rat kidney nuclear extracts. Purified TEF1 was shown to interact with the proximal regulatory element of TGFα promoter and was able to activate transcription in HeLa cell nuclear extracts in an *in vitro* transcription experiment (209). Other recent reports where sequence specific DNA affinity chromatography was employed in the purification of transcription factors include a) copurification of two polypeptides from rat liver that were identified as histone H1d and H1c following microsequencing of peptides generated by endoproteinase Lys-C, b) purification of the human transcriptional cofactor complex CRSP (a complex with an approximate relative molecular mass of 700 kDa that contains nine subunits of molecular weights ranging from 33-200 kDa) that is required together with the TAF(II)s for transcriptional activation by Sp1, c) purification of a protein identified as CArG-box binding factor-A that binds to the SP6-κ promoter and d) purification of the heteromeric transcriptional activator MvaT of the *Pseudomonas mevalonii* mvaAB operon that encodes HMG-CoA reductase (EC 1.1.1.88) and HMG-CoA lyase (EC 4.1.3.4), enzymes that catalyse the initial reactions of mevalonate catabolism in this organism (210-213).

The extent of purification is commonly monitored by electrophoresis on SDS-polyacrylamide gels followed by silver staining since very low quantities of the protein are available for analysis. Silver staining has the disadvantage of introducing background staining in the presence of reducing agents like β-mercaptoethanol (β-MSH) (214). From the results section, it was evident that by excluding β-MSH, it was possible to visualise specific bands. Two polypeptides with electrophoretic mobilities on SDS-polyacrylamide gels of 50 and 67 kDa were obtained when complex II proteins were purified by DNA-affinity chromatography, however, the fraction was not 100 % homogeneous. In a study by Collins *et al* (56) the molecular weight of complex III proteins was estimated at 41 kDa from UV cross-linking experiments and South-Western blotting. Some of the lower molecular weight components could therefore be degradation products. Interestingly, during purification of TFIIA by affinity chromatography, precipitation with 5 or 20 % (w/v) trichloroacetic acid generated a number of minor bands with lower molecular weights than 38 kDa observed on SDS gels and the authors proposed that these bands could represent a pool of degradation products of the 38 kDa protein regarded as TFIIA (215). Thus it is possible that differences in the sizes of the low molecular weight protein could arise from proteolytic degradation during analysis by UV-cross linking, more so the 41 kDa polypeptide was associated with complex III proteins; in Chapter 2, complex III proteins were shown to be specific degradation products of complex II proteins. In addition, analysis of the second affinity fractions by SDS-PAGE followed by silver staining shows the presence of a smear below the 50 kDa band, supporting the suggestion that some proteolysis does occur during precipitation of proteins.

The aim of this section was to obtain the primary amino acid sequence of the protein in order to determine whether it is a known transcription factor or related to any known transcription factor. If this was not the case, then one avenue would be to synthesise primers for screening a cDNA library for the cDNA of interest. Initially we attempted to raise antibodies in mice and rabbits, although antibodies were present in sera from rabbits only, none of the antibodies reacted with complex II proteins as shown by EMSA studies. It is possible that antibodies of interest were not present because the antigenic determinant was not recognised as "foreign". More so, complex II proteins have been shown to be present in rodent cells (59) thus they might not be considered "foreign". Digestion with a protease prior to injection into mice/rabbits could have improved the immune response but this would have required the use of larger amounts of protein. In the literature, antibodies raised successfully against transcription factors usually involved the use of either synthetic peptides or recombinant proteins (216, 217).

The next attempt at identification and analysis of complex II proteins was to carry out N-terminal sequencing. From the results section it is apparent that the two polypeptides separated by electrophoresis on SDS-polyacrylamide gels were blocked at their N-termini. It is not uncommon for eukaryotic proteins to have blocked N-termini since they are known to undergo extensive post-translational modifications

II formation is zinc dependent (56). Transcription factors with Zn finger motifs include galactose-dependent activator (GAL4) protein, steroid receptors, *Krüppel* and *Hunchback Drosophila* proteins, c-Krox and Sp1 transcription factors (reviewed in 220). It is interesting that proteins like the Sp1 and c-Krox transcription factors recognise GC-rich regulatory sequences in promoters and introns of a variety of genes (39), in the same way that complex II proteins recognise a relatively GC-rich element in the $\alpha2(I)$ procollagen promoter. Although the MALDITOF-ms results implicate complex II proteins as a Zn finger protein, it should be emphasised that these results are still very preliminary. Work in progress is focused on obtaining the *m/z* ratios from peptides digested with other proteases like Lys-C and to pool some of the peptides of interest for sequencing by tandem mass spectroscopy.

To conclude, this Chapter reported on the successful purification by conventional chromatography of the complex II protein into two major polypeptides of 67 and 50 kDa in size on SDS polyacrylamide gels. Attempts at characterisation of the protein were hindered by very poor yields of the protein and failure to obtain antibodies. Although preliminary, it would appear that these polypeptides contain zinc finger motifs, suggesting that DNA binding activity of complex II is a zinc dependent process. In addition to conventional chromatography, ongoing research has incorporated the use of molecular biology tools to search for the cDNA encoding complex II proteins using the yeast one-hybrid system. The one-hybrid system allows screening of an appropriate cDNA library for the transcription factor of interest with an added advantage that the positive cDNA clone can be overexpressed in yeast without an additional subcloning step. This technique has been used successfully to clone BFCOL (48). The one-hybrid system differs from the two-hybrid system in that the latter uses cloned transcription factor motifs, for example the DNA-binding motif, to study protein-protein interactions. Also, further MALDITOF-ms analyses should yield informative data in terms of the possible primary structure of the complex II protein.

CHAPTER 4

MODULATION OF α2(I) PROCOLLAGEN GENE EXPRESSION BY CONSTITUTIVELY ACTIVE C-FOS

4.1 INTRODUCTION

Transformation of cells is very often accompanied by changes in the expression of a number of cellular proteins. Some of these proteins include extracellular matrix proteins such as fibronectin, type I and type VI collagen, enzymes such as collagenase and urokinase, cytoskeletal proteins, such as vinculin, and regulatory proteins such as β-glycan, integrin-associated protein, and myosin kinase (221). Andreú *et al* (222) have demonstrated that genes encoding TIMP3 and α2(I) procollagen were repressed by EGF-induced transformation. The expression of both these genes is believed to inhibit oncogenic transformation. The expression of type I procollagen in transformed cells has been a subject of many studies; its expression is altered following transformation by viral oncogenes (such as *v-mos, v-src, ras, v-fos* and E1A), chemical carcinogens, and ultraviolet light (see Chapter 1 for details). There is, however, no evidence to suggest that neoplastic or virally transformed cells synthesise a collagen type not found in normal tissues. A peculiar observation was made in WI-38 human embryonic lung fibroblasts transformed by γ-irradiation (CT-1 fibroblasts), in which case transformation hardly affected collagen synthesis, resulting in less than a 20 % decrease in the levels of type I collagen synthesis (32,33). Further characterisation of these fibroblasts revealed that they have decreased *c-fos* mRNA levels, even after induction with serum (89). These cells, however, have been shown to express elevated levels of *c-Ki-ras* 2 mRNA and overexpressed *c-myc* mRNA, characteristic of immortalised cells (90). This decrease in the level of *c-fos* mRNA does not appear to be a general feature of transformed phenotypes since in another WI-38 derived cell line transformed by SV40 (SVWI-38 fibroblasts), no repression of *c-fos* was observed. The SVWI-38 cell line, however, displayed the consistent feature of transformed fibroblasts with regards to type I procollagen mRNA and protein levels. The proto-oncogene *c-fos* was first discovered as the cellular homologue of the oncogene carried by two murine retroviruses, Finkel, Biskis and Jinkins (FBJ) and Finkel, Biskis and Reilly (FBR), first isolated from murine osteosarcomas (119, 120). Deregulated expression of the *c-fos* DNA in transgenic mice has been shown to interfere with normal bone development, suggesting that it plays an important role in bone cell physiology (174). The *c-fos* proto-oncogene, an early-immediate gene and a member of a multi-gene family that includes *fra-1*, *fra-2*, and *fos*B, codes for a protein with a molecular weight of ~42 kDa. Since the protein is highly phosphorylated, it usually migrates as a broad band between 54 and 62 kDa on SDS-polyacrylamide gels (127,117). Fos has been shown to function as a transcription factor, with either trans-activator or trans-repressor activity. It is synthesised in the cytoplasm, transported to the nucleus and together with the Jun family of proteins regulates the expression of a wide range of genes by binding as heterodimers to the AP-1 consensus sequences present in those genes (127). These proteins, however, can also modulate the expression of genes lacking AP-1 motifs via multi-protein-protein interactions involving other proteins such as the glucocorticoid receptor (223). These interactions will consequently affect diverse aspects of cell growth, differentiation and development.

The observed repression of *c-fos* levels in CT-1 fibroblasts was shown not to reflect changes in DNA elements in the promoter (89). Multiple regulatory sites in the promoter were identified, consistent with published data (117). These sites include: (a) a serum response element (SRE), (b) a TPA-response element (TRE), and (c) a cAMP response element (CRE) and a *c-sis*-PDGF inducible factor element (SIE). Induction of cells with any of these effectors has been shown to modulate the levels of *c-fos* mRNA and protein. The activity of the wild type promoter was similar for both parental WI-38 and transformed CT-1 fibroblasts, thus ruling out the possibility that γ-irradiation might have induced promoter DNA damage. Radiation, however, is known to alter the transcriptional activity of a number of genes; both broad-spectrum and near monochromatic (334 nm, 365 nm and 405 nm) UVA (320-380 nm) and near-visible radiations have been shown to strongly activate accumulation of mRNA corresponding to the nuclear oncogene *c-fos* (117,224). Thus UVA is capable of modulating genes containing active AP-1 based enhancer elements in the promoter region (117,224).

The fact that CT-1 fibroblasts retained the ability to express almost normal levels of the $\alpha 2(I)$ procollagen gene, suggested that repression of *c-fos* may play a role in the expression of the $\alpha 2(I)$ procollagen gene. The approach adopted in this study was to introduce a constitutively active *c-fos* cDNA into CT-1 fibroblasts and examine changes in the expression of the $\alpha 2(I)$ procollagen gene. Changes in $\alpha 2(I)$ procollagen gene expression were analysed by examining mRNA levels, the stability of the $\alpha 2(I)$ procollagen message, and to identify Fos response elements in the promoter and the first intron of the $\alpha 2(I)$ procollagen gene.

4.2 RESULTS

4.2.1 Stable transfection of *c-fos* into CT-1 fibroblasts

To study the regulation of the $\alpha2(I)$ procollagen gene expression by *c-fos*, a human embryonic lung fibroblast line transformed by γ-irradiation (CT-1 fibroblasts), was used. CT-1 fibroblasts have been shown previously to have decreased *c-fos* mRNA levels (89). Stimulation with serum and phorbol esters failed to increase *c-fos* to normal levels. This cell line is also ideal for studying the expression of $\alpha2(I)$ procollagen gene since, although transformed as judged by their ability to induce tumours in nude mice, these fibroblasts still express about 80 % of the $\alpha2(I)$ chain when compared to the parental cell line (32,33). Cells were transfected with a vector carrying the *c-fos* cDNA (Figure 4.1A) under the control of the murine leukaemia virus Long Terminal Repeat (MuLV-LTR), lacking the 3′ untranslated region to ensure that the transcript is stable, and therefore allowing the establishment of a constitutively active *c-fos* expressing cell line (173). The advantage of using this vector lies in the fact that it contains the neomycin resistance gene driven by the SV40 promoter for selection in eukaryotic cells and the ampicillin resistance gene for selective growth in *E.coli*. In addition to transfection with a *c-fos* cDNA, an empty vector (Figure 4.1B) was also transfected in parallel in order to ensure that the effects observed in the presence of the exogenous cDNA were not due to other sequences in the vector. The procedure for transfection and subsequent screening of neomycin resistant cells is discussed in section 6.2.1. Briefly, cells were transfected with 20 μg plasmid DNA using the calcium-phosphate precipitation method. Selection for stably transfected cells was carried out by treating cells with 200 μg/ml G418 until colonies started to appear (2-3 weeks), this took about 10 days. Colonies were trypsinised, plated onto 35 mm dishes and allowed to grow in the presence of G418 until confluency.

4.2.2. Screening of cells for integrated *c-fos* cDNA plasmid

In order to screen for cells that have integrated the transfected DNA, chromosomal DNA was extracted from 100mm dishes at 80 % confluency as described in section

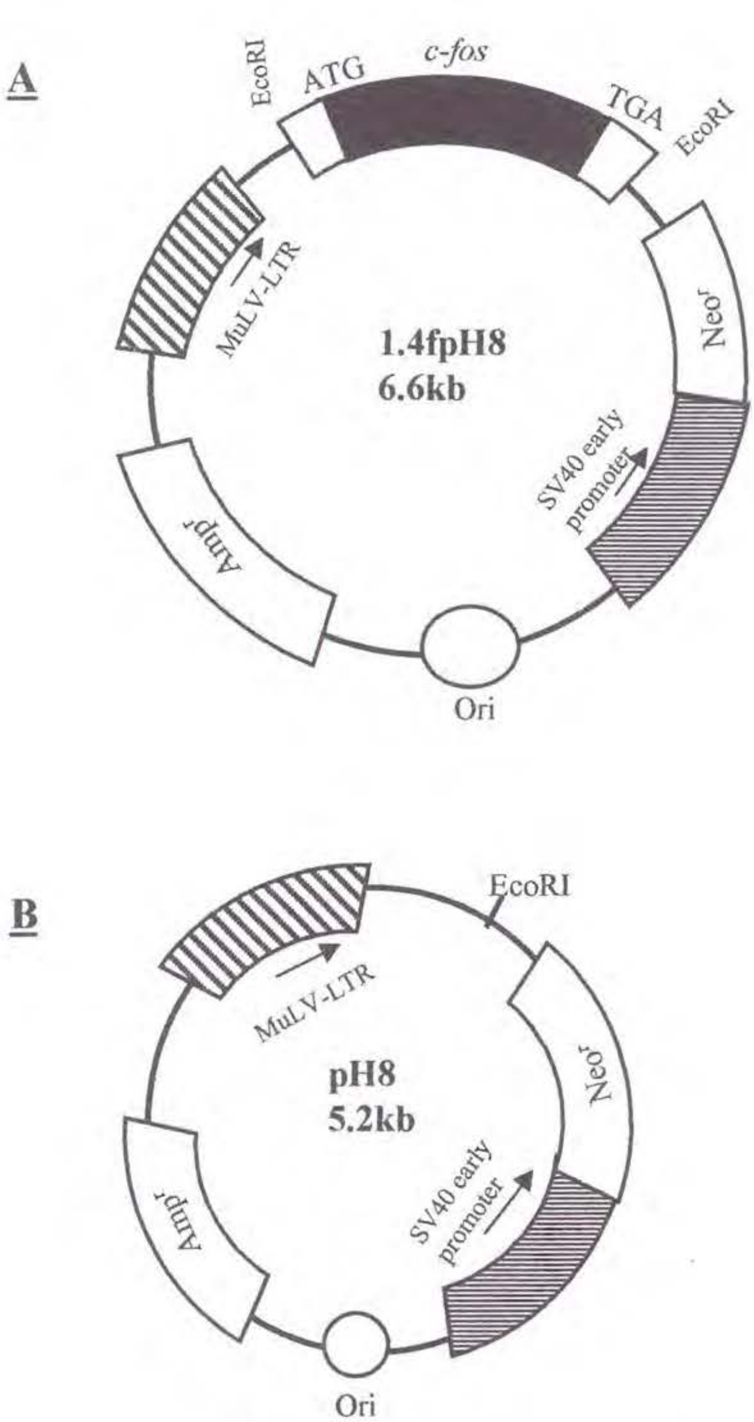


Figure 4.1: Constructs used in stable transfections. (A) 1.4fpH8: the 1.4 kb human *c-fos* cDNA inserted into the EcoRI site of the pH8 vector; the *c-fos* cDNA is under the control of the murine leukaemia virus long terminal repeat (**MuLV-LTR**), it also contains a neomycin (**Neo^r**) resistance gene under the control of the SV40 early promoter for selection in eukaryotic cells, whilst the ampicillin (**Amp^r**) resistance gene is for selection in bacteria. **(B)** pH8 vector: a self-ligated pH8 vector; this vector is similar in every respect to the 1.4fpH8 vector except that it lacks the *c-fos* cDNA. Plasmids were a generous gift from Dr Kuroki (173).

6.4.5. PCR primers were designed such that they allowed distinction of PCR products derived from the endogenous *c-fos* gene and the exogenous gene (Figure 4.2A); the primer set amplifies exon III, intron 3 and part of exon IV. Eight different *c-fos* clones gave two PCR products (Figure 4.2C) of 270 bp and 385 bp for the cDNA and endogenous gene, respectively. In those cells transfected with the empty vector and in untransfected cells, only the 385 bp PCR product was obtained (Figure 4.2B). Subsequent experiments were carried out with two or three Fos clones, selected at random since the different clones did not show any obvious differences.

4.2.3 Expression of *c-fos* mRNA and protein in stably transfected fibroblasts

It was important to show that the integrated cDNA was expressed prior to functional analysis of its effect on procollagen gene expression. Both the presence of the transcript and protein was determined. Total RNA was extracted from CT-1 cells grown to 80 % confluency, deprived of serum for 48 hours, followed by induction with 10 % (v/v) serum for 30 min. The stably transfected Fos clones were not subjected to serum stimulation since the *c-fos* gene is under the control of the MuLV-LTR and is therefore not induced by serum. RNA was fractionated on agarose formaldehyde gels (section 6.5.2), immobilised on a nylon membrane and probed with ^{32}P-dCTP labelled *c-fos* cDNA insert. Figure 4.3A depicts the levels of *c-fos* transcript in different cell lines, WI-38 human embryonic lung fibroblasts were used as a positive control for normal *c-fos* expression. Consistent with previous observations, both CT-1 and CT-PH fibroblasts have decreased *c-fos* mRNA levels. The expression of *c-fos* in transfected cells was higher than in the WI-38 fibroblasts. SVWI-38 fibroblasts were included in this experiment since they are also derived from the WI-38 parental cell line except they were transformed with SV40. In contrast to CT-1 fibroblasts, SVWI-38 fibroblasts expressed levels of *c-fos* transcript similar to the parental cell line (after correcting for differences in loading, Figure 4.3B and 4.3C), implying that transformation of WI-38 fibroblasts does not result in the shut down of *c-fos* expression. Again this observation highlights differences in the two transformed WI-38 fibroblasts. The membrane was also probed with β-actin to correct for loading differences (Figure 4.3B).

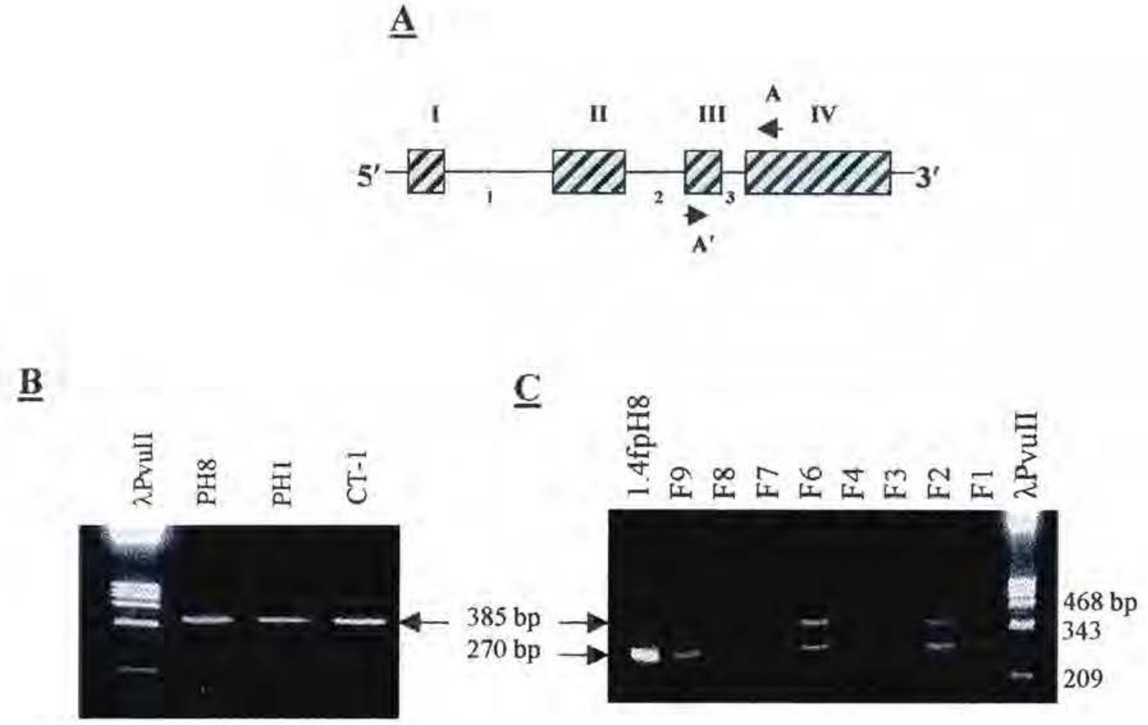


Figure 4.2: PCR screening for stably transfected *c-fos* colonies. (A) A schematic representation of the *c*-fos gene; PCR primers A and A′ were designed to amplify exon III, intron 3 and part of exon IV. **A**; 5′ TTA TCT CCA GAA GAA GAA GA 3′, **A′**; 5′ CAA GGG AAG CCA CAG ACA TCT 3′. **I-IV**; exons, **1-3**; introns. Chromosomal DNA was extracted from 80% confluent fibroblasts stably transfected with Fos and pH8 plasmids as described in section 6.4.5. 50 ng DNA was amplified using 10 pmol of primers A and A′ in a reaction mixture containing PCR buffer and 5 units of Takara *Taq*™ DNA polymerase. The optimum PCR conditions for amplification were 94 °C for 1 min, 60 °C for 1 min and 72 °C for 1 min for 30 cycles. PCR products were analysed by electrophoresis on 1.5% agarose gels and bands visualised by ethidium bromide staining. **(B)** PCR products (endogenous 385 bp) from DNA extracted from CT-1 and CT-1 fibroblasts stably transfected with the empty vector (PH). **(C)** PCR products (endogenous 385 bp and exogenous 270 bp) from DNA extracted from the different colonies stably transfected with the *c-fos* cDNA vector (F1-F9).

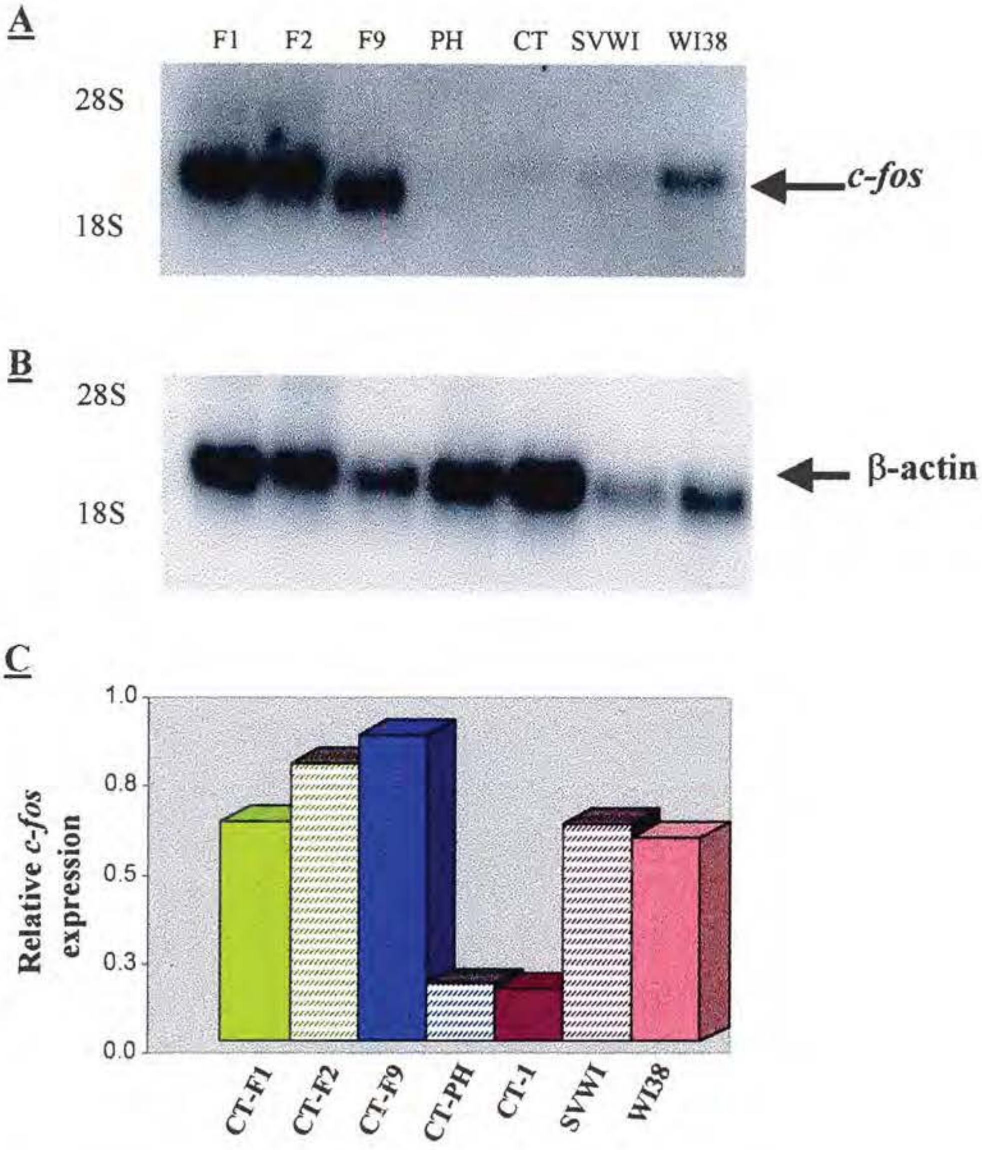


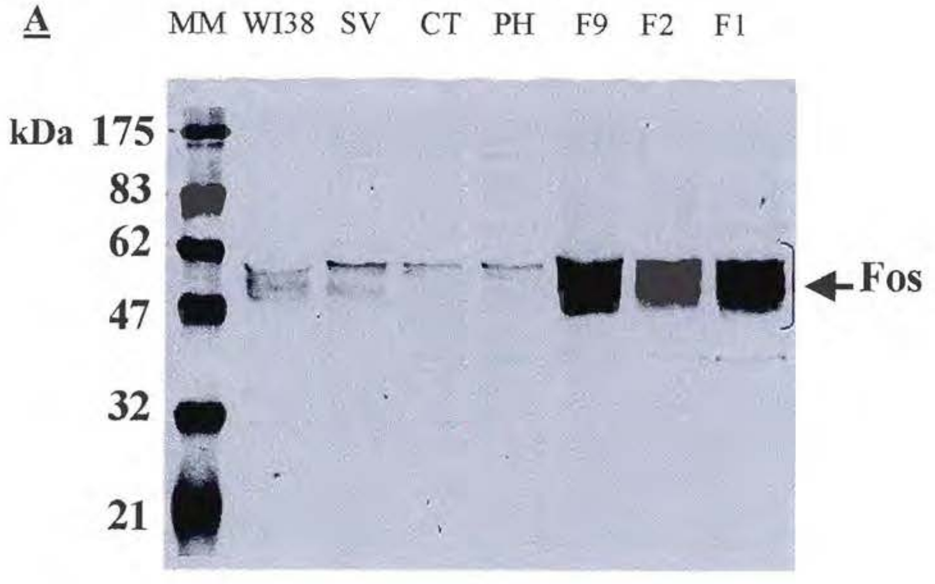
Figure 4.3: **Differential expression of *c-fos* in fibroblasts.** Fibroblasts were deprived of serum for 48 hours, except for the Fos transfected clones, stimulated with 10% serum for 30 min, and total RNA extracted using the one step method of Chomczynski and Sacchi (243) as described in section 6.5.1. RNA was resolved by electrophoresis on a 1% (w/v) agarose formaldehyde gel, stained with ethidium bromide and immobilised on a nylon membrane. **A)** The membrane was hybridised to a ^{32}P-labelled *c-fos* cDNA probe. **B)** The membrane was stripped in a 0.1% (w/v) SDS solution heated to 99 °C and rehybridised to a ^{32}P-labelled β-actin cDNA probe. **C)** Graphical representation of the relative *c-fos* mRNA levels in the indicated fibroblasts; the relative *c-fos* mRNA was calculated from the ratio of the band intensity of *c-fos* to β-actin obtained by densitometric scanning of the autoradiographs.

To determine the protein levels in these clones, adherent cells were extracted in a boiling Tris-SDS solution, incubated at 90 °C for 10 min, centrifuged, the supernatant mixed with 2x Laemmli loading buffer (section 6.7) and proteins resolved by electrophoresis on a SDS-12 % polyacrylamide gel (section 6.6.4). The Fos protein was detected by immunostaining after transfer to a nitro-cellulose membrane and probed with an antibody against Fos. As can be seen from figure 4.4A, *c-fos* clones expressed more than 4-fold the Fos protein produced in the serum induced WI-38 cell line. Very little Fos was detected in CT-1 and CT-PH fibroblasts. A Coomassie blue-stained gel was used to compare loading differences (Figure 4.4B).

Microscopic examination of the Fos clones showed no obvious changes in the morphology of the cells. Changes in the rate of cell growth was determined by cell counts using a Coulter Counter over a period of four days. No changes in the doubling time of cells was observed (Figure 4.5), implying that constitutive expression of Fos in these cells does not affect cell growth or morphology.

4.2.4. Modulation of $\alpha2(I)$ procollagen gene expression by constitutively active Fos

Cell transformation is usually accompanied by changes in the expression of type I procollagen. In most cases transformation leads to reduced type I collagen levels and the two alpha chain genes are not necessarily affected to the same extent. In the case of CT-1 fibroblasts, however, collagen synthesis was not affected drastically. The objective of this study was to determine whether the extremely low levels of Fos in these cells was in any way related to the inability to downregulate the $\alpha2(I)$ procollagen gene. To this end, the changes in type I procollagen gene expression, specifically the $\alpha2(I)$ chain were studied in the *c-fos* transfected fibroblasts. Fibroblasts were serum starved for 48 hours and stimulated with 10 % (v/v) serum, total RNA was extracted, electrophoresed on 1 % (w/v) agarose formaldehyde gels, blotted onto nylon membrane and probed with both a *c-fos* and $\alpha2(I)$ procollagen probes. The induction with serum is necessary because *c-fos* and other early immediate genes are normally expressed transiently following stimulation of cells



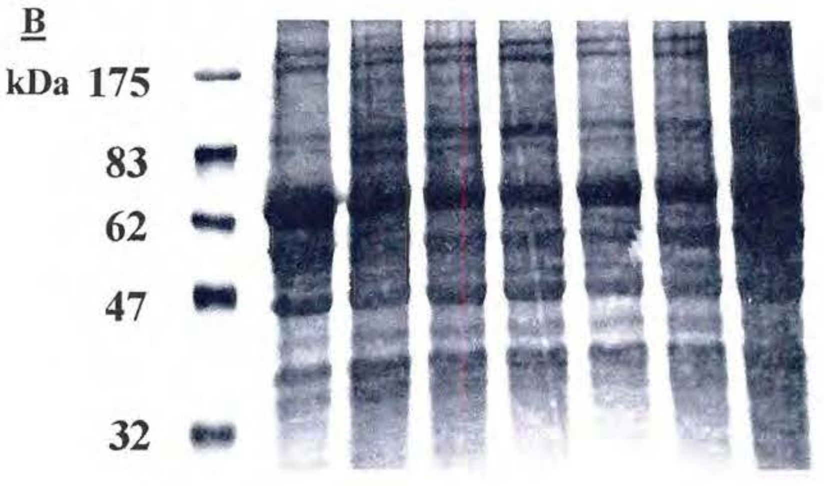


Figure 4.4: **Fos protein levels in stably transfected CT-1 fibroblasts**. Fibroblasts, except for F1 and F9 fibroblasts, were deprived of serum for 48 hours followed by a 30 min induction with 10% (v/v) serum, harvested with a boiling Tris-SDS solution, the suspension was centrifuged and the supernatant was retained; 100 µg of protein was mixed with Laemmli buffer (section 6.7) and resolved by electrophoresis on a SDS-12% polyacrylamide gel. **A)** Proteins were transferred onto a nitrocellulose membrane, blocked with 5% (w/v) fat free milk in Tris-buffered saline, 0.01% (v/v) Tween 20 (TBST) and probed with a rabbit anti-Fos antibody (1:250 dilution). The membrane was washed and incubated with an anti-rabbit IgG-horseradish peroxidase (HRP) conjugate (1:2000 dilution). Bands were visualised by staining with a 4-chloro-1-naphthol/methanol solution in the presence of 0.02% (v/v) H_2O_2 as described in section 6.6.7.2 **B)** Coomassie blue stained gel of 100 µg protein resolved by electrophoresis on a SDS-12% polyacrylamide gel.

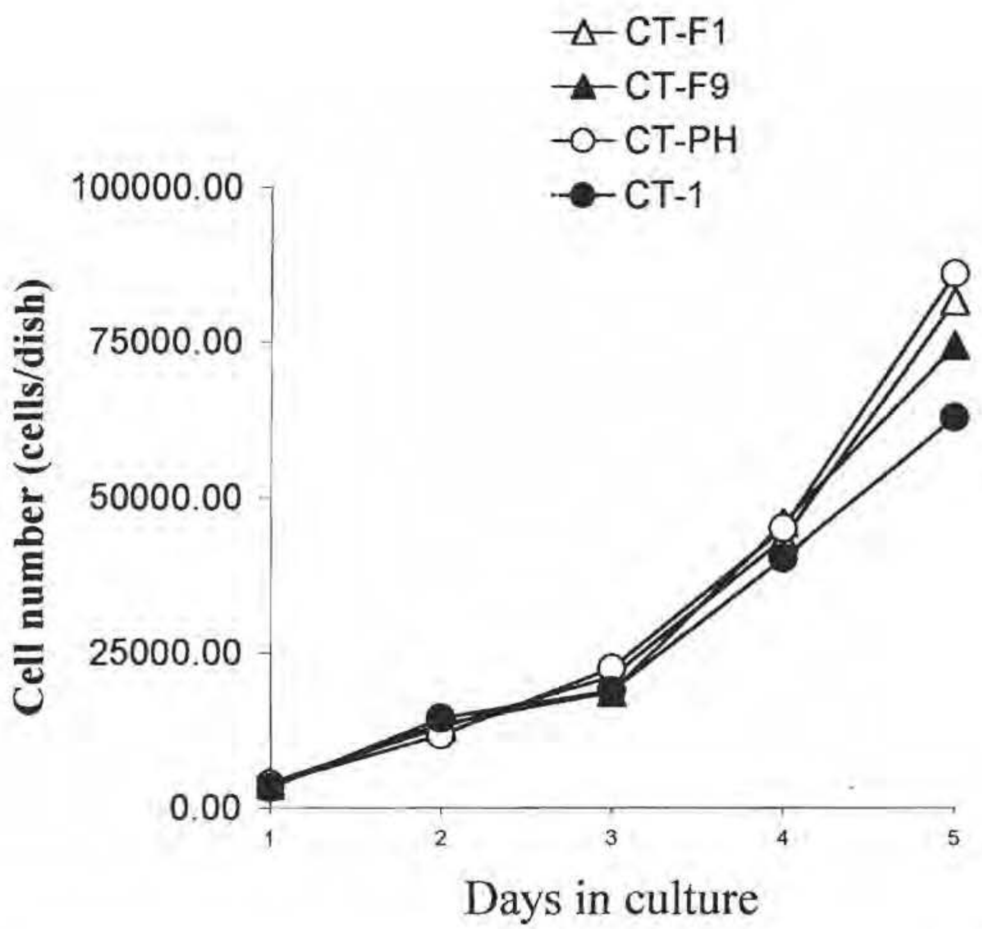


Figure 4.5: The growth profile of CT-1 fibroblasts overexpressing Fos. Fibroblasts were seeded at ~3.0 x 10^3 cells/ml in 35 mm dishes as described in section 6.1.1; each day, adherent cells were trypsined and 200μl counted using a Coulter counter. The cell number was calculated and expressed as cells/dish.

with growth factors. In order to have a much more comparable system, all fibroblast cultures were induced with serum to activate *c-fos* and any of the other early response genes that require activation by growth factors, e.g. Jun family of proteins which are required for heterodimerisation with Fos to bind to AP-1 consensus sequences. From the data on figure 4.6A and 4.6B, it is evident that constitutive expression of *c-fos* results in the down-regulation of $\alpha2$(I) procollagen mRNA levels. It is evident from figure 4.6C that comparable amounts of RNA were loaded as shown by the ethidium bromide stained membrane following transfer of the RNA from the gel to the nylon membrane.

4.2.5. Stability of $\alpha2$(I) procollagen mRNA in CT-1 fibroblasts

The regulation of gene expression can occur at different levels; these include transcription factor binding, RNA polymerase activity, mRNA processing, and mRNA stability and transport to the cytoplasm. It was clear that high levels of Fos reduced the levels of $\alpha2$(I) procollagen mRNA. The next step was to establish what mechanism(s) is involved in this modulation of procollagen gene expression. Transformation of Rat-1 fibroblasts with mutant *ras* has been shown to down-regulate and destabilise the $\alpha1$(I) procollagen mRNA (21). Since Fos acts downstream of Ras, it was pertinent to establish whether down-regulation of $\alpha2$(I) procollagen gene was a result of enhanced degradation of the transcript in the presence of Fos. Cells were treated with 5μg/ml actinomycin D and total RNA extracted at 0, 2, 4, 6, 8, 16 and 24 hour time points. Figures 4.7A-D represent Northern blots of total RNA probed with both $\alpha2$(I) procollagen and β-actin. It should be noted that since β-actin is also sensitive to actinomycin D treatment, it was important to ensure equal loading of RNA samples as illustrated by the adjoining membranes stained with ethidium bromide. In all cell lines analysed, procollagen RNA levels increased up to 8 hours followed by a slow decrease at 16 hours. In addition to the decrease in the $\alpha2$(I) mRNA, the β-actin transcript was also affected. By 24 hours, there was a general reduction in total RNA (18S and 28S) and this was accompanied by a decreasing cell viability. This is in line with published data on the half-life of $\alpha2$(I) procollagen mRNA, which has been estimated to be between 16 and

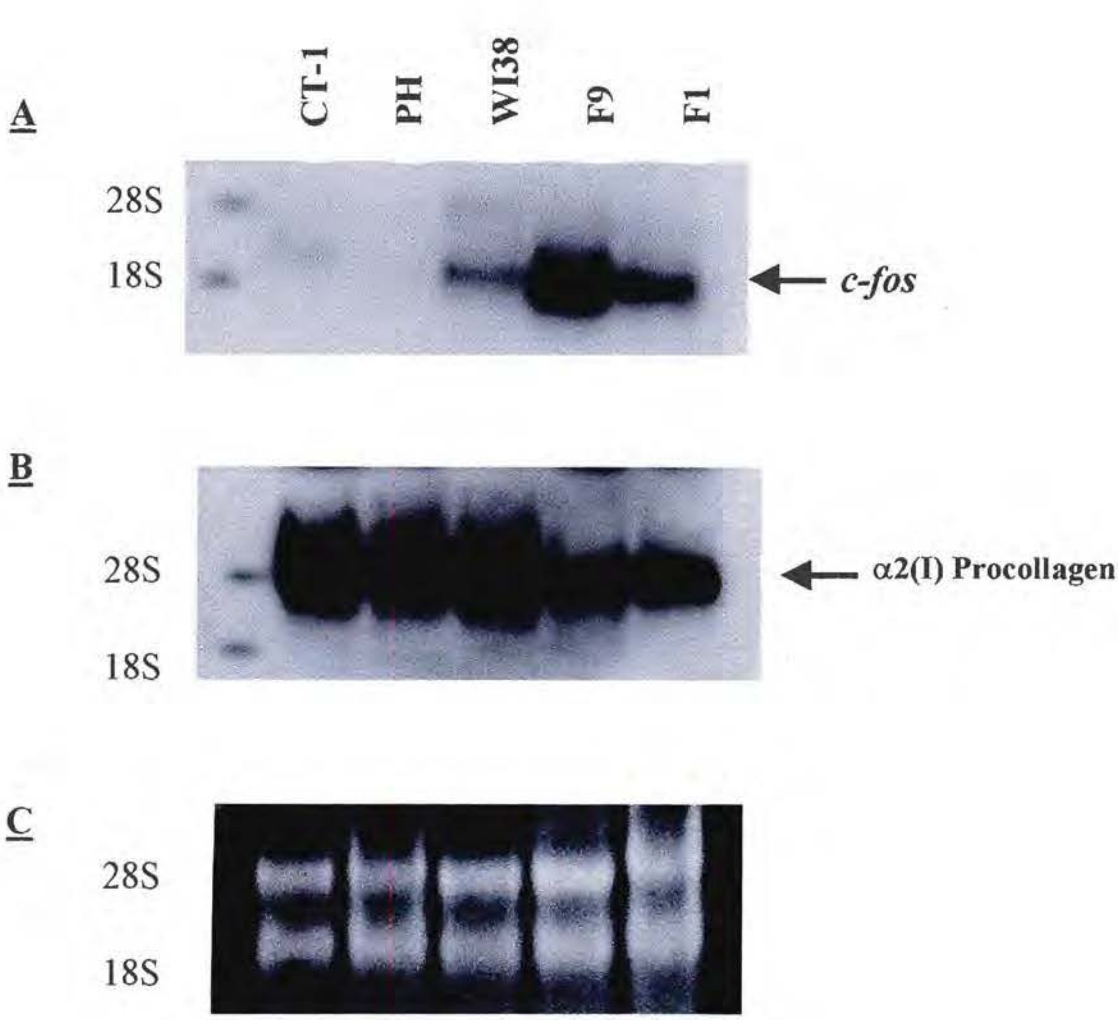


Figure 4.6: **Modulation of α2(I) procollagen mRNA levels by *c-fos*.** Fibroblasts were serum starved for 48 hours and induced with 10% (v/v) serum and total RNA extracted using the one-step method described in section 6.5.1. 10μg of RNA was fractionated by electrophoresis on a 1% (w/v) agarose formaldehyde gel followed by blotting onto a nylon membrane. **(A)** *c-fos* mRNA in fibroblasts induced with 10% (v/v) serum; the blot was hybridised to a ^{32}P-labelled *c-fos* cDNA probe for 16 hours at 42 °C and a high stringency wash was performed with 0.1x SSC (section 6.7) containing 0.1% (w/v) SDS for 15 min at 56 °C; **(B)** α2(I) procollagen mRNA levels in fibroblasts; the membrane was hybridised to a ^{32}P-labelled α2(I) procollagen cDNA probe and washed with 0.1x SSC containing 0.1% (w/v) SDS for 30 min at 65 °C ; **C)** The membrane was stained with ethidium bromide after transfer and prior to hybridisation to confirm equal loading.

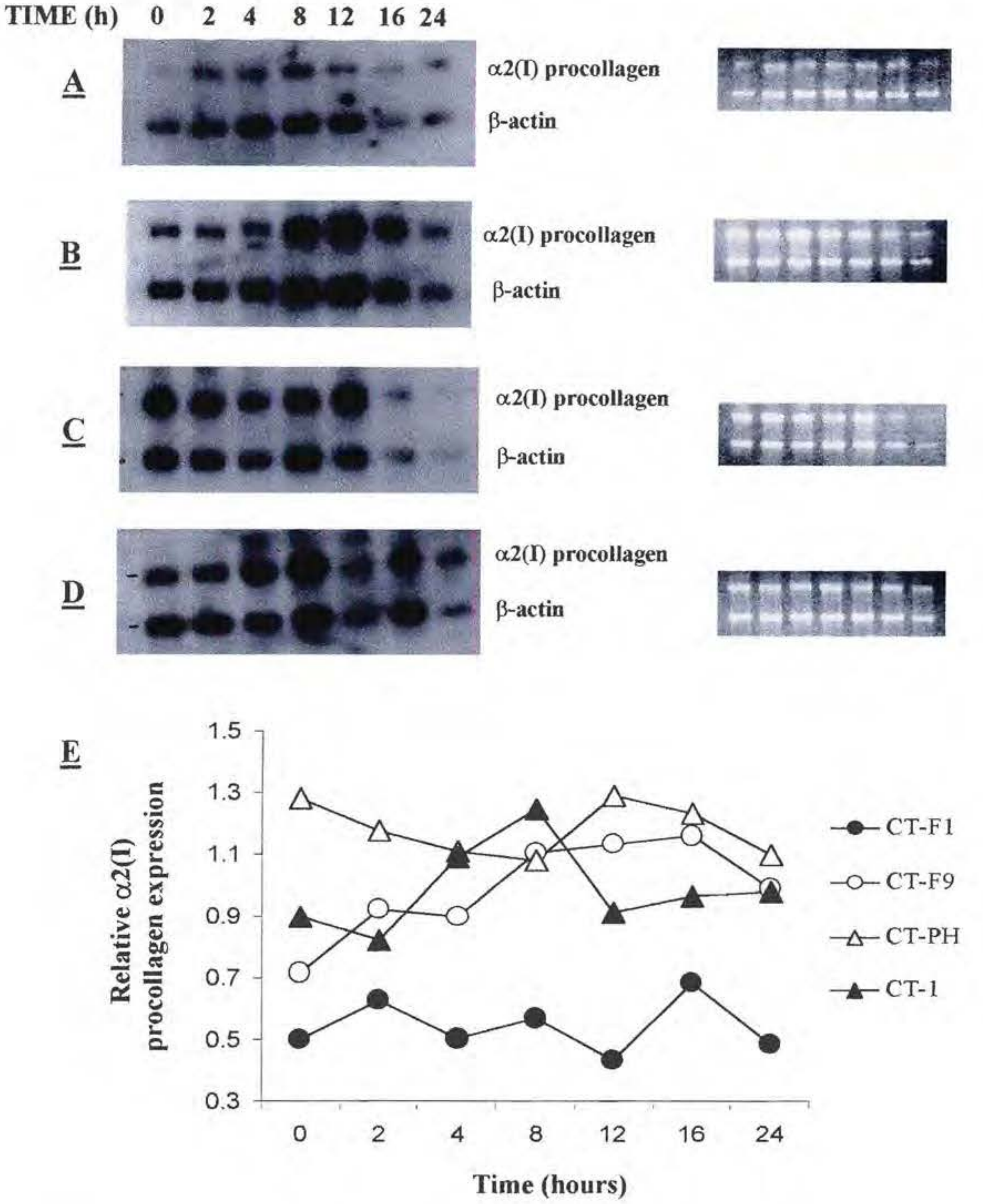


Figure 4.7: The stability of the α2(I) procollagen mRNA. Fibroblasts were treated with 5 μg/ml actinomycin D for 0, 2, 4, 6, 8, 16, and 24 hours and total RNA was extracted at the indicated time points. RNA was separated on 1% (w/v) agarose formaldehyde gels, blotted onto nylon membranes and double probed with α2(I) procollagen and β-actin cDNA probes. Northern blot results for **A)** CT-F1, **B)** CT-F9, **C)** CT-PH and **D)** CT-1 fibroblasts; the ethidium bromide stained membranes are shown on the right hand side. **E)** Relative α2(I) procollagen expression; the band intensities were obtained by densitometric scanning of the autoradiographs and the relative expression of the α2(I) procollagen gene was calculated from the ratios of α2(I) procollagen mRNA to β-actin mRNA and plotted against incubation time in actinomycin D.

18 hours. These results suggest that constitutive expression of *c-fos* does not alter the half-life of the α2(I) procollagen mRNA.

4.2.6 Transcriptional regulation of the α2(I) procollagen gene by Fos

4.2.6.1 Analysis of α2(I) procollagen proximal promoter for the presence of Fos response elements

Results obtained in section 4.2.5 suggested that the effects of Fos are not mediated at the level of message stability, thus ruling out a post-transcriptional event in the form of increased mRNA turnover. Transcriptional control is mostly related to the ability of different transcription factors to assemble at the promoter to allow formation of a transcription preinitiation complex through a series of events that ultimately lead to transcription of a specific gene. It was imperative to establish which of the promoter elements modulate Fos mediated repression of the α2(I) procollagen gene. The α2(I) procollagen proximal promoter constructs fused to the CAT reporter gene were used in transient transfections experiments. Cells were transfected with 10-20 μg of plasmid DNA using the calcium-phosphate precipitation method as described in section 6.2. The initial focus was on the proximal promoter since: 1) it has been shown to be critical in determining tissue/cell specific expression of the gene, 2) this promoter fragment contains an AP-1 like binding element, and 3) a novel transcriptional repressor has been mapped in this region of the α2(I) procollagen promoter (13,35,42,51,57,59). The promoter activity of the -343 COLCAT construct was not significantly changed in CT-F9 and CT-PH fibroblasts compared to CT-1 fibroblasts. The relative CAT activity of the CT-F1 fibroblasts on the other hand, increased significantly by about 1.5-fold relative to CT-1 fibroblasts (Figure 4.8). It appears that the proximal promoter was transactivated in the one Fos overexpressing cell line. This result could be explained by the presence of an AP-1-like sequence at -255 of the proximal promoter, which has been shown to bind AP-1 proteins (42). To further exclude the possibility that Fos transactivated a novel repressor (described in chapter 2), DNA-protein assays were carried out using an oligonucleotide which is recognised by the repressor and a well characterised CCAAT recognition sequence.

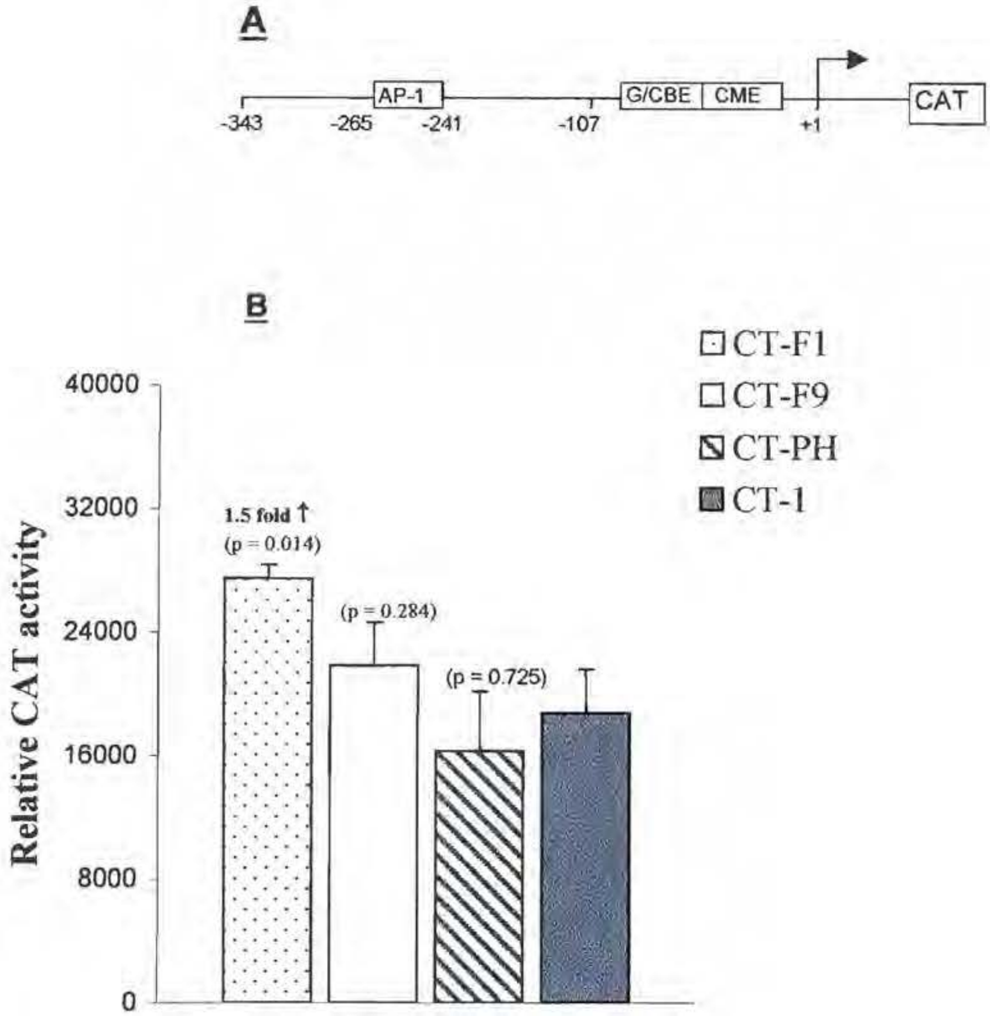


Figure 4.8: Proximal promoter activity of the α2(I) procollagen gene. Cells were seeded in 100 mm dishes, grown to 40% confluency and transfected with 20μg of the -343 COLCAT promoter construct using the calcium- phosphate precipitation method described in the section 6.2. 5μg of the cmvβgal plasmid was included in the transfection mixture to control for transfection effeciency. Cells were harvested after 48 hrs post-transfection and the CAT activity determined. The activity of the promoter is expressed as the relative CAT activity calculated from the ratio of CAT to βgalactosidase activity. The activity of the proximal promoter was significantly higher in CT-F1 fibroblasts relative to CT-1 fibroblasts, no significant differences in the activity of the proximal promoter were observed between CT-PH and CT-1 fibroblasts.These results represent the mean ± SD of four separate experiments performed in triplicate; p values were calculated using the student *t* test.

DNA-protein interactions at these promoter elements remained unchanged (Figure 4.9). DNA-protein interactions at the AP-1 like element were also assayed (Figure 4.10); it was difficult to obtain a very clear pattern of complex formation even after phenol:chloroform extraction of the labelled probe. However, competition with an AP-1 consensus sequence from the collagenase promoter as well as competition with the $\alpha2(I)$ procollagen AP-1 like sequence had no effect on DNA binding activity (data not shown). Competition was observed with an Sp1 oligonucleotide; there is one Sp1 binding site in the probe used in the electrophoretic mobility shift assay. These binding results further confirmed the absence of a Fos-response element in the proximal $\alpha2(I)$ procollagen promoter that could account for the observed down-regulation of the $\alpha2(I)$ procollagen gene.

4.2.6.2 Cloning and transcriptional activity of the -2300 bp $\alpha2(I)$ procollagen promoter

The proximal promoter (-343 COLCAT) was uninformative in terms of Fos response elements and it was therefore necessary to study further upstream and downstream elements of the gene. In order to carry out these assays it was necessary to clone the promoter elements into p8CAT. Figure 4.11 outlines the cloning strategy for the 2300 bp promoter fragment of the $\alpha2(I)$ procollagen gene in p8CAT. Briefly, ΔSph pNJE16 plasmid, an Sph digested fragment of a 4.0 kbp EcoRI 5′ clone of the $\alpha2(I)$ procollagen gene cloned into the pUC vector was digested with EcoRV and HindIII to release a 2300 bp fragment. The latter DNA insert was cloned into p8CAT cut with SmaI and HindIII to generate the -2300 COLCAT plasmid. Cells were transfected with this plasmid and the CAT activity assayed. The activity of this promoter construct was generally low in all cell lines in comparison to the activity of the proximal promoter. As can be seen from figure 4.12, there was a significant increase in the activity of the promoter, a 4- and 2-fold increase was observed for the CT-F1 and CT-F9 fibroblasts, respectively. From these results it would appear that the 2300 bp procollagen promoter fragment does not contain Fos response elements with a negative effect on the transcriptional activity of the gene.

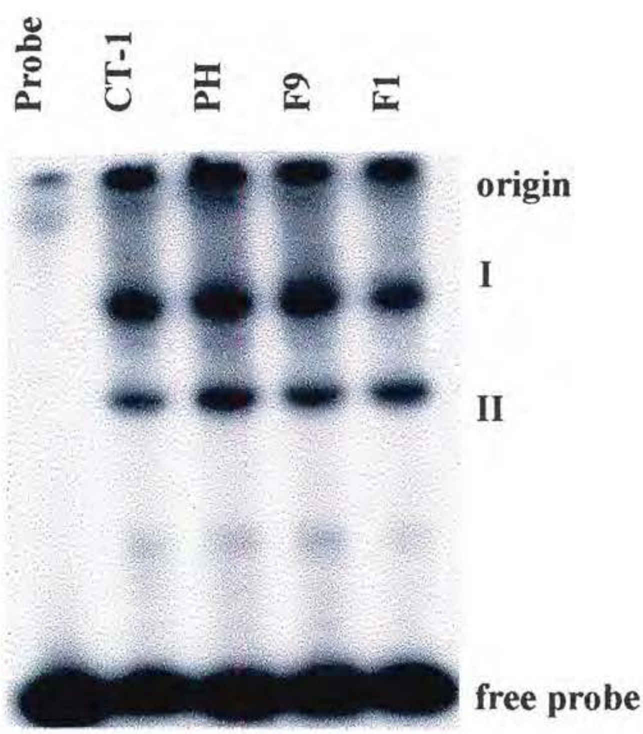


Figure 4.9: **DNA-protein interactions at the G/CBE-CME element of the α2(I) procollagen promoter.** Nuclear proteins were extracted from confluent fibroblast cultures as described in section 6.6.1 and incubated in the presence of a ^{32}P-labelled -107 to -60 promoter fragment that encompasses the G/CBE and CME regions of the α2(I) procollagen promoter (described in chapter 2). DNA-protein complexes were resolved by electrophoresis on a non-denaturing 5% polyacrylamide gel and the dried gel exposed to X-ray film for 16 hours at -70 °C. **I** and **II**; DNA-protein complexes.

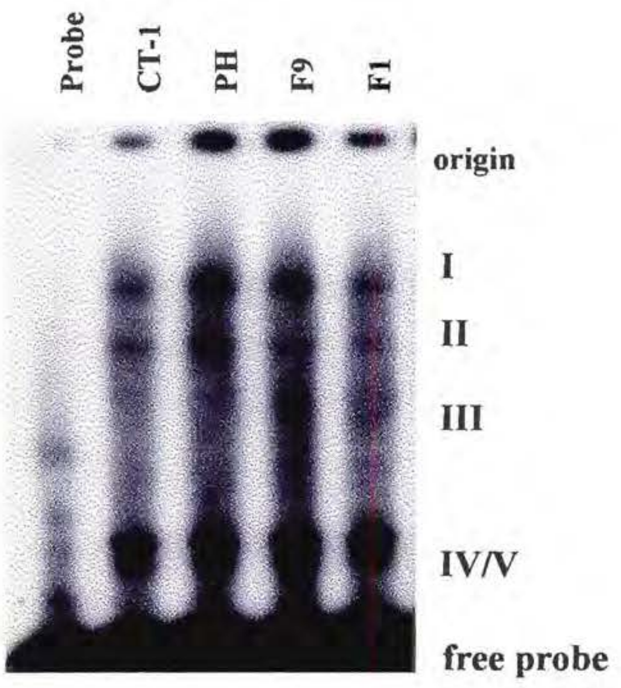


Figure 4.10: DNA binding activity of the -278 to -183 region of the α2(1) procollagen promoter. Nuclear proteins were prepared as described in section 6.6.1. 4μg protein was incubated with a ^{32}P-labelled -278 to -183 of the α2(I) procollagen promoter fragment encompassing the AP-1 like consensus sequence and an Sp1 consensus sequence. DNA-protein complexes were separated by electrophoresis on a non-denaturing 5% polyacrylamide gel. The gel was exposed to X-ray film for 16 hours at -70 °C. Complex I was competed by an Sp1 oligonucleotide.

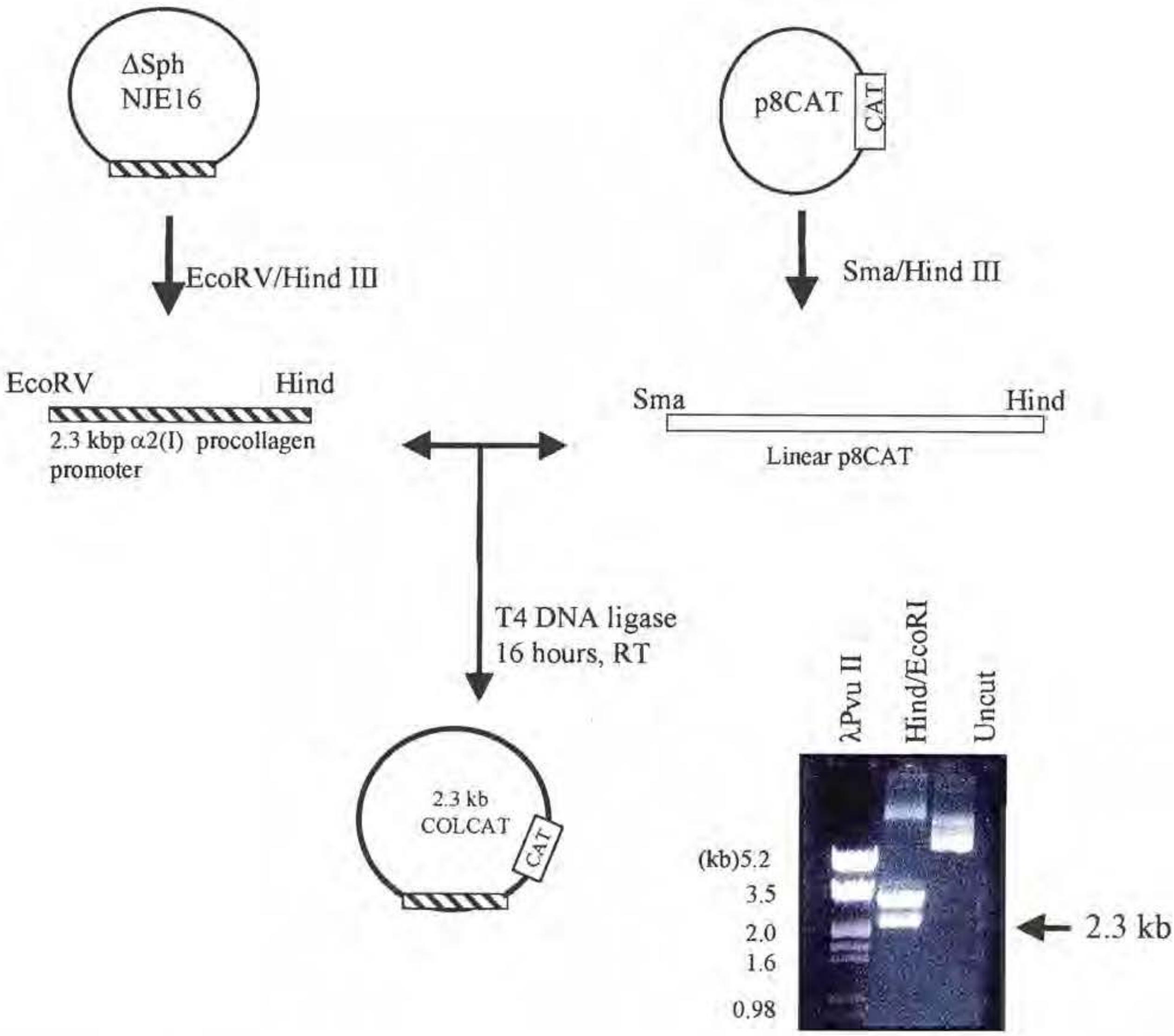
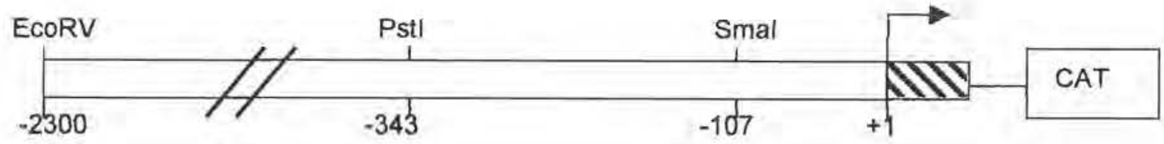


Figure 4.11: **Outline of the cloning strategy for the -2.3 kb α2(I) procollagen promoter fragment into p8CAT.** A 2.3 kb Hind/EcoRV fragment of the α2(I) procollagen promoter was released from ΔSph NJE16, (a genomic clone containing a 4.0 kb EcoRI fragment which harbours the first exon and approximately 3.8 kb of 5′ flanking sequences of the α2(I) procollagen gene digested with SphI to release 200 bp of the first exon sequences). The promoter fragment was subcloned into p8CAT reporter vector (see the Appendix for a map of p8CAT). The gel insert is a restriction enzyme digest of the 2.3 kb COLCAT promoter construct digested with Hind III and EcoRI.

A



B

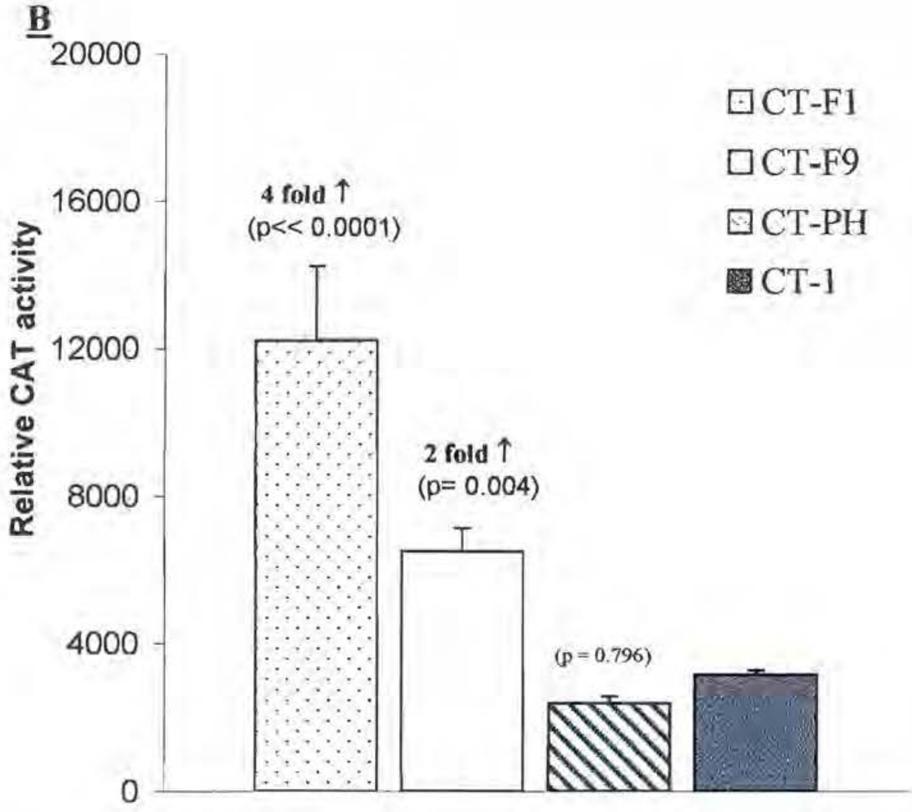


Figure 4.12: Transcriptional activity of the -2300 COLCAT construct. A) A schematic representation of the 2300 bp $\alpha2(I)$ procollagen promoter-CAT construct. **B)** 20 μg of plasmid DNA was co-transfected with 5 μg of cmvβgal plasmid using the calcium phosphate precipitation method (section 6.2). Transfected cells were harvested and CAT activity determined. The CAT activity of the promoter is expressed relative to βgalactosidase activity. The activity of the -2300 COLCAT construct increased significantly by 4- and 2-fold for the CT-F1 and CT-F9 fibroblasts, respectively. These results represent the mean $\pm$ SD of four separate experiments done in triplicate; p values were calculated using the student *t* test.

4.2.6.3 Transcriptional activity of the promoter/intron-COLCAT construct

Previous reports have demonstrated that the first intron of the human α2(I) procollagen gene contains negative regulatory elements while an enhancer element was identified in the mouse first intron of the same gene (61,63). The first intron of the α2(I) procollagen gene has been shown to contain a number of transcription factor binding sites, including two AP-1 consensus elements at nucleotides 1112 and 1085, but no specific function has been correlated with these sites (63). A sequence, which closely resembles the recognition site for the adenovirus replication protein, nuclear factor 1 is located between nucleotide +192 and +205, 5′ to the intron/exon junction. In addition, the first intron has been shown to reduce the activity of the α2(I) promoter activity 4-fold in transient transfection assays (63). The rationale was to therefore clone the first intron into a reporter plasmid and carry out transient transfection assays to establish whether this region of the α2(I) procollagen gene contains any sites which are required for the observed down-regulation of this gene in fibroblasts overexpressing Fos.

A pCOLα2(1.0 + 1.8 kb)bgh plasmid, described previously in the literature (63), which contains 1000 bp of sequences upstream of the transcription start site, 250 bp of exon 1, 1550 bp of the first intron, and 2100 bp of bovine growth hormone (bgh) sequence, cloned into pUC18, was digested with XbaI and EcoRV to release the 2800 bp insert of the α2(I) procollagen gene (Figure 4.13). The insert was treated with Klenow DNA polymerase to introduce blunt ends at all XbaI sites and subsequently cloned into pBluescript digested with EcoRV to generate pα2(I)I COLBS plasmid. The pα2(I)I COLBS plasmid was digested with EcoRI to determine the orientation of the different clones. Figure 4.13A illustrates the different fragments released after EcoRI digestion: the presence of the 2960, 1800 and 1000 bp fragments indicate a clone in the positive orientation. The appropriate plasmid clone was digested with BamHI and SalI to release the 2800 bp of α2(I) procollagen sequences and the insert subcloned into p8CAT digested with BamHI and SalI to generate the pα2(I)I COLCAT plasmid. To verify the size of the inserts, plasmid clones were digested with EcoRI to release the following fragments: 390, 1016, 1800, and 4110 bp (Figure 4.13B). The correct plasmid construct was confirmed by digesting plasmid DNA with BamHI and SalI to release the 2800 bp insert (Figure 4.13C) and by sequencing with a p8CAT reverse primer and the COL+1 primer (see Table 2.2). Plasmid DNA was transfected into the different Fos cell lines and the CAT activity was determined as described in section 6.2. Figure 4.14 illustrates the transfection results for the pα2(I)I COLCAT construct; it is evident that both Fos overexpressing cell lines have a higher promoter activity compared to CT-1 and CT-PH fibroblasts, however, the increase was not significant, a 1.6-fold increase ($p = 0.246$) was observed for the CT-F9 fibroblasts. These results indicate that AP-1 sites present in the first intron of the α2(I) procollagen gene do not play any role in the observed repression of the gene in cells overexpressing Fos. An interesting observation was the overall decrease in the transcriptional activity of the COLCAT construct containing the first intron in all cell lines, compare figures 4.12 and 4.14; these results are consistent with observations of Sherwood *et al* (63).

The use of transient transfection experiments is used to map transcription factor binding sites that play a role in the regulation of gene expression. In contrast, run-off transcription assays are commonly used to determine the rate at which a specific gene is transcribed and therefore do not provide an indication of the *cis*-acting elements that determine transcription of the gene under study. The disadvantage of using the transient transfection system is that different batches and purity of the cloned promoter constructs might introduce variability in the results, thus careful controls should be selected to correct for shortcomings of the assay. Regardless, the advantage of transient transfection assays is mainly the ability to effectively map transcription factor binding sites for potential activators and repressors.

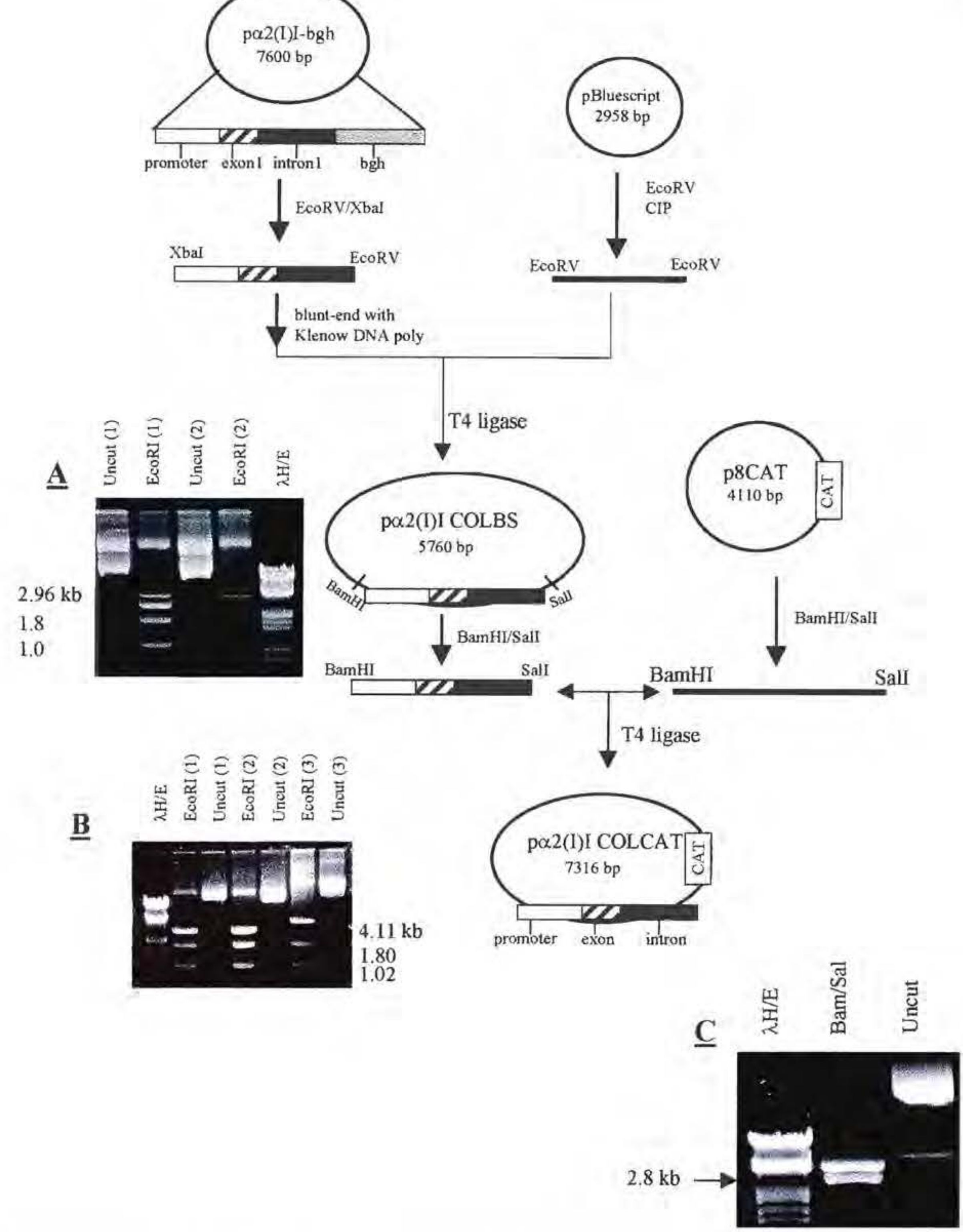


Figure 4.13: **Cloning of the α2(1) procollagen promoter/intron.** The pα2(I)I-bgh vector containing 1kb of the α2(I) procollagen promoter, 0.25kb of exon I, 1.8kb of intron 1, and 2.1kb of the bovine growth hormone (bgh) gene, was digested with XbaI and EcoRV to release a 2.8kb α2(I) procollagen fragment. The DNA was blunt-ended with Klenow DNA polymerase and ligated into pBluescript to generate pα2(I)I COLBS; the correct orientation of the cloned insert was confirmed by EcoRI restriction digest to release 2.96, 1.8 and 1.0 kb fragments **(A)** pα2(I)I COLBS was digested with BamHI and SalI and the insert was subcloned into p8CAT to generate pα2(I)I COLCAT; the correct clone was confirmed by restriction enzyme digestion with EcoRI to release 4.11, 1.80, 1.02, and 0.39 kb fragments **(B)** or BamHI and SalI to release an insert of 2.8 kb **(C)**.

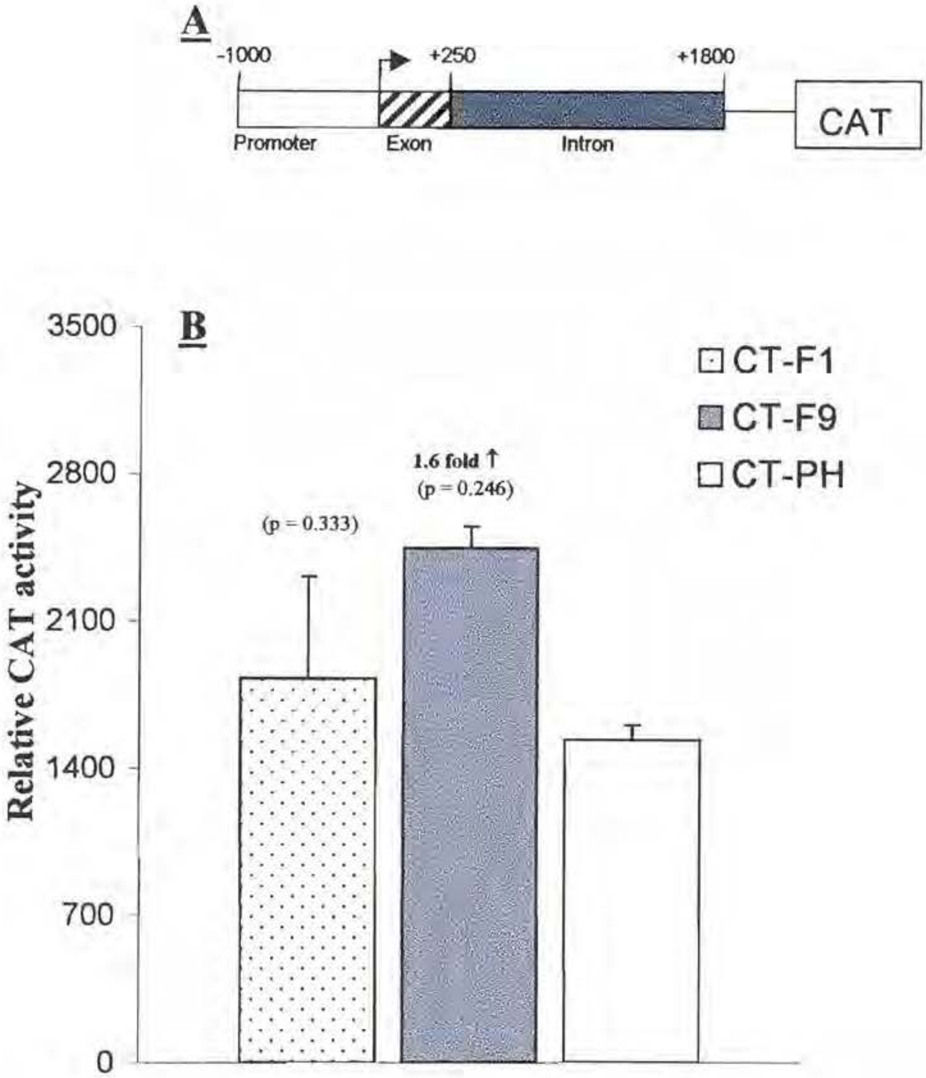


Figure 4.14: Transcriptional activity of the pα2(I)I COLCAT construct. (A) The pα2(I)I COLCAT plasmid was constructed as described in figure 4.12; the plasmid contains 1.0, 0.25, and 1.8 kb of promoter, exon and intron sequences, respectively. **(B)** 20μg of plasmid DNA was co-transfected with the cmvβgal plasmid into fibroblasts using the calcium phosphate precipitation method (see section 6.2). Cells were harvested and CAT activity determined. The CAT activity of the promoter is expressed relative βgalactosidase activity. The activity of the α2(I)I COLCAT construct was 1.6 fold high for CT-F9 fibroblasts as compared to the CT-PH fibroblasts. These results represent the mean ± SD of four separate experiments done in triplicate; p values were calculated using the student *t* test.

4.3 DISCUSSION

The extracellular matrix (ECM) is engaged in the transmission of a large variety of specific signals that directly influence growth, migration and differentiation of cells (225). In addition, the ECM can influence cellular behaviour through changes in the three-dimensional organisation of the cytoskeleton and activation of second messenger and protein kinase pathways (225). Some of the known oncogenes and tumour suppressor genes have been shown to interfere with intracellular pathways conveying signals arising from the ECM (225). Type I collagen is one of the most abundant proteins of the ECM, and its expression is altered after cellular transformation. Surprisingly, transformation of fibroblasts by γ-irradiation, as is the case in CT-1 fibroblasts, had very little effect on the expression of both the $\alpha2$(I) and $\alpha1$(I) procollagen genes (13,33). Radiation is known to induce transcription of a number of genes including the *c-fos* gene, however, in CT-1 fibroblasts, expression of *c-fos* was markedly reduced (89,117,224). Previous studies in these cells demonstrated that treatment of CT-1 fibroblasts with serum or phorbol esters failed to induce *c-fos* mRNA levels (89). Subsequent analysis of the promoter showed no changes in the regulatory elements and the activity of a transfected promoter was similar in parental and transformed fibroblasts (89).

The question asked in this study was whether failure to inhibit $\alpha2$(I) procollagen gene expression in CT-1 cells can be correlated with repressed levels of *c-fos*. To address this issue, a constitutively active *c-fos* plasmid was introduced into the CT-1 fibroblast line and examined for changes in the expression of the $\alpha2$(I) procollagen gene. Consistent with a report by Mercier *et al* (226) whereby overexpression of *c-fos* in rat epithelial cells had no effect on the doubling time of cells, in this study, constitutive expression of *c-fos* did not alter the doubling time of CT-1 fibroblasts. However, in a separate report, constitutive expression of a *c-fos* transcript in which the 3′ untranslated region responsible for its rapid turnover has been deleted, activated the transforming activity of *c-fos* with a consequent increase in the growth of the transformed phenotype (122,227). A study by Grigoriadis *et al* (142) also demonstrated that the expression of *c-fos* from an H-2K class I major histocompatibility complex (MHC) promoter causes osteosarcomas in transgenic mice; high levels of Fos were shown to perturb the normal growth control of osteoblastic cells with bone-forming capacity leading to their transformation. Thus depending on the cell type, overexpression of *c-fos* will have diverse effects on the growth rate of cells.

Analysis of the $\alpha2(I)$ procollagen gene transcripts clearly demonstrated a significant reduction by ~50 % in the Fos expressing clones. This result is consistent with the findings of Kuroki *et al* (173) that constitutive expression of *c-fos* inhibits type I collagen synthesis in osteoblasts. Transformation of Rat-1 fibroblasts by the *v-fos* oncogene has been shown to decrease the levels of both $\alpha1(I)$ and $\alpha2(I)$ procollagen mRNA; this effect could be reversed with the use of revertant cell lines that expressed a functional *c-fos* oncoprotein (22). Other related studies have demonstrated that overexpression of mutant *ras* downregulated both the $\alpha1(I)$ and $\alpha2(I)$ procollagen mRNA levels (21), which was correlated with increased DNA binding activity at the intronic AP-1 site (64). These results suggested a possible link between the effects of *ras* mutations and type I procollagen gene expression in fibroblasts, and further demonstrated that *c-fos* is a downstream mediator of *ras* in the signal transduction pathway. Interestingly, the *c-Ki-ras* gene has been shown to be amplified in CT-1 fibroblasts thus accounting for elevated expression of the gene in these cells. Thus it is possible that extracellular signals transduced downstream of the *ras* pathway fail to induce *c-fos* expression and this effect translates into very little change in the expression of the $\alpha2(I)$ procollagen gene in CT-1 fibroblasts. The observed decrease in the levels of $\alpha2(I)$ procollagen mRNA in Fos expressing clones support the suggestion that failure of CT-1 fibroblasts to show significant changes in the $\alpha2(I)$ procollagen mRNA and protein is related to failure of CT-1 fibroblasts to trigger the expression of *c-fos*. Treatment of mouse NIH 3T3 cells with the tumour promoter, 12-O-tetradecanoylphorbol-13-acetate (TPA) also resulted in a decrease in the transcription of $\alpha2(I)$ procollagen gene (104). Sobel *et al* (103) reported that treatment of JB-6 epidermal cells with TPA results in inhibition of type 1 collagen synthesis at a pretranslational level. TPA is known to activate protein kinase C, an upstream mediator of the signal leading to increased expression of *c-fos* (102,117). These results are all in agreement with our observations on the effect of Fos on α2(I) procollagen gene expression in that high levels of *c-fos* are accompanied by a decrease in expression of type I procollagen genes.

The regulation of gene expression is a multi-step process, it can be effected at different stages: (a) alteration of gene structure by changing the conformation of chromatin and demethylation of DNA, (b) transcription factor binding to specific *cis*-elements for participation in the formation of the pre-initiation complex, (c) RNA polymerase activity, (d) nuclear RNA processing, (e) mRNA stability, and (f) mRNA transport to and translation in the cytoplasm. Results obtained in this study demonstrated a decrease in the expression of the α2(I) procollagen gene in fibroblasts overexpressing Fos. To establish how this is accomplished, the study examined possible levels of regulation, namely, mRNA stability and promoter activity.

Although type I procollagen genes generally have a fairly long half-life, ~18 hours, agents that affect the expression of this gene can do so by destabilising the transcript via activation of a RNA binding protein which is responsible for regulating the turnover of mRNA (228-230). The stability of α2(I) procollagen message was determined by incubating cells with actinomycin D and total RNA extracted at different time points. This approach is often utilised to study the stability of mRNAs with much shorter half-lives because of the general toxicity of actinomycin D. It is nevertheless still widely used since it generates the same data as other methods recommended for messages with longer half-lives, e.g. classic pulse-chase experiments. Furthermore, it gives an indication of any changes in message stability, of which the rate can be determined using more accurate protocols. Treatment of fibroblasts with actinomycin D caused an initial increase in the α2(I) procollagen mRNA levels within the first 8 hours, after which levels started to show a decline, however, the trend was similar for all fibroblast lines analysed. After 16 and 24 hours there was a marked decrease in all mRNAs analysed including the 28S and 18S RNA, this decline coincided with a decline in the viability of cells. It is therefore evident that constitutive expression of *c-fos* did not downregulate the α2(I) procollagen mRNA levels by increasing its rate of turnover. Slack *et al* (21) have demonstrated that overexpression of wild type *ras* had no effect on the levels of both α1(I) and α2(I) procollagen mRNA, whilst mutant *ras* decreased the expression of both genes as well as destabilised the α1(I) procollagen message in Rat-1 fibroblasts. Although *c-fos* is a downstream mediator of *ras*, it appears that overexpression of both genes results in much more complex changes in the cell.

Since the stability of the α2(I) procollagen mRNA was not affected by overexpression of Fos, it seemed possible that, perhaps, a change in transcriptional activity could account for the decreased expression of this gene. The α2(I) procollagen proximal promoter contains a number of transcription factor binding sites and this promoter fragment has been shown to be sufficient to direct cell-specific expression of this gene (35). In contrast to the observed levels of the α2(I) procollagen mRNA in the Fos clones, the activity of the promoter remained unchanged in CT-F9 and marginally increased in the CT-F1 clone. Thus implying that although this promoter segment contains a putative AP-1 site and two repressor binding sites, none of these regions play a role in the Fos response. The promoter was transactivated but the DNA binding activity of transcription factors that recognise the AP-1 site and the previously described CME, revealed no differences in DNA-protein interactions. It was therefore apparent that the proximal promoter lacked Fos response elements, suggesting that the response elements could be located elsewhere in the gene. No negative Fos response elements were identified in the 2300 bp promoter fragment, however, a significant increase in the transcriptional activity of this promoter fragment was observed. The implication of these results are that the 2300 bp promoter possibly contains more AP-1 sites or *cis*-elements which have a positive effect on the transcription of the α2(I) procollagen gene in the presence of high levels of Fos. To date, no specific Fos response elements have been delineated in the 2300 bp region of the human α2(I) procollagen gene promoter.

Sherwood *et al* (63) have reported that the activity of the human α2(I) procollagen promoter is reduced by inclusion of the first intron in transient transfection experiments. As mentioned earlier, the first intron of the human α2(I) procollagen gene contains a number of regulatory elements, *inter alia,* two AP-1 consensus sequences, a NF-1 like sequence and an alternating GT stretch (63). Alternating dT and dG residues are known to form a Z-DNA conformation; Z-DNA is found in association with a number of known enhancer and regulatory sequences, thus it has the potential to function in transcriptional regulation (231,232). In order to map the Fos response elements in the first intron of the human α2(I) procollagen gene, the intron plus a 1000 bp promoter fragment were cloned into a reporter plasmid and transfected into stably transfected Fos clones.

The presence of the first intron did not result in a decrease in the transcriptional activity of the promoter in CT-1 and CT-PH fibroblasts in comparison to Fos overexpressing clones. These results indicate that constitutive expression of *c-fos* has a positive effect on the transcriptional activity of the α2(I) procollagen promoter in the region between -2300 and +1800. The observed down-regulation of the α2(I) procollagen gene expression by constitutively active *c-fos* is possibly an indirect effect in that *c-fos* might initiate changes in the genetic programme and some of these changes will result in an overall decrease in the level of the α2(I) procollagen mRNA. In tight skin (TSK) mice myocardial fibroblasts, however, there was an increase in the expression of α1(I) procollagen gene and the increase was attributed to a reduction in the interaction of a negative regulatory sequence located between -675 and -804 with the AP-1 transcription factor, implying that the transcription factor AP-1 is involved in transcriptional repression (234). In a recent report, Bakin and Curran (233) demonstrated that the transformation of normal 208F rat fibroblasts by constitutive expression of *c-fos* was a result of increased expression and activity of the DNA methylcytosine transferase (*dnmt-1*) enzyme, where it was shown that the overexpression of *dnmt-1* in the absence of *c-fos* resulted in transformation of cells. These observations indicate that although *c-fos* increased the expression and activity of the DNA methylcytosine transferase, it did not directly contribute to cell transformation.

These results lend further support to the suggestion that downregulation of the $\alpha2(I)$ procollagen gene in fibroblasts overexpressing Fos is not a direct effect on the procollagen gene, at least at the level of transcription of the gene and the stability of the mRNA. Furthermore, since most signals arising from the cell membrane trigger the expression of immediate early genes such as *c-fos* and *c-jun*, a number of other genes could be affected by overexpression of Fos and because Fos is present in abundance, this might lead to formation of more diverse and stable heterodimers with the different Jun family of proteins. Chung *et al* (42) have reported that constitutive expression of *c-jun* down-regulates the transcriptional activity of the $\alpha2(I)$ procollagen promoter in fibroblasts. The transcriptional activity of the human androgen receptor (hAR) has been shown to depend on *c*-Jun; *c*-Jun has a selective function of mediating AR dependent transactivation, however, dimerisation of *c*-Jun with *c*-Fos represses this effect (235). Interestingly, *c*-Fos also antagonises the *junD* gene positive autoregulatory loop thereby inhibiting the promoter activity of the *junD* gene (236). The expression of *junD* has been shown to be constitutive in quiescent cells and therefore inhibition of its expression by *c-fos* upregulates the expression of genes required in cell proliferation (236). These results, taken together, seem to point out the significance of the different Fos:Jun heterodimers in transcriptional regulation of a variety of genes. It is therefore possible that the constitutive expression of *c-fos* in CT-1 fibroblasts alters the composition of Fos:Jun heterodimers in favour of repression of the $\alpha2(I)$ procollagen gene, among other things. In addition, Hai and Curran (237) have shown that members of the ATF/CREB family form selective cross-family heterodimers with members of the Fos/Jun family and these complexes can bind to either CRE or AP-1 consensus sequences in gene promoters, although they exhibit a preference for the CRE site. This cross-family dimerisation is complicated by the observation that Fos or Jun further cross-couple with NF-κB/p65 to modulate gene expression at sites other than the AP-1 or NF-κB sites (100,101).

At this stage the conclusion that can be drawn from these results is that one of the many effects of overexpression of Fos is the down-regulation of the $\alpha2(I)$ procollagen gene, this effect was shown not to be a result of a high turnover of the $\alpha2(I)$ procollagen message or a decrease in the transcriptional activity of the gene ruling out increased transrepression via AP-1 binding sites.

The primary objective of this study was to correlate expression levels of *c-fos* with the extent of $\alpha2(I)$ procollagen expression in CT-1 fibroblasts. From the results obtained it is apparent that failure of cells to respond to signals that trigger the expression of *c-fos* accounts for the high $\alpha2(I)$ collagen levels in these transformed cells. From the literature, CT-1 fibroblasts have been shown to express high levels of *K-ras*, therefore the inability to trigger Fos arises from deficiencies in the signal cascade downstream of Ras and by-passing the signal transduction pathway, it was possible to observe changes in the expression levels of $\alpha2(I)$ procollagen. Fos, however, is central in the signal transduction pathway and overexpression of a constitutively active Fos will shift the balance in the concentration and composition of Fos:Jun heterodimers resulting in the induction and /or inhibition of a number of genes that will modulate the levels of $\alpha2(I)$ procollagen. At this stage it has become very clear that Fos does not have a direct effect on the activity of transcription factors that bind to the promoter and intron, between -2300 and +1800, of the $\alpha2(I)$ procollagen gene. It would be interesting to examine events in the signal cascade downstream of Ras that lead to changes in the levels of Fos and ultimately down-regulation of the $\alpha2(I)$ procollagen gene.

CHAPTER 5

CONCLUSION

5.1 Background information

Previous studies have reported on different *trans*-acting factors and their cognate *cis*-acting elements that participate in the regulation of the 2(I) procollagen gene (see chapter 1). A number of transcription factors have been identified, purified and/or cloned, amongst these are Sp1, BFCOL, CBF and IF-1/2. An interesting finding from these reports was that, although the human and mouse α2(I) procollagen promoters share some sequence identity, certain factors such as NF-1 and IF-1/2, have been shown to bind the mouse promoter only. Sequence identities do not extend to the first intron, which has also been shown to contain regulatory elements.

The expression of the α2(I) procollagen gene is significantly affected by transformation with SV40 but not by γ-irradiation (13,33,57,59). These differences were a subject of further analysis in this study. Parker *et al* (13) have suggested that these differences in the expression of the α2(I) procollagen gene arise from differences in *trans*-acting factors that bind between the -107 and -60 region of the α2(I) procollagen promoter.

5.2 Objectives of this study

- To demonstrate that transcription factors that form complex I and II at the basal -107 α2(I) procollagen promoter have a functional role in the regulation of this gene;
- To purify and identify the transcription factor(s) that forms complex II on the CME in the α2(I) procollagen promoter.;
- To study the regulation of the α2(I) procollagen gene by *c-fos* in CT-1 transformed fibroblasts.

5.3 Main findings

Earlier studies have suggested the differential regulation of the α2(I) procollagen gene in different cell lines was attained by *trans*-acting factor switch at the basal promoter (13,59). Two or three major complexes were shown to form when nuclear proteins extracted from either collagen or non-collagen producing cells were incubated with the basal α2(I) procollagen promoter. In this study, however, it was demonstrated that only two major complexes form on this promoter fragment. The third complex is a specific cleavage product of complex II proteins. Protease treatment of nuclear proteins or DNA-protein complexes generates degradation products that still retain DNA binding activity or the DNA-protein complex is not destroyed. Complex III proteins, were observed in nuclear extracts prepared in the absence of a cocktail of protease inhibitors or in the absence of leupeptin in the extraction buffers. The functional significance of this protease resistant complex remains to be studied. The important conclusion was that both complex I and complex II proteins are ubiquitous.

To study the functional role of complex I and complex II proteins, site-directed mutagenesis was employed in order to introduce point mutations, in the G/CBE and CME elements of the α2(I) procollagen promoter and to examine changes in promoter activity. G/CBE mutations resulted in a significant decrease in promoter activity; both mutations in the inverted CCAAT box and the 5′ adjacent, inverted GGAGG box almost totally abolished promoter activity. These results were observed in both CT-1 fibroblasts, a collagen producing cell line, and SVWI-38 fibroblasts, a non-collagen producing cell line. The decrease in promoter activity was observed in experiments carried out with both the proximal or basal promoter constructs. Furthermore, the study demonstrated that mutations in the G/CBE interfered with formation of DNA-protein interactions *in vitro*. It can be concluded that the CCAAT box and adjacent sequences that constitute the G/CBE element are required for the activation and regulation of expression of the α2(I) procollagen gene. It is possible that factors that bind at this element are crucial for the formation of the transcription pre-initiation complex through interactions with the general transcription machinery.

The G/CBE element has been shown previously to bind a protein that belongs to the CBF family of transcription factors, the focus of this study was therefore shifted towards the CME. Although the CME is absent in the mouse promoter, proteins that bind to this element are present in rodent cell lines, and as mentioned earlier, these proteins are ubiquitous. Mutations were introduced in the CME in order to study the effects of these mutations on the activity of the $\alpha2(I)$ procollagen promoter. A significant increase in promoter activity was observed when the mutated promoter constructs were transfected into CT-1 and SVWI-38 fibroblasts, with a 3-fold increase in the activity of the promoter observed in SVWI-38 fibroblasts. These results suggested that the CME binds a transcription factor that could be involved in the repression of the $\alpha2(I)$ procollagen gene. *In vitro* DNA binding assays did not suggest an increased binding of nuclear proteins in the presence of mutations in the CME, ruling out the possibility these mutations could increase the binding activity of a more potent activator and hence the increase in the activity of the promoter constructs. In addition, the increase was more significant in SVWI-38 fibroblasts as compared to CT-1 fibroblasts. An interesting finding was that the increased promoter activity of the mutated constructs was restricted to the proximal promoter and not the basal promoter, the latter constructs showed a decrease in the activity of the promoter. The explanation for this contradiction could simply be that the repressor requires interaction with upstream elements or additional factors that bind to upstream elements. It can be deduced that the CME binding protein is a repressor in agreement with previous suggestions. At this stage, the correlation that can be made between the above findings and the degradation of complex II to III is the possibility that CME binding proteins that are not in association with DNA might be susceptible to protease degradation the major product of which is still able to bind DNA. Fibroblasts will therefore either express or not express $\alpha2(I)$ procollagen depending on the extent of degradation/processing of complex II to III. An interesting observation was the apparent low ratio of complex II:III in CT-1 extracts compared to SVWI-38 extracts; this raises the possibility of existence of differences in the stability and perhaps the structure of the CME binding protein in the two cell lines. Since complex II proteins are also present in non-procollagen expressing cell lines, it is a clear indication of the complex nature of the regulation of procollagen gene expression. The regulation of gene expression is achieved at different levels such as activation of the gene structure, access and interaction of transcription factors with their recognition elements, and post-translational modifications of transcription factors, therefore CME binding proteins might require modifications in order for them to modulate expression of the $\alpha2(I)$ procollagen gene.

The functional significance of a specific *cis*-acting element is usually deduced from *in vivo* experiments using a reporter system and transfection. In order to establish whether a specific transcription factor that binds to this element is indeed an activator or repressor, it is important to purify and/or clone it and carry out *in vitro* transcription assays. In this study, the complex II protein was purified using conventional biochemical techniques. The protein was further purified by SDS-polyacrylamide gel electrophoresis. Two polypeptides of 67 and 50 kDa were isolated after two steps of DNA affinity purification. The polypeptides were eluted and subjected to either automated amino acid sequencing or MALDITOF-ms analysis. Sequencing was not successful since the polypeptides appear to be N-terminally blocked. Since this procedure requires relatively large amounts of material, it was not pursued any further. Polypeptides digested with trypsin were therefore analysed using MALDITOF-ms and the *m/z* ratios were used to search the databank for any matches. The only close match obtained with four peptide fragments was of a mouse zinc finger protein ZFP-27 (MKR4 protein). Collins *et al* (56) have reported that the DNA-binding activity of complex II proteins was zinc dependent. It can be concluded that the complex II protein belongs to the zinc family of transcription factors. Work is in progress to obtain the complete sequence of the protein to further characterise its DNA binding and activation domains.

The expression of a number of cellular proteins is altered as a result of transformation. The most common feature of transformation of fibroblasts, the primary type I collagen-producing cell, is a reduction in the expression of type I procollagen. Numerous reports have demonstrated the effect of viral oncogene products on the expression of type I procollagen genes. In this study, a human embryonic lung fibroblast cell line transformed by γ-irradiation (CT-1 fibroblasts), was used as a model to study the effect of transformation on the expression of the α2(I) procollagen gene. CT-1 fibroblasts have been shown to express normal levels of α2(I) procollagen gene, however, these cells have lost the ability to express *c-fos*, even after stimulation with positive effectors such as serum and phorbol esters (89). This study reports on the role of constitutive expression of *c-fos* in the modulation of α2(I) procollagen gene expression. CT-1 fibroblasts transfected with a *c-fos* cDNA construct that allows constitutive expression of *c-fos* show downregulated α2(1) procollagen mRNA levels. No change in the stability of the message or changes in the rate of message turnover were observed; the turnover of the message remained similar to that of normal cells.

In an attempt to characterise the Fos response in CT-1 fibroblasts, different regions of the promoter and the first intron were analysed for the presence of Fos-response elements. From the results it is evident that the 2300 bp of sequences 5′ of the transcription start site lack negative Fos response elements, instead the activity of this promoter construct was transactivated in the two Fos overexpressing cells lines. Subsequent analysis of the promoter activity of the α2(I) procollagen promoter construct containing sequences between -1000 and +2000 bp (including exon 1 and intron 1) did not reveal the presence of any Fos response elements that had a negative effect on promoter activity. Interestingly, the -2300 and -343 promoter constructs were transactivated in *c-fos* transfected clones and the promoter-intron construct was only transactivated in the CT-F9 clone. From these results, it can be concluded that AP-1 elements present in the proximal promoter and the first intron are important in transactivation of the α2(I) procollagen gene. Whilst the Fos response remains an enigma, the repressed α2(I) procollagen mRNA levels in fibroblasts overexpressing *c-fos* appears to be an indirect effect, at least in as far as transcription from the 2300 bp promoter is concerned. In a recent report, the overexpression of Fos was shown to result in increased expression and activity of DNA methylcytosine transferase accounting for the observed transformed phenotype in rat 208F fibroblasts. Since *c-fos* is at the end of the signal transduction pathway, it is possible that overexpression of this gene is likely to influence the expression of a variety of other genes that function to negatively regulate the expression of α2(I) procollagen gene. It is well documented that Fos modulate genes through formation of heterodimers with the Jun family of proteins, and depending on the composition of the dimer, the effect will be either an increase or a decrease in the expression of a specific gene. Diamond *et al* (223) have shown that a single *cis*-element can cause opposite effects on transcription; the glucocorticoid response element was inactive in the absence of *c-jun*, whereas a positive glucocorticoid response was observed in the presence of *c-jun*, and a negative effect was obtained in the presence of *c-jun* and relatively high levels of *c-fos*.

5.4 Significance of findings

The normal expression of the α2(I) procollagen gene relies on the proper interactions between transcription factors and regulatory elements of the gene. Changes in the secondary structure of the gene as well as the structure and stability of transcription factors will therefore play a significant role in the levels and the rate of expression of the α2(I) procollagen gene. Events such as DNA methylation have been associated with alteration in the gene expression pattern in cells, however, some of the transcription factors such as Sp1 can interact with either methylated and demethylated DNA. This study sought to evaluate the role of two specific *cis*-regulatory elements present in the basal promoter, 107 bp upstream of the transcription start site, of the α2(I) procollagen gene.

The study established the importance of the GGAGG/CCAAT (G/CBE) box in transcriptional activation of the human α2(I) procollagen gene. These observations are consistent with previous studies carried out with the mouse gene; the role of this element extents beyond cells that produce type I collagen. Furthermore, the G/CBE appears to function in the absence of upstream elements implying that the cognate factor possibly form protein-protein interactions with the basal transcription machinery, but transcriptional activation is reduced. The implications of these results are that the presence of mutations in the G/CBE region of the promoter will result in loss of transcriptional activity of the $\alpha2(I)$ procollagen gene. On the other hand, the downstream CME element appears to bind a factor that has a negative effect on transcriptional activity of the $\alpha2(I)$ procollagen gene. This inhibition appeared dominant in a non-$\alpha2(I)$ collagen producing cell line. In addition, the factor that recognises the CME has been found to be ubiquitous, it is present in a number of cell lines ranging from normal and transformed fibroblasts, haematopoietic cells, and epithelial cells. Inhibition of binding of this factor to the CME by mutagenesis increased transcription of the $\alpha2(I)$ procollagen gene only in the presence of upstream elements. The CME binding protein might function in transcriptional repression, but it is not a potent repressor. The presence of a transcriptional repressor of the $\alpha2(I)$ procollagen gene in a variety of cell lines can be exploited in the design of drugs for the treatment of collagen and collagen-related diseases, obviously depending on whether the protein is under- or overexpressed. A drug designed to inhibit the CME binding protein would result in moderate activation of $\alpha2(I)$ procollagen gene expression, an application that would be beneficial in the management of scleroderma characterised by overproduction of the $\alpha1(I)$ procollagen chain; such an approach would artificially maintain the 2:1 stoichiometry for a functional type I collagen protein. More studies, however, need to be conducted to further characterise the CME binding protein.

CHAPTER 6

MATERIALS AND METHODS

6.1 Cell culture

The cell lines used in this study were:

Human embryonic lung fibroblasts, WI38 (ATCC CCL-75)

WI-38 fibroblasts transformed by γ-irradiation, CT-1 (32)

WI-38 fibroblasts transformed with SV40, SVWI-38 (238)

Human monocytic cell line, U937 (ATCC CRL 1593)

6.1.1 Adherent cell culture

Adherent cells (all the fibroblast cell lines) were cultured in 35, 100, or 150mm tissue culture dishes in DMEM supplemented with 10 % (v/v) foetal bovine serum (FBS), 100 μg/ml streptomycin and 250 U/ml penicillin, in a humidified 95 % air/5 % CO_2 incubator (Queue) at 37 °C. Cells were grown to 90 % confluency and trypsinised in 0.05 % (w/v) trypsin in PBS containing 10 mM EDTA. Transformed cells were split 1:5 whilst normal fibroblasts were split 1:3 for subculturing.

6.1.2 Suspension cell culture

U937 cells were cultured in 75 or 150cm tissue culture flasks in RPMI-1640 medium supplemented with 10 % (v/v) FBS, 100 μg/ml streptomycin and 250 U/ml penicillin in a humidified 95 % air / 5 % CO_2 incubator, at 37 °C. Cells were split 1:10 at a cell density of 10^6 cells/ml for propagation of cultures.

6.1.3 Culture of serum deprived cells

Cells were trypsinised, split according to experimental conditions and cultured in DMEM supplemented with 0.1 % (v/v) FBS and antibiotics. After 48 hours, cells were incubated in fresh medium supplemented with 10 % (v/v) FBS for 30 minutes, harvested and used for either protein or RNA extraction.

U937 cells were cultured in 75 or 150cm tissue culture flasks in RPMI-1640 medium supplemented with 10 % (v/v) FBS, 100 μg/ml streptomycin and 250 U/ml penicillin in a humidified 95 % air / 5 % CO_2 incubator, at 37 °C. Cells were split 1:10 at a cell density of 10^6 cells/ml for propagation of cultures.

6.1.3 Culture of serum deprived cells

Cells were trypsinised, split according to experimental conditions and cultured in DMEM supplemented with 0.1 % (v/v) FBS and antibiotics. After 48 hours, cells were incubated in fresh medium supplemented with 10 % (v/v) FBS for 30 minutes, harvested and used for either protein or RNA extraction.

6.1.4 Treatment of cells with actinomycin D

Cells at 70-80 % confluency were treated with 5μg/ml actinomycin D for 0, 2, 4, 8, 16 and 24 hours and incubated in a humidified incubator at 37 °C. At the indicated time points, cells were harvested for total RNA extraction.

6.2 Cell transfection

In all transfection experiments, the calcium-phosphate precipitation protocol (239) was used. The following mixture was prepared for a 100 mm tissue culture dish:-

Solution	Stable	Transient
DNA construct[1]	10 μl (20 μg)	10 μl (10-20 μg)
Control plasmid[2]	-	5 μl (2-5 μg)
2M $CaCl_2$ (section 6.7)	60 μl	60 μl
1x TTE (section 6.7)	350 μl	345 μl
2x HBS[3] (section 6.7)	420 μl	420 μl

1~ plasmid DNA used depended on whether the transfection was stable (a cDNA expression vector was transfected) or transient (a promoter-reporter construct was transfected);
2~ the plasmid cmvβgal was co-transfected in all transient transfection experiments as a control for transfection efficiency;
3~ this solution was added dropwise with gentle agitation to a mixture containing the other listed components. Preparation of all solutions is described in section 6.7.

The mixture was incubated at room temperature for 30 minutes and should turn opaque implying successful formation of the DNA/calcium phosphate precipitate.

G418 . Resistant colonies were trypsinised and subcultured into 35mm tissue culture dishes for propagation.

6.2.2 Transient transfection of fibroblasts

Fibroblasts grown to 40 % confluency were transfected with 20 μg plasmid DNA using the calcium-phosphate precipitation method as described in section 6.2. Sixteen hours after transfection the medium was changed and the cells allowed to grow for 48 before being harvested for the assays as described in sections 6.2.3, 6.2.4, 6.2.5 and 6.2.6.

6.2.3 Extraction of protein from transfected cells

The medium was removed from transfected cells and the cells washed once with phosphate buffered saline (PBS). Cells were scraped with a cell scraper and pelleted by centrifugation at 13 000 rpm at 4 °C in a Biofuge 13R microcentrifuge (Heraeus, Sepatech) for 1 minute. The pellet was resuspended in 150 μl of 0.25 M Tris-HCl, pH 8.0 and lysed by three cycles of freeze-thawing in liquid nitrogen with a rapid vortexing step after the thaw cycle. The cytoplasmic fraction was obtained after centrifugation in a microcentrifuge at 13 000 rpm at 4 °C for 5 minutes. The post-nuclear supernatant was retained for enzyme assays.

6.2.4 BioRad protein assay

The protein concentration was determined using the standard BioRad assay, a modified Lowry colorimetric assay for protein determination (240). Briefly, extracts were mixed with the BioRad solution and incubated at room temperature for 10 minutes. The absorbance was measured at λ_{595nm} (Spectrophotometer PM2DL, Zeizz) and the protein concentration was determined from a standard curve performed with bovine serum albumin.

6.2.5 β-galactosidase assay

The reaction mixture for the β-galactosidase assay was as follows:-

Cell extract	50-100 μg protein

O-nitrophenyl-β-D-galactosidase (ONPG), 4mg/ml	50 μl
Mg^{2+} buffer, 100x (section 6.7)	3 μl

0.1M Sodium phosphate buffer (section 6.7) was added to a final volume of 300 μl. The reaction mixture was incubated at 37 °C until the appearance of a yellow colour. It was important to incubate the β-galactosidase assay until a yellow colour formed and the time ranged from 10 min to 2 hours, in which case the A_{420nm} was within the linear range of 0.2 to 0.8 AU. In certain instances, however, the incubation was carried out for 16 hours at 37 °C in the event where the transfection efficiency was low. The A_{420nm} was maintained within the linear range. The reaction was terminated by addition of 500 μl of 1M Na_2CO_3 and the A_{420nm} was measured. The extracts from cells transfected with water instead of plasmid DNA were used as blanks. The β-galactosidase activity was expressed as A_{420nm} per μg protein.

6.2.6 Chloroamphenicol acetyl transferase (CAT) assays

CAT activity was measured in a reaction mixture containing the following :-

Cell extract	50-100 μg protein
n-Butyryl CoA, 5mg/ml	5 μl
^{14}C-Chloroamphenicol	2.5 μl

(D-Threo[dichloroacetyl-1-^{14}C]Chloroamphenicol, 54 mCi/mmol, Amersham)

0.25M Tris-HCl, pH 8.0 was added to a final volume of 125 μl.

The tubes were incubated at 37 °C for 2 hours and the reaction terminated by addition of 300 μl xylene, mixed by vortexing and centrifuged to separate the two phases. The butyrylated chloroamphenicol separated into the xylene phase (top phase), while the unbutyrylated chloroamphenicol remained in the aqueous phase (bottom phase). The xylene phase was collected and re-extracted with 100 μl of 0.25M Tris-HCl, pH 8.0. The top, xylene phase (~190 μl) was transferred to a scintillation vial and a scintillation cocktail was added. The incorporation of the ^{14}C label was measured by counting in a Tri-Carb 4640 liquid scintillation counter (Packard). The CAT activity was expressed as cpm per μg protein. The relative CAT activity was expressed as the ratio of the CAT activity to β-galactosidase activity.

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

6.3.1 Preparation of competent cells

E. coli DH5α or XLBlue cells, were inoculated into 5 ml of Luria-Bertani (LB) broth (section 6.7) and grown overnight at 37 °C with shaking at 250 rpm (New Brunswick Scientific Co. Inc.). The overnight culture was used to inoculate 20 ml LB in a sterile 50 ml centrifuge tube giving a starting optical density$_{600nm}$ (OD_{600nm}) of about 0.1 as determined on a Beckman DU 650 Spectrophotometer. Cells were allowed to reach the exponential growth phase (OD_{600nm} of 0.4-0.6) at 37 °C with shaking at 250 rpm, and pelleted at 2 000 rpm in a benchtop centrifuge (Beckman Model TJ-6 Centrifuge) at 4 °C for 20 minutes. The pellet was gently resuspended in a 10x packed cell volume with ice-cold 50 mM $CaCl_2$ solution, incubated for 1-2 hours and pelleted by centrifugation at 1 000 rpm for 20 minutes at 4 °C. The pellet was gently resuspended in a small volume (3-5x PCV) of 50 mM $CaCl_2$-containing 50 % (v/v) glycerol and 100-200 μl aliquots were stored at -70 °C until required.

6.3.2 Transformation of competent cells

An aliquot of competent cells (50-100 μl) was mixed with 1ng of plasmid DNA or 10 μl of ligation mixture and incubated on ice for 1 hour. The cells were heat-shocked at 42 °C for 1 minute, 1 ml LB was added and the cells were incubated at 37 °C in a shaking incubator for 1 hour. 50-100 μl of the transformation mixture was plated on LB-agar plates containing 50 μg/ml ampicillin. Plates were incubated at 37 °C for 12-16 hours. Transformation efficiency was confirmed by transformation with 1 ng of pUC19 plasmid; competent cells should give 10^7-10^8 colonies per μg plasmid DNA.

6.4 DNA manipulation

6.4.1 Cloning of DNA fragments into plasmid vectors

DNA fragments were digested with appropriate restriction enzymes, the reaction mixture and incubation conditions were as described by the manufacturers of the enzymes. DNA fragments were separated by electrophoresis on 1-1.5 % (w/v) agarose gels and the gels stained with ethidium bromide. DNA bands of interest were excised from the gel and eluted using gene-elute columns (section 6.4.2.2) or Qiaex beads (section 6.4.2.1).

6.4.1.1 Preparation of DNA for blunt-end ligation

DNA fragments with 5′ overhangs were incubated with 1U of Klenow DNA polymerase in the presence of the polymerase buffer (supplied by the manufacturers of the enzyme), 0.5 mM dATP, dGTP, dTTP and dCTP in a final volume of 20 μl at 30 °C for 30 minutes to 1 hour. DNA was extracted once with phenol:chloroform:isoamyl alcohol (25:24:1), once with chloroform:isoamyl alcohol (24:1) and precipitated by the addition of 0.1 volume of 3 M sodium acetate, pH 4.0 and 2.5 volumes of 96 % (v/v) ethanol and placed at -20 °C for 1-2 hours. The precipitate was pelleted by centrifugation at 13 000 rpm at 4 °C for 20 minutes, the pellet washed with 70 % (v/v) ethanol, dried and dissolved in TE. The vector for blunt-end ligation was digested with either SmaI or EcoRV to generate blunt ends, treated with 1 unit calf intestinal phosphatase (CIP) at 37 °C for 1 hour and extracted with phenol:chloroform:isoamyl alcohol as described above. The DNA was precipitated by the addition of ethanol, washed with 70 % (v/v) ethanol, dried and dissolved in TE.

6.4.1.2 Preparation of DNA for cohesive-end ligation

DNA was digested with the required restriction enzymes and eluted as described in section 6.4.2.1 or 6.4.2.2. The vector was treated with CIP as described in section 6.4.1.1.

Ligation reactions contained 1U of T4 DNA ligase, 0.25 mM ATP, T4 DNA ligase buffer and the indicated ratios of vector DNA to insert DNA in a final volume of 10 μl. All ligations were incubated for 16-24 hours at room temperature.

6.4.2 DNA elution methods

6.4.2.1 Qiaex DNA extraction method

DNA fragments were separated by electrophoresis on low-melting agarose gels in TAE buffer (section 6.7) and stained with ethidium bromide. DNA bands of interest were excised and the gel treated as described by the manufacturers of the Qiaex gel extraction kit (Qiagen).

6.4.2.2 Gene-elute column DNA extraction method

DNA was fractionated by electrophoresis on high melting agarose gels in TBE (section 6.7), excised and eluted using gene-elute columns. The procedure was as described by the manufacturers (Supelco, Inc). Briefly, DNA was placed on a column equilibrated with TE, the column placed in a 1.5 ml Eppendorf tube and centrifuged in a microcentrifuge at 13 000 rpm for 10-30 minutes depending on the size of the gel. The DNA was used directly in the subsequent steps or it was precipitated by the addition of 0.1 volumes of 3M sodium acetate and 2.5 volumes of 96 % (v/v) ethanol.

6.4.2.3 Crush soak method

DNA was separated by electrophoresis on a non-denaturing polyacrylamide gel, stained with ethidium bromide and the band of interest excised. The gel was cut into small pieces and incubated in 300 μl of extraction buffer (0.5 mM ammonium acetate, 10 mM magnesium acetate and 1mM EDTA, pH 8.0) at 37 °C in a shaking incubator for 16 hours. The buffer was removed and the remaining gel pieces washed with 100 μl of the extraction buffer. Supernatants were pooled and extracted once with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1), once with chloroform:isoamyl alcohol (24:1) and precipitated with 0.1 volumes 3M sodium acetate and 2.5 volumes of 96 % (v/v) ethanol solution. The pellet was collected by centrifugation, washed with 70 % (v/v) ethanol, dried and dissolved in 20 µl TE (section 6.7).

6.4.3 Rapid plasmid DNA preparation

Plasmid DNA was extracted using a modified alkaline lysis method of Birnboim and Dolly (241). Overnight colonies were inoculated into 3 ml of LB supplemented with 50 µg/ml ampicillin and allowed to grow at 37 °C for 12-18 hours in a shaking incubator (250 rpm). 2 ml of the overnight culture was transferred into 1.5 ml Eppendorf tubes, centrifuged at 13 000 rpm at 4 °C for 2 minutes, the cell pellet resuspended in 100 µl of solution 1 (25 mM Tris-HCl, pH 8.0, 10 mM EDTA and 50 mM glucose), lysed with 200 µl freshly prepared solution 2 (0.25M NaOH and 1.0 % (w/v) SDS) and mixed by gentle inversion of the tubes (4-6 times). To the suspension, which at this stage should be very viscous, was added 200 µl of solution 3 (3M KAc, pH 4.8) the tube was inverted 4-6 times and centrifuged at 13 000 rpm at 4 °C for 5 minutes to pellet the membranes and cellular debris. 400-500 µl of the supernatant was mixed with an equal volume of 100 % (v/v) isopropanol, centrifuged at 13 000 rpm at 4 °C for 10 minutes to collect the precipitate, the pellet was washed with 70 % (v/v) ethanol and dried under vacuum. The dried pellet was resuspended in 40-50 µl of TE and subjected to digestion with restriction endonucleases as required.

6.4.4 Purification of plasmid DNA

Purification of plasmid DNA was performed using the Qiagen plasmid purification kits for mini or maxi preparations. Glycerol stocks were used to inoculate 5ml or 200 ml of LB supplemented with 50 µg/ml ampicillin. Cells were grown at 37 °C in a shaking incubator at 250 rpm for 16 hours, harvested by centrifugation and used for plasmid DNA isolation as recommended by the manufacturers of the kits.

6.4.5 Chromosomal DNA extraction

Cells were cultured in 150 mm tissue culture dishes until they reached 80-90 % confluency, harvested by brief trypsinisation, collected by centrifugation and washed once with phosphate buffered saline. The cell pellet was resuspended in 100 µl digestion buffer (100mM NaCl, 10 mM Tris-HCl, pH 8.0, 25 mM EDTA, pH 8.0, 0.5 % (w/v) SDS) supplemented with freshly made 0.1 mg/ml proteinase K. The samples were incubated at 50 °C for 12-18 hours in Eppendorf tubes (tubes were sealed with parafilm) with shaking. 2 µl of 10 mg/ml RNase was added and incubation continued at 50 °C for 10 minutes. The samples were extracted with an equal volume of phenol:chloroform:isoamyl alcohol (25.24.1) and centrifuged at top speed for 10 minutes. The aqueous (top) phase was transferred to a clean tube and to it was added 0.5 volumes of 7.5M ammonium acetate and 2 volumes of 96 % (v/v) ethanol; the DNA formed a stringy precipitate which was either collected by centrifugation at top speed for 2 minutes or it was spooled using a pasteur pipette. The DNA was washed once with 70 % (v/v) ethanol, air dried and dissolved in TE by heating at 65 °C if required.

6.4.6 DNA sequencing

DNA sequencing was performed using the Sanger dideoxy chain termination method (242). 1-2 µg purified plasmid DNA was mixed with 100 pmol of primer and 1.5 µl of 1M NaOH in a final volume of 11.5 µl, mixed by vortexing and incubated at 65 °C for 5 minutes. The reaction was transferred to a 37 °C waterbath, neutralised by the addition of 1.5 µl of 1M HCl and 2 µl of the annealing buffer followed by a further incubation at 37 °C for 10 minutes. The annealing reaction was cooled to room temperature for 5 minutes. Labelling and subsequent reactions were performed as recommended by the manufacturers (Pharmacia). 3 µl of the final sequencing reaction mixture containing stop buffer was loaded onto a urea 6 % polyacrylamide gel and fragments separated by electrophoresis in 1x TBE buffer. The gel was always pre-electrophoresed for 1 hour to a temperature of 45-50 °C prior to loading samples. The gel was dried and exposed to X-ray film for 16 hours at room temperature.

temperature for 5 minutes. Labelling and subsequent reactions were performed as recommended by the manufacturers (Pharmacia). 3 μl of the final sequencing reaction mixture containing stop buffer was loaded onto a urea 6 % polyacrylamide gel and fragments separated by electrophoresis in 1x TBE buffer. The gel was always pre-electrophoresed for 1 hour to a temperature of 45-50 °C prior to loading samples. The gel was dried and exposed to X-ray film for 16 hours at room temperature.

6.4.7 Polymerase Chain Reaction (PCR) amplification of DNA

A standard PCR reaction contained 25-50 ng of template (purified chromosomal or plasmid DNA), 200 μM dNTP mixture (dATP, dGTP, dCTP and dTTP), 1x *Taq* DNA polymerase reaction buffer containing $MgCl_2$, 100 pmol of primers and 0.5-1 unit of Takara *Taq* DNA polymerase in a final volume of 50-100 μl. The PCR reaction was over-layered with ~30 μl of mineral oil. Although PCR conditions varied with respect to the annealing temperature, a typical PCR programme (Hybaid OmniGene PCR machine) would be :-

Cycle no.	Temp.	Time
Cycle 1	94 °C	2 minutes
Cycle 2-25	94°C	30 seconds
	55°C	30 seconds
	72°C	1 minute
Cycle 26	72°C	10 minutes

5-10 μl of PCR products were analysed by electrophoresis on agarose gels followed by ethidium bromide staining.

6.4.8 Analysis of DNA by electrophoresis on agarose gels

1-1.5 % (w/v) agarose was dissolved in 0.5-1x TBE or TAE buffer, the agarose was dissolved by boiling in a microwave, after cooling to ~50 °C, ethidium bromide was contained 0.25-0.5 μg of DNA, 300 μM dNTP mixture of dATP, dGTP and dTTP, 20 μCi of dCTP [α-^{32}P] (specific activity 300 Ci/mmol, Amersham), 1x Klenow DNA polymerase reaction buffer and 1-2 units of Klenow DNA polymerase in a final volume of 25 μl. The reaction was incubated for 1 hour at room temperature or at 30 °C and the reaction was terminated by heating at 65 °C for 10 minutes. The reaction volume was made up to 150 μl with sterile distilled water and loaded onto a Sephadex G-50 column (packed in a standard pasteur pipette plugged with autoclaved glass-wool). 15 x 150 μl fractions were collected by elution with sterile distilled water and the radioactivity of each fraction determined by Cerenkov counting. Fractions from the first radioactive peak were pooled, 10 μl of the pooled fractions mixed with a liquid scintillation cocktail and the radioactivity measured by counting in a liquid scintillation counter. The specific activity of the labelled DNA was expressed as cpm/μg DNA.

6.4.9.2 Random prime labelling of DNA

The random prime labelling protocol was followed as described by the manufacturers of the kit (Amersham). 25 ng of DNA was heated for 5-10 minutes at 95 °C, snap-cooled on ice for 5 minutes and 5 μl of primer, 10 μl of reaction buffer, 50 μCi of [α-^{32}P]-dCTP and 2.5 units of Klenow DNA polymerase were added to a final volume of 50 μl. The reaction mixture was incubated for 1 hour in a 37 °C waterbath, the reaction was terminated by heating at 65 °C for 10 minutes and labelled DNA separated from unlabelled nucleotides on a Sephadex G-50 column. The specific activity of the labelled DNA calculated as described in section 6.4.9.1.

6.5 RNA extraction and analysis

6.5.1 RNA extraction

The one-step method of Chomczynski and Sacchi (243) was used to extract RNA from adherent fibroblasts. Fibroblasts were cultured to 80-90 % confluency in 100mm tissue culture dishes, washed once with PBS and the cells lysed by the addition of 500 µl of guanidinium isothiocyanate (GITC) solution D (section 6.7). The suspension was transferred to an Eppendorf tube, to it was added 50 µl of 3M sodium acetate, pH 4.0 and mixed by vortexing. 500 µl of water saturated phenol was added and the mixture was vortexed until all lumps had disappeared before the addition of 100 µl of chloroform:isoamyl alcohol (49:1). The suspension was mixed by shaking for 10 seconds followed by incubation on ice for 15 minutes. The mixture was then centrifuged at 13 000 rpm at 4 °C for 20 minutes. The top aqueous layer was mixed with an equal volume of 100 % (v/v) isopropanol, incubated at -20 °C for 2 to 16 hours and the RNA pelleted by centrifugation at 13 000 rpm for 20 minutes. The RNA pellet was washed in 70 % (v/v) ethanol, vacuum dried, resuspended in 50 µl of sterile DEPC treated distilled water and incubated at 60 °C for 10 minutes to dissolve the RNA pellet. The concentration of RNA was measured spectrophotometrically at λ_{260}.

6.5.2 Electrophoresis of RNA on agarose formaldehyde gels

Horizontal 1 % (w/v) agarose formaldehyde gels were prepared by boiling 0.4g of agarose in 27 ml of sterile distilled water and 4 ml of a 10x RNA running buffer (section 6.7); these proportions are adequate for the preparation of a minigel. The agarose was allowed to cool to ~60 °C and 8.5 ml of formaldehyde was added (NB: at this stage, it was important to work in the fume hood to avoid formaldehyde fumes). The agarose formaldehyde solution was poured into a gel casting tray and allowed to set at room temperature in the fume hood. 10-20 µg of RNA was mixed with 3.3 µl of sterile DEPC treated distilled water, 1.5 µl of 10x RNA running buffer, 7.5 µl of formamide, 2.7 µl of formaldehyde and 1.5 µl of RNA loading buffer (section 6.7) in a final volume of ~26 µl. The RNA mixture was heated at 55 °C for 10 minutes, loaded onto the agarose formaldehyde gel and fractionated by electrophoresis at 30 mA in 1x RNA running buffer. The buffer was mixed at 30 minutes time intervals. The gel was stained with ethidium bromide and rinsed with five changes of distilled water, or until the 28S and 18S bands could be visualised on the UV box in the absence of any background fluorescence.

6.5.3 Northern blotting of RNA onto nylon membrane

After visualisation of the 28S and 18S RNA bands under UV light, the RNA was transferred by blotting onto nylon membranes (Amersham Hybond-N) in 10x SSC using the sandwich method of Sambrook *et al* (244). The membrane was viewed under UV light with the RNA side facing down, to confirm successful transfer of the RNA. The membrane was then rinsed in 6x SSC (section 6.7), the excess liquid was removed by gentle blotting with Whatman 3M filter paper and the RNA fixed onto the membrane by UV cross-linking (Spectrolinker, XL-1000 UV crosslinker, Spectronics Corporation). The membrane was stored at 4 °C until required.

6.5.4 Hybridisation to radiolabelled probes

Cross-linked membranes were prehybridised in 5-10ml hybridisation buffer per 100 mm^2 membrane (section 6.7) in a hybridisation oven (Hybaid Dual hybridisation oven) at 42 °C for at least 4 hours. Random prime [α-^{32}P]-dCTP labelled DNA was denatured at 94 °C for 5-10 minutes, snap-cooled in ice for 5 minutes and added to the prehybridisation solution at 1-2 X 10^6 cpm/ml for a probe with a specific activity of between 10^8-10^9 cpm/μg DNA. Hybridisation was continued at 42 °C for 18-24 hours in a rotating hybridisation oven. The hybridisation solution was decanted and the membrane was washed twice in 2x SSC, 0.1 % (w/v) SDS at room temperature for 15 minutes each, followed by a high stringency wash in 0.1x SSC, 0.1 % (w/v) SDS at 56-65 °C for 30 minutes (depending on the probe). It was sometimes necessary to repeat the high stringency wash until the background was significantly reduced (monitored by a Geiger counter). The membrane was sealed in a plastic bag and exposed to X-ray film for at least 16 hours at -70 °C using intensifying screens.

6.6 Protein analysis

6.6.1 Preparation of nuclear proteins

Preparation of nuclear proteins was performed according to the method of Dignam *et al* (245). Cells were cultured to 90 % confluency in either 150 mm tissue culture dishes or flasks (see section 6.1.1 and 6.1.2) and adherent cells were harvested by scraping with a rubber policeman whilst suspension cultures were harvested by centrifugation with a Beckman JA 14 rotor at 18 000 xg at 4 °C for 10 minutes. The cell pellet was resuspended in 5 packed cell volumes of Dignam buffer A (section 6.7) and incubated on ice for 10 minutes to allow cells to swell. Cells were pelleted by centrifugation at 18 000 xg at 4 °C for 10 minutes, the pellet was resuspended in 2 packed cell volume of Dignam buffer A (section 6.7) and homogenised with 10-20 strokes using a Dounce homogeniser. Nuclei were collected by centrifugation at 21 000 xg (13 000 rpm with a Beckman JA20 rotor) at 4 °C for 20 minutes. Nuclei were resuspended in Dignam buffer C (section 6.7), homogenised and incubated on ice for 30 minutes with gentle stirring to facilitate extraction of nuclear proteins. The homogenate was centrifuged at 21 000 xg at 4 °C for 30 minutes and the supernatant dialysed against 50 volumes of Dignam buffer D (section 6.7). The dialysate was centrifuged at 21 000 xg (Beckman JA 20 rotor) at 4 °C for 30 minutes and the supernatant stored at -70 °C in aliquots of 50-100 μl. For purification purposes, the supernatant was stored in 1 ml aliquots.

6.6.2 Analysis of DNA binding activity of nuclear proteins using the electrophoretic mobility shift assay (EMSA)

Approximately 4 μg of crude nuclear extract was incubated with 1-2 μg of polydI.dC and 4 μl of incubation buffer (section 6.7) in a volume of 20-25 μl at room temperature for 10 minutes. Approximately 10^4 cpm of [α-^{32}P]-dCTP end-labelled probe was added (equivalent to ~4 ng of DNA) and the reaction incubated on ice for a further 20 minutes. Two microlitres of 0.2 % (w/v) bromophenol blue was added and DNA-protein complexes resolved by electrophoresis on a non-denaturing 5 % polyacrylamide gel. Briefly, a non-denaturing polyacrylamide gel was cast using standard procedures, the gel was pre-electrophoresed in 0.5x TBE buffer at 150 V at 4 °C for ~30 minutes prior to sample loading. The gel was electrophoresed at 150 V at 4 °C for approximately 2.5 hours, depending on the length of the probe, dried and exposed to X-ray film for a minimum of 16 hours at -70 °C using intensifying screens.

6.6.3 Purification of transcription factors

For the purification of transcription factors, conventional chromatography techniques were utilised and at each purification step, the activity of the transcription factor was monitored by EMSA for DNA binding activity. Throughout the purification, the 114 bp probe encompassing the α2(I) procollagen promoter between -107 and -60 was used as a probe in the EMSA reactions. All purification steps were performed at 4 °C.

6.6.3.1 Heparin-agarose purification

Commercial heparin-agarose was suspended in CB buffer (section 6.7) and degassed at room temperature for about 1 hour. A 15 cm x 0.7 cm heparin-agarose column was packed and equilibrated with CB buffer. Crude nuclear proteins loaded onto the column and eluted stepwise with CB buffer containing 0.1, 0.2, 0.4 or 1.0 M KCl. The elution of proteins from the column was monitored at λ_{280nm} using a UV detector. The volume of each fraction varied according to the starting material. Fractions were assayed for DNA binding activity by EMSA (6.6.2).

6.6.3.2 DNA affinity purification

6.6.3.2.1 Preparation of the DNA affinity matrix

The preparation of the DNA affinity matrix was performed according to the method of Kerrigan and Kadonaga (206). 20 µl of 10x T4 polynucleotide kinase buffer (supplied by the manufacturers of the enzyme) was added to a mixture containing 440 µg of each oligonucleotide (total volume of 130 µl), and oligonucleotides were annealed by incubating for 2 minutes at 88 °C, 10 minutes at 65 °C, 10 minutes at 37 °C, and 5 minutes at room temperature. The mixture was divided in half in separate 1.5 ml Eppendorf tubes. To each 75 µl aliquot, 15 µl of 20 mM ATP (pH 7.0), 10 µl of 10 U/µl T4 polynucleotide kinase (100 U total) were added and incubated for 2 hours at 37 °C. The kinase was inactivated by adding 50 µl of 10 M ammonium acetate and 100 µl water, heated for 15 minutes at 65 °C and allowed to cool to room temperature. To purify phosphorylated oligonucleotides, the mixture was precipitated with 750 µl of 96-100 % (v/v) ethanol, mixed by inversion and the pellet was collected by centrifugation for 15 minutes at room temperature. The pellet was resuspended in 225 µl TE buffer, extracted with 250 µl of 25:24:1 of phenol:chloroform:isoamyl alcohol. The aqueous phase (top layer) was transferred into a new tube, extracted with 250 µl of 24:1 chloroform:isoamyl alcohol, and the aqueous phase transferred into a new tube. Oligonucleotides were precipitated with 25 µl of 3M sodium acetate and 750 µl of 96-100 % (v/v) ethanol, and the pellet was collected by centrifugation at high speed for 15 minutes. The pellet was washed once with 800 µl of 70 % (v/v) ethanol, dried and resuspended in 65 µl of water. To this oligonucleotide mixture was added 10 µl of 10x T4 ligase buffer (supplied by the manufacturers of the enzyme), 20 µl of 20 mM ATP (pH 7.0), and 5 µl of 6000 U/ml T4 DNA ligase (30 Weiss units) and the mixture was incubated at room temperature for $\geq$ 16 hours. Ligation was monitored by agarose gel electrophoresis. Ligated oligonucleotides were purified by extraction with 100 µl of TE buffered phenol, the aqueous phase extracted with 100 µl of 24:1 chloroform:isoamyl alcohol, and precipitated with 33 µl of 10M ammonium acetate and 133 µl of 100 % (v/v) isopropanol. The mixture was incubated for 20 minutes at -20 °C, centrifuged and the pellet dissolved in 225 µl of TE buffer. To this mixture was added 25 µl of 3M sodium acetate and 750 µl of 96-100 % (v/v) ethanol, centrifuged and the pellet washed twice with 70 % (v/v) ethanol. The dried pellet was dissolved in 50 µl of water and stored at -20 °C.

4g of commercial CNBr-activated Sepharose 4B was placed in a 50 ml polypropylene tube, hydrated with 10 ml of 1M HCl and mixed gently by flicking and inverting the tube. After 1 minute, the slurry was transferred to a 60 ml coarse-sintered glass funnel, washed and the beads allowed to swell by gradually pouring 500 ml of 1mM HCl through the funnel. The resin was washed with 100 ml of water and with 100 ml of 10mM potassium phosphate, pH 8.0. The resin was immediately transferred to a 50 ml polypropylene screw-cap tube and 4 ml of 10 mM potassium (pH 8.0) was added until the resin had a consistency of a thick slurry. To couple oligonucleotide multimers to CNBr-Sepharose, the two 50 µl aliquots of DNA were added immediately to the thick slurry and incubated on a rotating wheel overnight at room temperature. Working in the hood, the resin was transferred to a 60 ml coarse-sintered glass funnel and washed with 2 x 100 ml washes of water and one 100 ml wash with 1M ethanolamine hydrochloride, pH 8.0 (0.2M Glycine could be used in place of ethanolamine hydrochloride). The resin was transferred to a 15 ml polypropylene screw-cap tube and 1M ethanolamine hydrochloride (pH 8.0) was added until the mixture was a smooth slurry, the tube incubated on a rotating wheel for 2-4 hours at room temperature. This is step is crucial for the inactivation of unreacted CNBr-activated Sepharose. The resin was transferred to a 60 ml coarse-sintered glass funnel and washed with 100 ml of 10mM potassium phosphate, pH 8.0, 100 ml of 1M potassium phosphate, pH 8.0, 100 ml of 1M KCl, 100 ml of water, and a 100 ml of column storage buffer (section 6.7). The resin was stored at 4 °C and it is stable for one year.

6.6.3.2.2 DNA affinity chromatography

Heparin-agarose fractions containing the transcription factor of interest, in this case complex II proteins, were pooled for further purification using a sequence specific DNA affinity matrix (section 6.6.3.2.1). The affinity matrix was packed into a 7 cm x 0.5 cm column in EMSA buffer (section 6.7) and equilibrated in EMSA buffer containing 0.1M KCl. Proteins were pre-incubated on ice for 30 minutes before loading in the presence of 0.1 μg/μl polydI.dC in order to reduce non-specific interactions with the matrix. Proteins were loaded onto the column at gravity flow and the column was washed with 20 ml of EMSA buffer containing 0.1M KCl. Since complex II proteins eluted in 0.4 M KCl, it was necessary to dilute heparin-agarose fractions to ~0.1 M KCl prior to DNA affinity chromatography. The column was washed with the EMSA buffer, proteins were eluted with increasing KCl concentrations and fractions were assayed for DNA binding activity by EMSA. The affinity resin was regenerated at room temperature with 5 ml of column regeneration buffer (section 6.7), the resin was stirred with a glass rod and the buffer allowed to flow out of the column. This step was repeated and 10 ml of column storage buffer was added, allowed to flow through, and 10 ml of storage buffer added again, the bottom of the column closed, the top covered and the column was stored at 4 °C.

6.6.4 Separation of proteins by electrophoresis on SDS-polyacrylamide gels

To cast discontinuous polyacrylamide gels containing SDS, the Mighty Small II SE 250 Dual Cooled Vertical Slab Gel Electrophoresis Unit (Hoefer Scientific Instruments) was used to form 8 cm wide x 7 cm long gels. Protein samples were mixed with Laemmli sample buffer (section 6.7) in the presence (reducing) or absence (non-reducing) of β-mercaptoethanol and separated by electrophoresis in SDS electrophoresis buffer (section 6.7) at 10 mA for about 2 hours. Gels were fixed and stained as described in section 6.6.5 and 6.6.6.

6.6.5 Silver staining of proteins separated by electrophoresis on SDS-polyacrylamide gels

A modification of the silver staining method of Merrill *et al* (214) was used to stain the gels. The silver nitrate staining protocol was used to monitor the extend of purification of transcription factors since this protocol can detect as little as 5 ng of protein per band. After electrophoresis, the gel was rinsed with distilled water, transferred into a glass container and fixed at room temperature for 30 minutes to 16 hours with gentle shaking in 2 % (w/v) $CuCl_2$ dissolved in 50 % (v/v) methanol-12 % (w/v) trichloroacetic acid. The gel was washed with 10 % (v/v) ethanol-5 % (v/v) acetic acid for 10 minutes, incubated in 0.01 % (w/v) KMn_6O_4 for 2 minutes, washed in the ethanol-acetic acid solution for 5 minutes followed by a single wash in 10 % (v/v) ethanol and distilled water for 5 minutes each. The gel was incubated in 0.1 % (w/v) $AgNO_2$ dissolved in distilled water for 5 minutes, rinsed with distilled water for about 20 seconds and then incubated in 10 % (w/v) K_2CO_3 dissolved in distilled water for 1 minute. The gel was rinsed with distilled water for 1 minute and the silver nitrate stain developed by incubation in 0.01 % (v/v) formaldehyde-2 % (w/v) K_2CO_3 solution until bands started to appear (about 5 minutes). The gel was rinsed with running distilled water, destained in 3 % (v/v) acetic acid for about 1 hour (optional) and stored in 1 % (v/v) glycerol; the gel was normally dried after a brief rinse in the glycerol solution. The incubation times were modified depending on the size and thickness of the gel. It was necessary to wear gloves at all times since this staining procedure is very sensitive and therefore even fingerprints on the gel would stain; all steps were performed at room temperature with gentle, continuous shaking and all solutions were prepared fresh prior to use.

6.6.6 Rapid Coomassie blue staining

SDS-polyacrylamide gels were fixed in a 25 % (v/v) isopropanol-10 % (v/v) acetic acid solution for 15 minutes and stained in a solution of 0.06 % (w/v) Commassie blue G-250 dissolved in 10 % (v/v) acetic acid for 1 hour. The gel was then destained by washing in 7 % (v/v) acetic acid-5 % (v/v) methanol until the background was clear; a sponge was placed in the destaining dish to accelerate the destaining process. The gel was rinsed in 1 % (v/v) glycerol solution and dried. Gels stained using this protocol were also subjected to silver nitrate staining in cases where bands were too weak.

6.6.7 Western blot analysis of proteins

6.6.7.1 Transfer of proteins onto polyvinyl difluoride (PVDF) membranes

Proteins were separated by electrophoresis on SDS-polyacrylamide gels, the gel rinsed twice in distilled water for 2 minutes each then soaked in CAPS transfer buffer (section 6.7) for 15-20 minutes. The PVDF membrane was placed in 20 ml 100 % (v/v) methanol for 5 minutes, the methanol was decanted and the membrane immediately rinsed in water for 5 minutes (it is important to ensure that the membrane is completely submerged in water). The membrane was then soaked in CAPS transfer buffer. A sandwich comprising of 2 x Whatman 3M filter paper, gel, PVDF membrane and 2 x Whatman 3M filter paper was assembled, the gel was placed on the negative terminal; proteins were transferred using the Mini Trans-Blot Electrophoretic Transfer Cell (BioRad) at 70V for 2 hours. After transfer, the membrane was stained in a 0.1 % (w/v) Coomassie blue G250 dissolved in a 50 % (v/v) methanol and 10 % (v/v) acetic acid solution (filtered on a No.2 Whatman filter paper), for 10 minutes, destained in 50 % (v/v) methanol, 10 % (v/v) acetic acid for 2 x 10 minutes each, rinsed in distilled water for 10 minutes, the bands of interest were excised and subjected to automated N-terminal sequencing.

6.6.7.2 Immunoblotting of proteins onto nitro-cellulose membranes

Proteins were separated by electrophoresis on polyacrylamide gels. The gel was rinsed with distilled water, soaked in Tris-glycine transfer buffer (section 6.7) and a sandwich comprising of 2 x Whatman 3M filter paper, gel (on the negative terminal side), nitrocellulose and 2 x Whatman 3M filter paper was assembled. Proteins were transferred onto the nitro-cellulose membrane at 100V for 2 hours. After transfer the membrane was incubated in a blocking solution composed of 5 % (w/v) non-fat milk in Tris buffered saline (section 6.7)-0.01 % (v/v) Tween 20 (TBST) for 2-16 hours. The membrane was incubated with a primary antibody for 1 hour at room temperature, washed twice with the milk-TBST solution for 15 minutes per wash, a secondary antibody-horseradish peroxidase conjugate was added at a 1:2000 dilution and the membrane incubated for 45 minutes at room temperature; all dilutions were made in the blocking solution. The membrane was washed twice with milk-TBST for 15 minutes per wash and a last wash was performed in TBS for 5 minutes. To develop protein-antibody complexes, the membrane was incubated in a staining solution (6 mg 4-chloro-1-naphthol dissolved in 2ml of ice-cold methanol, 6 μl of 30 % (v/v) H_2O_2 in a final volume of 10ml TBS) until bands appeared, that is, until the reactive areas turned purple. The membrane was rinsed in running distilled water; the colour development was complete within 30 minutes.

6.6.8 *In situ* trypsin digestion of proteins and extraction of peptides

Proteins were separated by electrophoresis on SDS-polyacrylamide gels, stained with Coomassie blue and the bands of interest excised. The gel was cut into small pieces, transferred into an Eppendorf tube, sufficient 100 % (v/v) acetonitrile added to cover gel pieces and incubated at room temperature for 10 minutes. The supernatant was discarded, gel pieces rehydrated in 5 mM Tris-HCl, pH 7.0 at room temperature for 10 minutes and gel pieces were washed twice in alternate washes with 100 % (v/v) acetonitrile and 5mM Tris-HCl, pH 7.0; the supernatant was discarded after each wash step. The final two washes were performed with 100 % (v/v) acetonitrile and gel pieces were incubated at 60 °C for 30 minutes in 10 mM DTT dissolved in 5 mM Tris-HCl, pH 7.0. This was followed by two x 10 minute washes with 100 % (v/v) acetonitrile and gel pieces were incubated at room temperature for 15 minutes in 50 mM iodoacetate dissolved in 5 mM Tris-HCl, pH 7.0. Gel pieces were washed twice with alternate changes of 100 % (v/v) acetonitrile and 5 mM Tris-HCl, pH 7.0 for 10 minutes each wash and a final wash with 100 % (v/v) acetonitrile at room temperature and twice at 37 °C. The digestion was performed with 4 % (w/v) trypsin, 0.2-0.4 μg of trypsin in 0.25 mM Tris-HCl, pH 7.0 was added per 5 μg protein, gel pieces were completely covered in digestion buffer and incubated at 37 °C for 6 to 16 hours To extract peptides, the supernatant was collected, the gel pieces washed alternately with changes of 100 % (v/v) acetonitrile for 15 minutes and 10 % (v/v) formic acid for 10 minutes. These washes were repeated 4 times and at each step, the supernatant was collected. The supernatants were pooled and dried under vacuum.

6.6.9 MALDITOF-ms analysis of peptides

Peptides were prepared as described in section 6.6.8 and the dried sample was resuspended in a 0.1 % (v/v) triflouroacetate:100 % (v/v) acetonitrile (2:1) solution (TFA:ACN). The peptide solution was mixed with sinapinic acid as a matrix, crystallisation was allowed to take place at room temperature (the presence of crystals was confirmed by light microscopy) and the crystals were subjected to MALDITOF-ms analysis using a Biospectrometry™ Research Station. A calibration mixture comprising bovine insulin (*m/z* 5734.59), *E coli* thioredoxin (*m/z* 11 674.4) and horseradish apomyoglobin (*m/z* 16 952.5) was used.

6.7 Solutions

TTE for transfections

1 mM Tris, pH 8.0

0.025 mM EDTA, pH 8.0

Filter sterilise (0.22 μ filters, Millipore) and store at room temperature.

$CaCl_2$ solution for transfections

2 M $CaCl_2$

10 mM Hepes

Adjust pH to 5.8, filter sterilise (0.22 μ), store at room temperature.

10x Hepes-buffered saline (HBS)

Hepes	1g
NaCl	1.6 g
KCl	0.074 g
Dextrose	0.20 g
1.45 mM Na_2HPO_4 (anhydrous)	0.103 g

Make up to 10 mL with distilled water and store at -20 °C. For transfection, the 10x HBS was diluted to just over 2x HBS immediately prior to use, the pH adjusted to 7.09 and the solution made up to volume. After filter sterilisation this solution could be stored at room temperature. Glucose can be used instead of dextrose.

100x Mg^{2+} buffer for β-galactosidase assay

0.1M $MgCl_2$

4.5M β-mercaptoethanol (β-MSH)

For 1 ml: 315 μl β-MSH , 100 μl $MgCl_2$, and 595 μl distilled water

Luria-Bertani (LB) broth

NaCl	1 g
Yeast extract	1 g
Tryptone	2 g

Make up to 200 ml with distilled water and sterilise by autoclaving. Store at room temperature.

LB plates

NaCl	2.5 g
Yeast extract	2.5 g

Tryptone 5 g

Agar 10 g

Weigh out the above reagents into a 500 mL reagent bottle, make up to 500 mL with distilled water and sterilise by autoclaving; the agar dissolves during autoclaving. Cool the LB-agar solution to about 50 °C and add ampicillin to a final concentration of 50 µg/mL. Pour into petri dishes, allow the agar to set, cover dried plates with cling wrap and store at 4 °C until use.

1x TAE buffer

40 mM Tris

5 mM Sodium acetate

1 mM EDTA

Adjust pH to 7.8 with glacial acetic acid. Sterilise by autoclaving and store at room temperature.

1x TBE buffer

90 mM Tris, pH 8.0

90 mM Boric acid

2.5 mM EDTA

Sterilise by autoclaving and store at room temperature.

DNA loading buffer

50 % (v/v) Glycerol

1x TBE

0.5 % (w/v) Bromophenol blue

0.5 % (w/v) Xylene cyanol

Store in 1 mL aliquots at -20 °C.

GITC Solution D

4 M Guanidinium isothiocyanate (GITC)

25 mM Sodium acetate

0.5 % (w/v) Sarkosyl

Sterilise by autoclaving. Store in a dark bottle at room temperature. Add 36 μl of β-mercaptoethanol per 5 mL of GITC solution D prior to use.

RNA running buffer, 10x

200 mM MOPS, pH 7.0

50 mM sodium acetate

10 mM EDTA

RNA loading buffer

1 mM EDTA

0.25 % (w/v) Bromophenol blue

0.25 % (w/v) Xylene cyanol

50 % (v/v) Glycerol

Store in 1 mL aliquots at -20 °C.

20x SSC

3 M NaCl

0.3 M Sodium citrate, pH 7.0, sterilise by autoclaving and store at room temperature.

Hybridisation buffer (~10 mL)

Dextran sulphate	1 g
Distilled water	0.8 mL
20x SSC	2.5 mL

Dissolve by heating in a microwave, then add the following:-

500 mM Sodium pyrophosphate, pH 6.5	1 mL
100x Denhardt's solution	0.5 mL
Formamide	5 mL
10 mg/mL sonicated herring sperm DNA	100 μL
10 % (w/v) SDS	100 μL

50x Denhardt's solution

Ficoll 400	0.5 g
Polyvinylpyrolidone	0.5 g
BSA	0.5 g

Dissolve and make up to 50 mL in water. Store at -20 °C in 2 ml aliquots.

Dignam Buffer A

10 mM Hepes, pH 7.9

1.5 mM $MgCl_2$

10 mM KCl

0.5 mM DTT

1 μg/mL Leupeptin

1 μg/mL Pepstatin A

A solution containing the first three ingredients was sterilised by autoclaving and stored at 4 °C. Protease inhibitors and DTT were only added just before use.

Dignam buffer C

20 mM Hepes, pH 7.9

25 % (v/v) Glycerol

0.45 mM NaCl

1.5 mM $MgCl_2$

0.2 mM EDTA

0.5 mM DTT

0.5 mM PMSF

1 μg/mL Leupeptin

1 μg/mL Pepstatin A

A solution containing the first five ingredients was sterilised by autoclaving and stored at 4 °C. Protease inhibitors and DTT were only added just before use.

Dignam buffer D

20 mM Hepes, pH 7.9

25 % (v/v) Glycerol

0.1 mM KCl

0.2 mM EDTA

0.5 mM DTT

0.5 mM PMSF

1 μg/mL Leupeptin

1 μg/mL Pepstatin A

A solution containing the first four ingredients was sterilised by autoclaving and stored at 4 °C. Protease inhibitors and DTT were only added just before use.

5x Incubation (EMSA) buffer

100 mM Hepes, pH 7.9

250 mM KCl

1.0 mM EDTA

5 mM $MgCl_2$

20 % (w/v) Ficoll 400

The solution was stored in 1 mL aliquots at -20 °C; 2.5 μL of 100 mM DTT was added prior to use.

CB buffer

50 mM Tris-HCl, pH 7.9

0.1 mM EDTA

20 % (v/v) Glycerol

0.5 mM DTT

0.5 mM PMSF

1 μg/mL Leupeptin

1 μg/mL Pepstatin A

The buffer containing the first three ingredients was sterilised by autoclaving and stored at room temperature. Protease inhibitors and DTT were only added just before use.

Column storage buffer

10 mM Tris-HCl, pH 7.8

1 mM EDTA, pH 8.0

0.3 M NaCl

0.04 % (w/v) sodium azide

The buffered was stored at room temperature without sodium azide. A 4 % (w/v) of sodium azide stock solution was prepared and added just before use.

Column regeneration buffer

10 mM Tris-HCl, pH 7.8

1 mM EDTA, pH 8.0

2.5 mM NaCl

1 % (v/v) NP-40

The solution turns cloudy and separates into two phases, i.e. the NP-40 and the aqueous phases, upon storage. Mix by swirling and shaking just before use.

1x Laemmli buffer

125 mM Tris-HCl, pH 6.8

1 % (w/v) SDS

5 % (v/v) Glycerol

1 % (v/v) β-Mercaptoethanol

0.03 % (w/v) Bromophenol blue

The buffer was stored at -20 °C in 1 mL aliquots.

1x SDS electrophoresis buffer

25 mM Tris-HCl, pH 8.3

192 mM Glycine

0.1 % (w/v) SDS

Store in a brown bottle at room temperature.

CAPS transfer buffer

10 mM CAPS, pH 10.5

10 % (v/v) methanol buffer

Tris-glycine transfer buffer

25 mM Tris

192 mM glycine

20 % (v/v) methanol

Tris-buffered saline

10 mM Tris-HCl, pH 8.0

150 mM NaCl

REFERENCES

1. Ramirez, F. and di Liberto, M. *Complex and diversified regulatory programs control the expression of vertebrate collagen genes.* **FASEB J**. (1990) 4: 1616-1623.
2. Van der Rest, M. and Garrone, R. *Collagen family of proteins.* **FASEB J**. (1991) 5: 2815-2823.
3. Cole, W.G. *Collagen genes: mutations affecting collagen structure and expression.* **Progress Nucleic Acid Res**. (1994) 47: 29-80.
4. Prockop, D.J. and Kivirikko, K.I. *Collagens: molecular biology, diseases, and potentials for therapy.* **Annu. Rev. Biochem**. (1995) 64: 403-434.
5. Karsenty, G. and de Crombrugghe, B. *Conservation of binding sites for regulatory factors in the coordinately expressed* $\alpha 1(I)$ *and* $\alpha 2(I)$ *collagen promoters.* **Biochem. Biophys. Res. Commun**. (1991) 177: 538-544.
6. Slack, J.L., Liska, D.J., and Bornstein, P. *Regulation of expression of the type I collagen genes.* **American J. Med. Genetics** (1993) 45: 140-151.
7. Karsenty, G. and Park, R. *Regulation of type I collagen gene expression.* **Inter. Rev. Immunol**. (1995) 12:177-185.
8. Bornstein, P. and Sage, H. *Structurally distinct collagen types.* **Annu. Rev. Biochem**. (1980) 49: 957-1003.
9. Trojanowska, M., Leroy, E.C., Eckes, B. and Krieg, T. *Pathogenesis of fibrosis : type I collagen and the skin.* **J. Mol. Med**. (1998) 76: 266-274.
10. Rojkind, M. and Martinez-Palomo, A. *Increase in type I and type III collagens in human alcoholic liver cirrhosis.* **Proc. Natl. Acad. Sci. USA** (1976) 73: 539-543.
11. Kahari, V.M., Vuorio, T., Nanto-Salonen, K. and Vuorio, E. *Increased type I collagen mRNA levels in cultured scleroderma fibroblasts.* **Biochim. Biophys. Acta** (1984) 781: 183-186.
12. Sobel, M.R., Yamamoto, T., de Crombruggle, B. and Pastan, Z. *Regulation of procollagen messenger ribonucleic acid levels in Rous sarcoma virus transformed duck embryo fibroblasts.* **Biochemistry** (1981) 20: 2678-2684.

13. Parker, M.I., Smith, A.A., Mundell, K., Collins, M., Boast, S. and Ramirez, F. *The abolition of procollagen gene expression in SV40-transformed fibroblasts is associated with trans-acting factor switching.* **Nucleic Acids Res**. (1992) 20: 5825-5830.
14. Yoshimura. M., Jimenez, S.A. and Kaji, A. *Effects of viral transformation on synthesis and secretion of collagen and fibronectin-like molecules by embryonic chick chondrocytes in culture.* **J. Biol. Chem**.(1981) 256: 9111-9117.
15. Allebach, E.S., Boettiger, D., Pacific, M. and Adams, S.L. *Control of type I and type II collagen and fibronectin gene expression in chondrocytes delineated by viral transformation.* **Mol. Cell Biol**. (1985) 5: 1002-1008.
16. Thenet, S., Benya, P.D., Demignot, S., Feunteun, J. and Adolphe, M. *SV40 immortalisation of Rabbit articular chondrocytes: alteration of differentiated functions.* **J. Cell Physiol**. (1992) 150: 158-167.
17. Sandmeyer, S., Smith, R., Kiehn, D. and Bornstein, P. *Correlation of collagen synthesis and procollagen messenger RNA levels with transformation in rat embryo fibroblasts.* **Cancer Res**. (1981) 41: 830-838.
18. Schmidt, A., Setoyama, C. and de Combrugghe, B. *Regulation of a collagen gene promoter by the product of viral mos oncogene.* **Nature** (1985) 314(6008) : 286-289.
19. Propst, F., Storm, S.M. and Rapp, U.R. *Protein serine/threonine kinases: mos and raf,* In: **The molecular basis of human cancer**, Neel, B.G. and Kumar, R., eds. New York, Futura Publishing Co. (1993): 173-201.
20. Eizenberg, O. and Oren, M. *Reduced levels of $\alpha 1(I)$ collagen mRNA in cells immortalised by mutant p53 or transformed by ras.* **Biochim**. **Biophys**. **Acta** (1991) 1129: 34-42.
21. Slack, J.L, Parker, M.I., Robinson, V.R. and Bornstein, P. *Regulation of procollagen gene expression by ras.* **Mol. Cell. Biol**. (1992) 12: 4714-4723.
22. Hoemann, C.D. and Zarbl, H. *Use of revertant cell lines to identify targets of v-fos transformation-specific alterations in gene expression.* **Cell Growth Diff**. (1990) 1:581-590.

23. Setoyama, C. Hatamochi, A., Peterkofsky, B., Prather, W. and de Combrugghe, B. *V-fos stimulates expression of the $\alpha 1(III)$ collagen gene in NIH3T3 cells.* **Biochem. Biophys. Res. Commun.** (1986) 136: 1042-1048.

24. Peters, D.M., Dowd, N., Brandt, C.and Compton, T. *Human papilloma virus E6/E7 genes can expand the lifespan of human corneal fibroblasts.* **In vitro Cell. Dev. Biol. Anim.** (1996) 32: 279-284.

25. Santala, P., Larjava, H., Nissinen, L., Riikonen, T., Maatta, A. and Heina, J. *Suppressed collagen gene expression and induction of $\alpha_2\beta_1$ integrin-type collagen receptor in tumourigenic derivatives of human osteogenic sarcoma (HOS) cell line.* **J. Biol. Chem.** (1994) 269: 1276-1283.

26. Hatamochi, A., Aumialley, M., Mauch, C., Chu, M.L., Timpl, R. and Krieg, T. *Regulation of collagen VI expression in fibroblasts. Effects of cell density, cell matrix interactions, and chemical transformation.* **J. Biol. Chem.** (1989) 264: 3494-3499.

27. Marsilio, E., Sobel, M.E. and Smith, B.D. *Absence of procollagen $\alpha 2(I)$ mRNA in chemically transformed rat liver epithelial cells.* **J. Biol. Chem.** (1984) 259: 1401-1404.

28. Smith, B.D. and Marsilio, E. *Methylation of $\alpha 2(I)$ collagen gene in chemically transformed rat liver epithelial cells.* **Biochem J.** (1988) 253: 269-273.

29. Perkofsky, B. and Prather, W. *Suppression of synthesis of pro-$\alpha 11(I)$ and production of altered pro-$\alpha 2(I)$ procollagen subunits in 4-nitroquinoline-1-oxide-transformed fibroblasts.* **J. Biol. Chem.** (1986) 261: 16818-16826.

30. Asokan, R., Puvanakrishnan, R., Ravichandran, L.V., Kokila, V., Reddy, G.K. and Dhar, S.C. *Purification and characterisation of collagens from rat fibrosarcoma induced by 3-methylcholanthrene.* **Mol. Cell Biochem.** (1993) 121: 99-107.

31. Chen, A. and Davis, B.H. *UV irradiation activates JNK and increases $\alpha 1(I)$ collagen gene expression in rat hepatic stellate cells.* **J. Biol. Chem.** (1999) 274: 158-164.

32. Namba, M., Nishitami, K. and Kimoto, T. *Characteristics of WI-38 cells (WI-38 CT-1) transformed by treatment with Co-60 gamma rays.* **Gann**. (1980) 71: 300-307.

33. Parker, M.I., Smith, A.A., Boast, S., Wan Su, M. and Ramirez, F. *Regulation of the human α2(I) procollagen gene by a trans-acting factor.* **Ann. N.Y. Acad. Sci.** (1990) 580: 451-453.

34. Hatamochi, A., Ono, M., Ueki, H. and Namba, H. *Regulation of collagen gene expression in transformed human fibroblasts: decreased type I and type III collagen RNA transcription.* **J. Invest. Dermatol**. (1991) 96: 473-477.

35. Boast, S., Su, M.W., Ramirez, F., Sanchez, M. and Avvedimento, E.V. *Functional analysis of cis-acting DNA sequences controlling transcription of the human type I procollagen genes.* **J. Biol. Chem**. (1990) 265: 13351-13356.

36. Miao, K., Potter, J.J., Anania, F.A., Rennie-Tankersley, L. and Mezey, E. *Identification of two repressor elements in the mouse α2(I) collagen promoter.* **Arch. Biochem. Biophys**. (1999) 36: 7-16.

37. Hatamochi, A., de Crombrugghe, B. and Krieg, T. *Purification of a novel factor which binds to the mouse α2(I) promoter.* **FEBS Lett**. (1993) 373: 325-331.

38. Inagaki, Y., Truter, S. and Ramirez, F. *Transforming growth factor β stimulates α2(I) procollagen gene expression through a cis-acting element that contains an Sp1 binding site.* **J. Biol. Chem**. (1994) 269: 14828-14834.

39. Tamaki, T., Ohnishi, K., Hartl, C., LeRoy, E.C. and Trojanowska, M. *Characterisation of a GC-rich region containing Sp1 binding site(s) as a constitutive responsive element of the α2(I) collagen gene in human fibroblasts.* **J. Biol. Chem**. (1995) 270: 4299-4304.

40. Ihn, H. and Trojanowska, M. *Sp3 is a transcriptional activator of the human α2(I) procollagen gene.* **Nucleic Acids Res** (1997) 25: 3712-3717.

41. Greenwel, P., Inagaki, Y., Hu, W., Walsh, M. and Ramirez, F. *Sp1 is required for the early response of α2(I) procollagen to TGF-β1.* **J. Biol. Chem**. (1997) 272: 19738-19745.

42. Chung, K., Agarval, A., Uitto, J. and Mauviel, A. *An AP-1 binding sequence is essential for regulation of the human $\alpha 2(I)$ procollagen (COL1A2) promoter activity by TGFβ.* **J. Biol. Chem.** (1996) 271: 3272-3278.
43. Kouba, D.J., Chung, K.Y., Nishiyama, T., Vindevoghel, L., Kon, A., Klement, J.F., Uitto, J. and Mauviel, A. *NF-κB mediates TNFα inhibitory effect on $\alpha 2(I)$ collagen (COL1A2) gene transcription in human dermal fibroblasts.* **J. Immunol.** (1999) 162: 4220-4234.
44. Oikarinen,, J., Hatamochi, A. and de Crombrugghe, B. *Separate binding sites for nuclear factor 1 and a CCAAT DNA binding factor in the mouse $\alpha 2(I)$ collagen promoter.* **J. Biol. Chem.** (1987) 262: 11064-11070.
45. Karsenty, G., Golumbek, P. and de Crombrugghe, B. *Point mutations and small substitution mutations in three different upstream elements inhibit the activity of the mouse $\alpha 2(I)$ collagen promoter.* **J. Biol. Chem.** (1988) 263: 13909-13915.
46. Rossi, P., Karsenty, G., Roberts, A.B., Roche, N.S., Sporn, M.B. and de Crombrugghe, B. *A nuclear factor 1 binding site mediates the transcriptional activation of a type I collagen promoter by transforming growth factor β.* **Cell** (1988) 52: 405-414.
47. Ihn, H., Ohnishi, K., Tamaki, T., LeRoy, E.C. and Trojanowska, M. *Transcriptional regulation of the human $\alpha 2(I)$ procollagen gene.* **J. Biol. Chem.** (1996) 271: 26717-26723.
48. Hasegawa, T., Takeuchi, A., Miyaishi, O., Isobe, K. and de Crombrugghe, B. *Cloning and characterisation of a transcription factor that binds to the proximal promoter of the two mouse type I procollagen genes.* **J. Biol. Chem.** (1997) 272: 4915-4923.
49. Higashi, K., Kouba, D.J., Song, Y.J., Uitto, J. and Mauviel, A. *A proximal element within the human $\alpha 2(I)$ collagen (COL1A2) promoter, distinct from the tumour necrosis factor α response element, mediates transcriptional repression by interferon-γ.* **Matrix Biol.** (1998) 116: 447-456.
50. Hasegawa, T., Zhou, X., Garrett, L.A., Ruteshouser, E.C., Maity, S.N. and de Crombrugghe, B. *Evidence for three major transcription activation elements in*

the proximal mouse proα2(I) procollagen promoter. **Nucleic Acids Res**. (1996) 24: 3253-3260.

51. Collins, M., Leaner, V.D., Madikizela, M. and Parker, M.I. *Regulation of the human α2(I) procollagen gene by sequences adjacent to the CCAAT box.* **Biochem. J**. (1997) 322: 199-206.
52. Chen, S.S., Ruteshouser, E.C., Maity, S.N. and de Crombrugghe, B. *Cell-specific in vivo DNA-protein interactions at the proximal promoters of the proα1(I) and the proα2(I) procollagen genes.* **Nucleic Acids Res** (1997) 25: 3261-3268.
53. Maity, S.N. and de Crombrugghe, B. *Role of the CCAAT-binding protein CBF/NF-Y in transcription.* **TIBS** (1998) 23:174-178.
54. Hasegawa, T., Zhou, X., Garret, L.A., Ruteshouser, C., Maity, S.N. and de Crombrugghe, B. *Evidence for three major transcription activation elements in the proximal mouse proα2(I) collagen promoter.* **Nucleic Acid Res**. (1996) 24: 3253-3260.
55. Hatamochi, A., Paterson, B. and de Crombrugghe, B. *Differential binding of a CCAAT DNA binding factor to the promoters of the mouse α2(I) and α1(III) collagen genes.* **J. Biol. Chem**. (1986) 261: 11310-11314.
56. Collins, M., Smith, A.A. and Parker, M.I. *Characterisation of two distinct families of transcription factors that bind to the CCAAT box region of the human COL1A2 gene.* **J. Cell Biochem**. (1998) 70: 455-467.
57. Collins, M. *Characterisation of the human α2(I) procollagen promoter-binding proteins.* **PhD Thesis** (1993) University of Cape Town.
58. Kim, I.S., Sinha, S., de Crombrugghe, B. and Maity, S.N. *Determination of functional domains in the C subunit of the CCAAT binding factor (CBF) necessary for formation of a CBF-DNA complex; CBF-B interacts with both the CBF-A and CBF-C subunits to form a heterotrimeric CBF molecule.* **Mol. Cell Biol**. (1996) 16: 4003-4013.
59. Leaner, V.D. *Transcriptional regulation of the α2(I) procollagen gene.* **PhD Thesis** (1997) University of Cape Town.

60. Sengupta, P.K. and Smith, B.D. *Methylation in the initiation region of the first exon suppresses collagen pro-α2(I) gene transcription.* **Biochim. Biophys. Acta** (1998) 1443: 75-89.
61. Rossi, P. and de Crombrugghe, B. *Identification of a cell-specific transcriptional enhancer in the first intron of the mouse α2(I) collagen gene.* **Proc. Natl. Acad. Sci. USA** (1987) 84: 5590-5594.
62. Pogulis, R.J. and Freytag, S.O. *Contribution of specific cis-acting elements to activity of the mouse pro-α2(I) collagen enhancer.* **J. Biol. Chem.** (1993) 268: 2493-2499.
63. Sherwood, A.L., Bottenus, R.E., Martzen, M.R. and Bornstein, P. *Structural and functional analysis of the first intron of the human α2(I) procollagen-encoding gene.* **Gene** (1990) 89: 239-244.
64. Slack, J.L, Parker, M.I. and Bornstein, P. *Transcriptional repression of the α1(I) collagen gene by ras is mediated in part by an intronic AP-1 site.* **J. Cell Biochem.** (1995) 58: 380-392.
65. Massagué, J. *The transforming growth factor β family.* **Annu. Rev. Cell Biol.** (1990) 6: 597-641.
66. Kivirikko, K.I. *Collagens and their abnormalities in a wide spectrum of diseases.* **Ann. Med.** (1993) 25: 113-126.
67. Kahari, V.M. *Activation of dermal connective tissue in scleroderma.* **Ann. Med.** (1993) 25: 511-518.
68. Weiner, F.R., Shah, A., Smith, P.J., Rubin, C.S. and Zern, M.A. *Regulation of collagen gene expression in 3T3-L1 cells. Effects of adipocyte differentiation and tumour necrosis factor α.* **Biochemistry** (1989) 28: 4094-4099.
69. Piquet, P.F., Collart, M.A., Grau, G.E., Sappino, A.P. and Vassalli, P. *Requirement of tumour necrosis factor for development of silica-induced pulmonary fibrosis.* **Nature** (1990) 344: 245-247.
70. Mahone, A., Kukkola, A., Risteli, L., Risteli, J. and Mäenpää. *Type I procollagen synthesis is regulated by steroids and related hormones in human osteosarcoma cells.* **J. Cell Biochem.** (1998) 68: 151-163.

71. Komm, B.S., Terpenting, C.M., Benz, D.J., Graeme, K.A., Gallegos, A., Korc, M., Greene, G.L., O'Malley, B.W. and Haussler, M.R. *Oestrogen binding, receptor mRNA, and biological response in osteoblast-like osteosarcoma cells.* **Science** (1988) 241: 81-84.

72. Turner, R.T., Colvard, D.S. and Spelberdg, T.C. *Oestrogen inhibition of periosteal bone formation in rat long bones: down-regulation of gene expression for bone matrix proteins.* **Endocrinology** (1990) 127: 1346-1351.

73. Silbiger, S., Lei, J. and Neugarten, J. *Estradiol suppresses type I collagen synthesis in mesangial cells via activation of activator protein-1.* **Kidney Int**. (1999) 55: 1268-1276.

74. Harrison, J.R., Petersen, D.N., Lichtler, A.C., Mador,, A.T., Rowe, D.W. and Kream, B.E. *1,25-Dihydroxyvitamin D3 inhibits transcription of type I collagen genes in the rat osteosarcoma cell line ROS17/2.8.* **Endocrinology** (1989) 125:327-333.

75. Kuroki, Y., Shiozawa, S., Kano, J. and Chihara, K. *Competition between c-fos and 1,25 $(OH)_2$ Vitamin D3 in the transcriptional control of type I collagen synthesis in MC3T3-E1 osteoblastic cells.* **J. Cell Physiol**. (1995) 164: 459-564.

76. Li, E., Bestor, T.H. and Jacnisch, R. *Targeted mutation of the DNA methyltransferase gene results in embryonic lethality.* **Cell** (1992) 69: 915-926.

77. Cross, S.H. and Bird, .A.P. *CpG islands and genes.* **Curr. Opin. Genet. Dev.** (1995) 5: 309-314.

78. Graesmann, M. and Graesmann, A. *DNA methylation, chromatin structure and the regulation of gene expression.* **EXS** (1993) 64: 404-424.

79. Yeivin, A. and Razin, A. *Gene methylation patterns and expression.* **EXS** (1993) 64: 523-568.

80. Singal, R., Ferris, R., Little, J.A., Wang, S.Z. and Ginder, G.D. *Methylation of the minimal promoter of an embryonic globin gene silences transcription in primary erythroid cells.* **Proc. Natl. Acad. Sci. USA** (1997) 94: 113724-13729.

81. Nan, X., Campoy, F.J. and Bird, A. *MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin.* **Cell** (1997) 88: 471-481.

82. Umezawa, A., Yamamoto, H., Rhodes, K., Klemsz, M.J., Maki, R.A. and Oshima, R.G. *Methylation of an ETS site in the intron enhancer of the keratin 18 gene participates in tissue specific repression.* **Mol. Cell Biol.** (1997) 17: 4885-4894.

83. Fernandez, M.P., Young, M.F. and Sobel, M.E. *Methylation of type II and type I collagen genes in differentiated and dedifferentiated chondrocytes.* **J. Biol. Chem.** (1985) 260: 2374-2378.

84. Thompson, J.P., Simkevich, C.P., Holmess, M.A., Kamg, A.H. and Raghow, R. *In vitro methylation of the promoter and enhancer of pro-alpha 1(I) collagen gene leads to its transcriptional inactivation.* **J. Biol. Chem.** (1991) 266: 2549-2556.

85. Burbelo, P.D., Horikoshi, S. and Yamada, Y. *DNA methylation and collagen IV gene expression in F9 teratocarcinoma cells.* **J. Biol. Chem.** (1990) 265: 4839-4843.

86. Rhodes, K., Rippe, R.A., Umezawa, A., Nehls, M., Brenner, D.A. and Breindl, M. *DNA methylation represses the murine $\alpha 1(I)$ collagen promoter by an indirect mechanism.* **Mol. Cell Biol.** (1994) 14: 5950-5960.

87. Kopp, M.U., Winterhalter, K.H. and Trueb, B. *DNA methylation accounts for the inhibition of collagen VI expression in transformed fibroblasts.* **Eur. J. Biochem.** (1997) 249: 489-496.

88. Guenette, D.K., Ritzenthalerm J.D., Foley, J., Jackson, J.D. and Smith, B.D. *DNA methylation inhibits transcription of procollagen $\alpha 2(I)$ promoters.* **Biochem. J.** (1992) 283: 699-703.

89. te Water Naudé, J.A.L. *Oncogene expression in transformed cells.* **PhD Thesis** (1994) University of Cape Town.

90. Mizuki, K., Nose, K., Okambo, H., Tsuchida, N. and Hayashi, K. *Amplification of c-ki-ras gene and aberrant expression of c-myc in WI-38 cells transformed in vitro by γ-irradiation.* **Biochem. Biophys. Res. Commun.** (1985) 128: 1037-1043.

91. Davidson, J.M., Zoia, O. and Lin, J.M. *Modulation of TGFβ1 stimulated elastin and collagen production and proliferation in porcine vascular smooth muscle and skin fibroblasts.* **J. Cell Physiol.** (1993) 155: 149-156.

92. Watson, S., Burnside, T. and Carver, W. *Angiotensin II-stimulated collagen gel contraction by heart fibroblasts: role of the AT1 receptor and tyrosine kinase activity.* **J. Cell Physiol.** (1998) 177: 224-231.

93. Chu, A.J. and Prasad, J.K. *Up-regulation by human recombinant TGFβ-1 of collagen production in cultured dermal fibroblasts is mediated by the inhibition of nitric oxide signaling.* **Am. J. Coll. Surg.** (1999) 188: 271-280.

94. Sohis-Herruzo, J.A., Hernandez, I., De la Torre, P., Garcia, I., Sanchez, J.A., Fernandez, I., Castellano, G. and Munoz-Yague, T. *G proteins are involved in the suppression of collagen α1(I) gene expression in cultured rat hepatic stellate cells.* **Cell Signal.** (1998) 10: 173-183.

95. Ichiki, Y., Edwin, A., Smith, E., Leroy, C. and Trojanowska, M. *Basic fibroblast growth facto inhibits basal and TGFβ induced collagen α2(I) gene expression in scleroderma and normal fibroblasts.* **J. Rheumat.** (1997) 24: 90-95.

96. Ahmad, M., Theofanidis, P. and Medform, R.M. *Role of activating protein-1 in the regulation of the vascular cell adhesion molecule-1 gene expression by TNFα.* **J. Biol. Chem.** (1998) 273: 4616-4621.

97. Han, Y., Meng, T., Murray, N.R., Fields, A.P. and Brasier, A.R. *Interleukin-1 induced NF-κB autoregulatory feedback loop in hepatocytes.* **J. Biol. Chem.** (1999) 274: 939-947.

98. Karin, M. and Hunter, T. *Transcriptional control by protein phosphorylation: signal transmission from the cell surface to the nucleus.* **Current Biol.** (1995) 5: 747-757.

99. May, M.J. and Ghosh, S. *Signal transduction through NF-κB.* **Immunol. Today** (1998) 19: 80-88.

100. Stein, B., Baldwin, A.S.,Jr., Ballard, D.W., Greene, W.C., Angel, P. and Herrlich, P. *Cross-coupling of NF-κB p65 and Fos/Jun transcription factors produces potentiated biological function.* **EMBO J.** (1993) 12: 3879-3891.

101. Ogretmen, B. and Safa, A.R. *Negative regulation of MDR1 promoter activity in MCF-7, but not in multidrug resistant MCF-7/Adr, cells by cross-coupled NF-κB/p65 and c-Fos transcription factors and their interaction with the CAAT region.* **Biochemistry** (1999) 38: 2189-2199.

102. Nishizuka, Y. *Studies and perspectives of protein kinase C.* **Science** (1986) 234: 305-312.

103. Sobel, M.E., Dion, L.D., Vuust, J. and Colburn, N.H. *Tumour-promoting phorbol esters inhibit procollagen synthesis at a pretranslational level in JB-6 mouse epidermal cells.* **Mol. Cell Biol.** (1983) 3: 1527-1532.

104. Rabin, M.S., Doherty, P.J. and Gottesman, M.M. *The tumour promoter PMA induces a program of altered gene expression similar to that induced by PDGF and transforming oncogenes.* **Proc. Natl. Acad. Sci. USA** (1986) 83: 357-360.

105. Stuiver, I., Shimizu, Y. and Shimizu, N. *Phorbol ester mediated expression of the collagen type I pro-α2 gene in mouse 3T3-L1 cells and its absence in a phorbol-12-myristate 13-acetate-non-responsive variant.* **Biochem. J.** (1991) 278: 369-373.

106. Fisher, G.J. and Voorhees, J.J. *Molecular mechanisms of photoaging and its prevention by retinoic acid: ultraviolet irradiation induces MAP kinase signal transduction cascades that induce AP-1 regulated matrix metalloproteinases that degrade human skin in vivo.* **J. Invest. Dermatol. Symp. Proc.** (1998) 3: 61-68.

107. Matrisian, L.M. *The matrix-degrading metalloproteinases.* **Trends Genet.** (1990) 6: 121-125.

108. Matrisian, L.M. *The matrix-degrading metalloproteinases.* **Bioassays** (1992) 14: 455-463.

109. Friedman, S.L., Roll, F.J., Boyles, J. and Bissell, D.M. *Hepatic lipocytes: the principal collagen producing cells of normal rat liver.* **Proc. Natl. Acad. Sci. USA** (1985) 82: 8681-8685.

110. Davis, B.H., Chen, A. and Beno, D.W.A. *Raf and mitogen-activated protein kinase regulate stellate cell collagen gene expression.* **J. Biol. Chem.** (1996) 271: 11039-11042.

111. Ledwith, B.J., Manam, S., Kraynak, A.R., Nichols, W.W. and Bradley, M.O. *Antisense-fos RNA causes partial reversion of the transformed phenotypes induced by the c-Ha-ras oncogene.* **Mol. Cell Biol.** (1990) 10: 1545-1555.

112. Chambers, A.F. and Tuck, A.B. *Ras-responsive genes and tumour metastasis.* **Crit. Rev. Oncog.** (1993) 4: 95-114.

113. Willigham, M.C., Yamada, K.M., Yamada, S.S., Ponysségur, J. and Pastan, J. *Microfilament bundles and cell shape are related to adhesiveness to substratum and are dissociable from growth control in cultured fibroblasts.* **Cell** (1997) 10:375-380.

114. Travers, H., French, N.S. and Norton, J.D. *Suppression of tumourigenicity in Ras-transformed fibroblasts by α2(I) collagen.* **Cell Growth Diff.** (1996) 7: 1353-1360.

115. Westermarck, J., Ilvonen, E. and Kähäri, V.M. *The protein phosphatase inhibitor okadaic acid suppresses type I collagen gene expression in cultured fibroblasts at the transcriptional level.* **Biochem. J.** (1995) 308: 995-999.

116. Meek, D.W. and Street, A.J. *Nuclear protein phosphorylation and growth control.* **Biochem. J.** (1992) 287: 1-15.

117. Angel, P. and Karin, M. *The role of Jun, Fos and AP-1 complex in cell proliferation and transformation.* **Biochim. Biophys. Acta** (1991) 1072: 129-157.

118. Hunter T. *Oncoprotein networks.* **Cell** (1997) 88: 333-346.

119. Finkel, J.P., Biskis, B.O. and Jinkins, P.B. *Virus induction of osteosarcoma in mice.* **Science** (1966) 151: 698-701.

120. Finkel, J.P., Reilly, C.A. and Biskis, B.O. *Viral etiology of bone cancer.* **Front. Radiat. Ther. Oncol.** (1975) 10:28-39.

121. van Beveren, C., van Straaten, F., Curran, T., M üller,, R. and Verma, I.M. *Analysis of FBJ-MuSV provirus and c-fos (mouse) gene reveals that viral and cellular fos gene products have different carboxy termini.* **Cell** (1983) 32: 1241-1255.

122. Miller, A.D., Curran, T. and Verma, I.M. *c-Fos protein can induce cellular transformation: A novel mechanism of activation of cellular oncogene.* **Cell** (1984) 36: 51-60.

123. Wisdom, R. and Verma, I.M. *Transformation by Fos proteins requires a C-terminal transactivation domain.* **Mol. Cell Biol.** (1993) 13: 7429-7438.

124. Lee, W.M.F., Lin, C. and Curran, T. *Activation of the transforming potential of the human fos proto-oncogene requires message stabilisation and results in increased amounts of partially modified fos protein.* **Mol. Cell Biol.** (1988) 8: 5521-5527.

125. Skinner, M., Qu, S., Moore, C. and Wisdom, R. *Transcriptional activation and transformation by FosB protein requires phosphorylation of the carboxy-terminal activation domain.* **Mol. Cell Biol.** (1997) 17: 2372-2380.

126. van Straaten, F., Müller, R.,Curran, T., van Beveren, C. and Verma, I.M. *Complete nucleotide sequence of a human c-onc gene: deduced amino acid sequence of the human c-fos protein.* **Proc. Natl. Acad. Sci. USA** (1983) 80: 3183-3187.

127. Distel, R.J. and Spiegelman, B.M. *Protooncogene c-fos as a transcription factor.* **Adv. Cancer Res.** (1990) 55: 37-55.

128. Curran, T., Miller, D., Zokas, L. and Verma, I.M. *Viral and cellular fos proteins: a comparative analysis.* **Cell** (1984) 36: 259-268.

129. Mehmet, H. and Rozengurt, E. *Regulation of c-fos expression in Swiss 3T3 cells: An interplay of multiple signal transduction pathways.* **British Med. Bulletin** (1991) 47: 76-86.

130. Rahmsdorf, H.J., Schöntal, A., Angel, P., Liftin, M., Rüther, U and Herrlich, P. *Post transcriptional regulation of c-fos expression.* **Nucleic Acids Res.** (1987) 15: 1643-1659.

131. Wilson, T. and Treisman, R. *Removal of polyA and consequent degradation of c-fos mRNA facilitated by 3'AU-rich sequences.* **Nature** (1988) 336: 396-399.

132. Curran, T. and Teich, N.M. *Candidate product of the FBJ murine osteosarcoma virus oncogene: characterisation of a 55,000 dalton phosphoprotein.* **J. Virol.** (1982) 42: 114-122.

133. Franza, B.R., Sambucetti, L.C., Cohen, D.R. and Curran, T. *Analysis of Fos protein complexes and Fos related antigens by high resolution two-dimensional gel electrophoresis.* **Oncogene** (1987) 1:213-221.

134. Curran, T., van Beveren, C., Ling, N. and Verma, I.M. *Viral and cellular fos proteins are complexed with a 39,000 dalton cellular protein.* **Mol. Cell Biol.** (1985) 5: 167-172.

135. Sassone-Corsi, P., Ransome, L.K., Lamp, W.W. and Verma, I.M. *Direct interaction between Fos and Jun nuclear oncoproteins: role of the leucine zipper domain.* **Nature** (1988) 336: 692-695.

136. Sassone-Corsi, P. *Goals for signal transduction pathways: linking up with transcriptional regulation.* **EMBO J.** (1994) 13: 4717-4728.

137. Field, S.J., Johnson, R.S., Mortensen, R.M., Papaioannou, V.E., Spiegelman, B.M. and Greenberg, M.E. *Growth and differentiation of embryonic stem cells that lack an intact c-fos gene.* **Proc. Natl. Acad. Sci. USA** (1992) 89: 9306-9310.

138. Rahm, M., Jin, P., Surregi, J. and Sejersen, T. *Elevated c-fos expression inhibits differentiation of L6 rat myoblasts.* **J. Cell Physiol.** (1989) 139: 237-244.

139. Wang, Z.Q., Grigoriadis, A.E., Mohle-Steinlein, U. and Wagner, E.F. *A novel target cell for c-fos induced oncogenesis: development of chondrogenic tumours in embryonic stem cell chimeras.* **EMBO J.** (1991) 10: 2437-2450.

140. Lee, M.S., Yang, J.H., Sacchi, Z., Arnstein, P., Chen, L.S., Jay, G. and Rhim, J.S. *Neoplastic transformation of a human keratinocyte cell line by the v-fos oncogene.* **Oncogene** (1993) 8: 387-393.

141. Smeyne, R.J., Schilling, K., Robertson, L., Luk, D., Oberdick, J., Curran, T. and Morgan, J.I. *fos-LacZ transgenic mice: mapping sites of gene induction in the central nervous system.* **Neuron** (1992) 8:13-23.

142. Grigoriadis, A.E., Schellander, K., Wang, Z.Q. and Wagner, E.F. *Osteoblasts are target cells for transformation in c-fos transgenic mice.* **J. Cell Biol.** (1993) 122: 685-701.

143. Preston, G.A., Lyon, T.T., Yin, Y., Lang, J.E., Solomon, G., Annat, L., Srinivasan, D.G., Alcorta, D.A. and Barrett, J.C. *Induction of apoptosis by c-fos protein.* **Mol. Cell Biol.** (1996) 16: 211-218.

144. Berkowitz, L.A., Riabowol, W.T. and Gilman, M.Z. *Multiple sequence elements of a single functional class are required for cAMP responsiveness of the mouse c-fos promoter.* **Mol. Cell Biol.** (1989) 9: 4272-4281.

145. Sadowski, H.B., Shuai, K., Darnell, J.E. and Gilman, M.Z. *A common nuclear signal transduction pathway activated by growth factor and cytokine receptors.* **Science** (1993) 261: 1739-1744.

146. Fisch, T.M., Prywes, R., Simon, M.C. and Roeder, R.G. *An AP-1 binding site in the c-fos gene can mediate induction by EGF and 12-O-tetradecanoyl phorbol-13-acetate.* **Mol. Cell Biol.** (1989) 9: 1327-1331.

147. Wagner, B.J., Hayes, T.H., Hoban,, C.J. and Cochran, B.H. *The SIF binding element confers sis/PDGF inducibility onto the c-fos promoter.* **EMBO J.** (1990) 9: 4477-4484.

148. Sheng, M., Dougan, S.T., McFadden, G. and Greenberg, M.E. *Calcium and growth factor pathways of c-fos transcriptional activation require distinct upstream regulatory sequences.* **Mol. Cell Biol.** (1988) 8: 2787-2796.

149. Shaui, K., Ziemiecki, A., Wilks, A.F., Harper, A.G., Sadowski, H.B., Gilman, M.Z. and Darnell, J.E. *Polypeptide signalling to the nucleus through tyrosine phosphorylation of Jak and Stat proteins.* **Nature** (1993) 366: 580-583.

150. Sassone-Corsi, P., Sisson, J.C. and Verma, I.M. *Transcription autoregulation of the proto-oncogene fos.* **Nature** (1988) 334: 314-319.

151. Schönthal, A., Büscher, M., Angel, P., Rahmsdorf, H.J., Ponta, H., Hattori, K., chiu, R., Karin, M. and Herrlich, P. *The Fos and Ju/AP-1 proteins are involved in the downregulation of Fos transcription.* **Oncogene** (1989) 4: 629-636.

152. Bateman, E. *Autoregulation of eukaryotic transcription factors.* **Prog. Nucleic Acid Res.** (1998) 60: 133-168.

153. Lucibello, F.C., Lowag, C., Neuberg, M. and Müller, R. *Trans-repression of the mouse c-fos promoter: a novel mechanism of fos-mediated trans-regulation.* **Cell** (1989) 59: 999-1007.

154. Roux, P., Blanchard, J.M., Fernandez, A., Lamb, N., Jeanteur, P. and Piechaczyk, M. *Nuclear localisation of c-Fos, but not v-Fos proteins, is controlled by extracellular signals.* **Cell** (1990) 63: 341-351.

155. Chiu, R., Bayle, W.J., Meek, J., Smeal, T, Hunter, T. and Karin, M. *The c-fos protein interacts with c-jun/AP-1 to stimulate transcription of AP-1 responsive genes.* **Cell** (1988) 54: 541-552.

156. Lee, W., Mitchell, P. and Tjian, R. *Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements.* **Cell** (1987) 49: 741-752.

157. Ransone, L.J. and Verma, I.M. *Nuclear proto-oncogenes fos and jun.* **Annu. Rev Cell Biol**. (1990) 6: 539-557.

158. Curran, T. *Fos and Jun: Oncogenic transcription factors.* **Tohoku J. Exp. Med**. (1992) 168: 169-174.

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

160. Vogt, P.K. and Boss, T.J. *jun: Oncogene and transcription factor.* **Adv. Cancer Res**. (1990) 55: 2-36.

161. Cohen, D.R. and Curran, T. *Analysis of dimerisation and DNA binding functions in Fos and Jun by domain-swapping: Involvement of residues outside of the leucine zipper/basic region.* **Oncogene** (1990) 5: 929-939.

162. Abate, C., Luk, D. and Curran, T. *Transcriptional regulation by Fos and Jun in vivo: Interaction among multiple activator and regulatory domains.* **Mol. Cell Biol**. (1991) 11: 3624-3632.

163. Sassone-Corsi, P., Lamph, W.W., Kamps, M. and Verma, I.M. *Fos-associated cellular p39 is related to nuclear transcription factor AP-1.* **Cell** (1988) 54: 553-560.

164. Nakabeppu, Y., Ryder, K. and Nathans, D. *DNA binding activities of three murine jun proteins: stimulation by fos.* **Cell** (1988) 55: 907-915.

165. Hirai, S.I., Ryseck, R.P., Mechte, F., Bravo, R. and Yaniv, M. *Characterisation of junD: a new member of the jun oncogene family.* **EMBO J**. (1989) 8: 1433-1439.

166. Zerial, M., Toschi, L., Ryseck, R.P., Schuermann, M., Muller, R. and Bravo, R. *The product of a novel growth factor activated gene, fosB, interacts with jun proteins enhancing their DNA binding activity.* **EMBO J**. (1989) 8: 805-813.

167. Cohen, D.R., Ferreira, P.C.P., Gentz, R., Franza Jr. and B.R.,Curran, T. *The product of a fos-related gene, fra-1, binds cooperatively to the AP-1 site with Jun: transcription factor AP-1 is comprised of multiple protein complexes.* **Genes Dev**. (1989) 3: 173-184.

168. Matsui, M.M., Tokuhara, M., Konuma, Y., Nomura, N. and Ishizaki, R. *Isolation of human fos-related genes and their expression during monocyte-macrophage differentiation.* **Oncogene** (1990) 5: 249-255.

169. Ryseck, R.P. and Bravo, R. *c-Jun, JunB and JunD differ in their binding affinities to AP-1 and CRE consensus sequences: effects of fos proteins.* **Oncogene** (1991) 6: 533-542.

170. Sambucetti, L. and Curran, T. *The fos protein complex is associated with DNA in isolated nuclei and binds to DNA cellulose.* **Science** (1986) 234: 1417-1419.

171. Renz, M., Vernier, B, Kurz, C. and Muller, R. *Chromatin association and DNA binding properties of the c-fos proto-oncogene product.* **Nucleic Acid Res**. (1987) 15: 277-292.

172. Setoyama, C., Frunzio, L., Liau, G., Murdyj, M. and de Crombrugghe, B. *Transcriptional activation encoded by the v-fos gene.* **Proc. Natl. Acad. Sci. USA** (1986) 83: 3212-3217.

173. Kuroki, Y., Shiozawa, S., Sugimoto, T. and Fujita, T. *Constitutive expression of c-fos gene inhibits type I procollagen synthesis in transfected osteoblasts.* **Biochem. Biophys. Res. Comm**. (1992) 182: 1389-1394.

174. Rüther, U, Garber, C., Kamitowski, D., Müller, R. and Wagner, E.F. *Deregulated c-fos expression interferes with normal bone development in transgenic mice.* **Nature** (1987) 325: 412-416.

175. Wang, Z.Q., Ovitt, C., Grigoriadis, A.E., Möhle-Steinlein, U., Rüther, U. and Wagner, E.F. *Bone and haematopoietic defects in mice lacking c-fos.* **Nature** (1992) 360: 741-745.

176. Wong, Z.Q., Liang, J., Schellander, K., Wagner, E.F. and Grigoriadis, A.E. *c-fos induced osteosarcoma formation in transgenic mice: cooperativity with c-jun and the role of endogenous c-fos.* **Cancer Res**. (1995) 55: 6244-6251.

177. Binetruy, B., Smeal, T. and Karin, M. *Ha-Ras augments c-jun activity and stimulates phosphorylation of its activation domain.* **Nature** (1991) 351: 122-127.

178. Sassone-Corsi, P., Der, C.J. and Verma, I.M. *ras-induced neuronal differentiation of PC12 cells: possible involvement of fos and jun.* **Mol. Cell Biol.** (1989) 9: 3174-3183.

179. Stacey, D.W., Watson, T., Kung, H.F. and Curran, T. *Microinjection of transforming ras protein induces c-fos expression.* **Mol. Cell Biol.** (1987) 7: 523-527.

180. Schönthal, A., Herrlich, P., Rahmsdorf, H.J. and Ponta, H. *Requirement for fos gene expression in the transcriptional activation of collagenase by other oncogenes and phorbol esters.* **Cell** (1988) 54: 325-334.

181. Deng, T. and Karin, M. *JunB differs from c-Jun in its DNA-binding and dimerisation domains, and represses c-Jun by formation of inactive heterodimers.* **Genes Dev.** (1993) 7: 479-490.

182. Bornstein, P. and Sage, H. *Regulation of procollagen gene expression.* **Prog. Nucleic Acid Res. Mol. Biol.** (1989) 37: 67-106.

183. Slack, J.L, Parker, M.I., Robinson, V.R. and Bornstein, P. *Regulation of procollagen gene expression by ras.* **Mol. Cell Biol.** (1992) 12: 4714-4723.

184. Hooft van Huijsdjuijen, R.A.M., Bolleckens, J., Dorn, A., Benoist, C. and Mathis, D. *Properties of a CCAAT box-binding protein.* **Nucleic Acids Res.** (1987) 15: 7265-7282.

185. Dreier, B., Fritz, S., Volkert, G., Bronchvagel, B., Lottspeich, F. and Fey, G.H. *A 32 kDa proteolytic fragment of transcription factor Stat3 is capable of specific DNA binding.* **J. Biol. Chem.** (1997) 272: 22118-11124.

186. Higuchi, R., Krummel, B. and Saiki, R.K. *A general method of in vitro preparation and specific mutagenesis of DNA fragments: Study of protein and DNA interactions.* **Nucleic Acids Res.** (1988) 16: 7351-7367.

187. Coustry, F., Sinha, S., Maity, S.N. and de Crombrugghe, B. *The two activation domains of the CCAAT-binding factor CBF interact with the dTAF$_{II}$110 component of the Drosophila TFIID complex.* **Biochem. J.** (1998) 331: 291-297.

188. Ogbourne, S. and Antalis, T.M. *Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes.* **Biochem. J.** (1998) 331: 1-14.

189. Chen, H., Campisi, J. and Padmanabhan, R. *SV40 large T antigen transactivates the human cdc2 promoter by inducing a CCAAT box binding factor.* **J. Biol. Chem.** (1996) 271: 13959-13967.

190. Rangan, V.S., Oskonian, B. and Smith, S. *Identification of an inverted CCAAT box motif in the fatty acid synthase gene as an essential element for mediation of transcriptional regulation by cAMP.* **J. Biol. Chem.** (1996) 271: 2307-2312.

191. Ziyadeh, F.N., Snipes, E.R., Watanabe, M., Alvarez, R.J., Goldfarb, S. and Haverty, T.P. *High glucose induces cell hypertrophy and stimulates collagen gene trasncription in proximal tubule.* **Am. J. Physiol.** (1990) 259: F704-F714.

192. Kindy, M.S., Chang, C.J. and Sonenshein, G.E. *Serum deprivation of vascular smoothe muscle cells enhances collagen gene expression.* **J. Biol. Chem.** (1988) 263: 11426-11430.

193. Vaage, J. and Lindblad, W.J. *Production of collagen type I by mouse peritoneal macrophages.* **J. Leuk. Biol.** (1990) 48: 274-280.

194. Kim, J. and Kemfer, B. *Phenobarbital alters protein-binding to the CYP2B1/2 phenobarbital-responsive unit in native chromatin.* **J. Biol. Chem.** (1997) 272: 29423-29425.

195. Kara, C.J and Glimcher, L.H. *In vivo footprinting of MHC class II genes: bare promoters in the bare lymphocyte syndrome.* **Science** (1991) 252: 709-712.

196. Sturn, R.A, Das, G. and Herr, W. *The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain.* **Genes Dev.** (1988) 2: 1582-1599.

197. Staudt, L.M., Clerc, R.G., Singh, H., LeBowitz, J.H., Sharp, P.A. and Baltimore, D. *Cloning of a lymphoid-specific cDNA encoding a protein binding the regulatory octamer DNA motif.* **Science** (1988) 241: 577-580.

198. Hoeffler, J.P., Meyer, T.E., Yun, Y., Jameson, J.H. and Habener, J.F. *Cyclic AMP-responsive DNA-binding protein: structure based on a cloned placental cDNA.* **Science** (1988) 242: 1430-1433.

199. Fletcher, C., Heintz, N. and Roeder, R.G. *Purification and characterisation of OTF-1, a transcription factors regulating cell cycle expression of a human histone H2b gene.* **Cell** (1987) 51: 773-781.

200. Scheidereit, C., Heguy, A. and Roeder, R.G. *Identification and purification of a human lymphoid specific octamer binding protein in vitro.* **Cell** (1987) 51: 783-793.

201. Usuda, Y., Kubota, A., Berk, A.J. and Handa, H. *Affinity purification of transcription factor IIA from HeLa cell nuclear extracts.* **EMBO J**. (1991) 10: 2305-2310.

202. Maity, S.N., Vuorio, T. and de Crombrugghe, B. *The B subunit of a rat heteromeric CCAAT-binding transcription factor shows a striking sequence identity with the yeast Hap2 transcription factor.* **Proc. Natl. Acad. Sci. USA** (1990) 87: 5378-5382.

203. Vuorio, T., Maity, S.N. and de Crombrugghe, B. *Purification and molecular cloning of the "A" chain of a rat heteromeric CCAAT-binding protein.* **J. Biol. Chem**. (1990) 265: 22480-22486.

204. Briggs, M.R., Kadonaga, J.T., Bell, S.P. and Tjian, R. *Purification and biochemical characterisation of the promoter specific transcription factor, Sp1.* **Science** (1986) 234: 47-52.

205. Ruteshouser, E.C. and de Crombrugghe, B. *Purification of BBF, a DNA-binding protein recognising a positive cis-acting element in the mouse $\alpha 1(III)$ procollagen promoter.* **J. Biol. Chem**. (1992) 267: 14398-14404.

206. Kerrigan, L.A. and Kadonaga, J.T. *Purification of sequence specific DNA-binding proteins by affinity chromatography.* In: **Current Protocols in Molecular Biology** Greene Publishing Associates, New York (1993).

207. Creighton, T.E. **Protein Structure: A Practical Approach Series** IRL Press (Oxford University Press), 2nd Edition, 1997.

208. Briggs, M.R., Kadonaga, J.T., Bell, S.P. and Tjian, R. *Purification and biochemical characterisation of the promoter specific transcription factor, Sp1.* **Science** (1986) 234: 47-52.

209. Wang, D. and Kudlow, J.E. *Purification and characterisation of TEF1, a transcription factor that controls the human transforming growth factor α promoter.* **Biochim. Biophys. Acta** (1999) 1449: 50-62.

210. Gai, B., Jaffe, H. and Kunos, G. *Histone H1 isoforms purified from rat liver bind nonspecifically to the nuclear factor 1 recognition sequence and serve as generalised transcriptional repressors.* **Acta Biochim. Pol.** (1997) 44: 579-589.

211. Ryn, S., Zhou, S., Ladurner, A.G. and Tjian, R. *The transcriptional cofactor complex CRSP is required for activity of the enhancer binding protein Sp1.* **Nature** (1999) 397: 446-450.

212. Benmark, M., Olsson, H., Heinigard, D. and Leanderson, T. *Purification and characterisation of a protein binding to the SP6κ promoter. A potential role of CArG box binding factor A in κ transcription.* **J. Biol. Chem.** (1998) 273: 18881-18890.

213. Rosenthal, R.S. and Rodwell, V.W. *Purification and characterisation of the heteromeric transcriptional activator MvaT of the Pseudomonas mevalonii mvaAB operon.* **Protein Sci.** (1998) 7: 174-184.

214. Merril, C.R., Goldman, D. and van Keuren, M.L. *Gel protein stains: silver stain.* **Methods in Enzymol.** (1984) 104: 441-447.

215. Usuda, Y., Kubota, A., Berk, A.J. and Handa, H. *Affinity purification of transcription factor IIA from HeLa cell nuclear extracts.* **EMBO J.** (1991) 10: 2305-2310.

216. Maity, S.N. and de Crombrugghe, B. *Biochemical analysis of the B subunit of the heteromeric CCAAT-binding factor.* **J. Biol. Chem.** (1992) 267: 8286-8292.

217. Rauschar, F.J. 3rd, Morris, J.F., Fredericks, W.J., Lopez Guisa, J., Balakrishman, C., Jost, M., Herlyn, M. and Rodeck, U. *Characterisation of monoclonal antibodies directed to the amino-terminus of the WT1, Wilms' tumour suppressor protein.* **Hybridoma** (1998) 17: 191-198.

218. Chong, B.E., Lubman, D.M, Rosenspire, A. and Miller, F. *Protein profiles and identification of high performance liquid chromatography isolated proteins of cancer cell lines using matrix assisted laser desorption/ionisation time-of-flight mass spectrometry.* **Rapid Commun. Mass Spectrom** (1998) 12: 1986-1993.

219. Chowdhury, K., Rohdewohld, H. and Gruss, P. *Specific and ubiquitous expression of different Zn finger protein genes in the mouse.* **Nucleic Acids Res.** (1988)16: 9995-10011.

220. Klug, A. and Rhodes, D. *"Zinc fingers": a novel protein motif for nucleic acid recognition.* **TIBS** (1987) 12: 465-469.

221. Schenker, T. and Trueb, B. *Down-regulated proteins of mesenchymal tumour cells.* **Exp. Cell Res.** (1998) 239: 161-168.

222. Andreú, T., Beckers, T., Thoemes, E., Hilgard, P. and von Melchner, H. *Gene trapping identifies inhibitors of oncogenic transformation.* **J. Biol Chem.** (1998) 273: 13848-13854.

223. Diamond, M.I., Miner, J.N., Yoshinaga, S.K. and Yamamoto, K.R. *Transcription factor interactions: selectors of positive and negative regulation from a single DNA element.* **Science** (1990) 249: 1266-1272.

224. Bose, B., Soriani, M. and Tyrell, R.M. *Activation of expression of the c-fos oncogene by UVA irradiation in cultured human skin fibroblasts.* **Photochem. Photobiol.** (1999) 69: 489-493.

225. Lochter, A. and Bissell, J. *Involvement of extracellular matrix constituent in breast cancer.* **Seminars in Cancer Biol.** (1995) 6: 165-173.

226. Mervier, T., Gaillard Senchez, I., Martel, P. and Seillan-Heberden, C. *Constitutive overexpression of c-fos protein in rat liver epithelial cells decreases TGF-β synthesis and increases TGF-β_1 receptors.* **Biochim. Biophys. Acta** (1995) 1266: 64-72.

227. Jenuwein, T., Muller, D.,, Curran, T. and Muller, R. *Extended lifespan and tumourigenicity of non-established mouse connective tissue cells transformed by the fos oncogene of FBR-MuSV.* **Cell** (1985) 41: 629-637.

228. Maatta, A. and Penttinen, R.P.K. *A fibroblast protein binds the 3′ untranslated region of pro-α1(I) collagen mRNA.* **Biochem. J.** (1993) 295: 691-698.

229. Maatta, A. and Penttinen, R.P.K. *Nuclear and cytoplasmic α1(I) collagen mRNA binding proteins.* **FEBS Lett.** (1996) 340: 71-77.

230. Stefanovic, B., Hellerbrand, C., Holcik, M., Briendl, M., Liebhaber, S.A. and Brenner, D.A. *Posttranscriptional regulation of collagen α1(I) mRNA in hepatic stellate cells.* **Mol. Cell Biol.** (1997) 17: 5201-5209.

231. Nordheim, A. and Rich, A. *Negatively supercoiled simian virus 40 DNA contains Z-DNA segments within transcriptional enhancer sequences.* **Nature** (1983) 303: 674-679.

232. Rich, A., Nordheim, A. and Wang, A.H.J. *The chemistry and biology of left-handed Z-DNA.* **Annu. Rev. Biochem.** (1984) 54: 791-846.

233. Bakin, A.V. and Curran, T. *Role of DNA 5-methylcytosine transferase in cell transformation by fos.* **Science** (1999) 283: 387-390.

234. Philips, N., Bashey, R.I. and Sergio, A.J. *Increased α1(I) procollagen gene expression in tight skin (TSK) mice myocardial fibroblasts is due to a reduced interaction of negative regulatory sequence with AP-1 transcription factor.* **J. Biol. Chem.** (1995) 270: 9313-9321.

235. Tillman, K., Oberfield, J.L., Shen, X.Q., Bubulya, A. and Shemshedini, L. *c-Fos dimerisation with c-Jun represses c-Jun enhancement of androgen receptor transactivation.* **Endocrine** (1998) 9: 193-200.

236. Berger, I. and Shaul, Y. *c-Fos antagonises the junD gene positive autoregulatory loop; a novel c-Fos role in promoter switching.* **Gene** (1998) 211: 375-382.

237. Hai, T. and Curran, T. *Cross-family dimerisation of transcription factors Fos/Jun and ATF/CREB alters DNA binding specificity.* **Proc. Natl. Acad. Sci. USA** (1991) 88: 3720-3724.

238. De Haan, J.B, Parker, M.I. and Gevers, W. *Effects of sodium butyrate on the synthesis and methylation of DNA in normal cells and their transformed counterparts.* **Cancer Res.** (1986) 46:713-716.

239. Graham, F.L. and van der Eb, A.J. *A new technique for the assay of infectivity of human adenovirus 5 DNA.* **Virology** (1973) 52: 450-467.

240. Lowry, O.H., Rosenbrough, N.J., Farr, A.L., and Randall, R.J. *Protein measurement with the Folin phenol reagent.* **J. Biol. Chem.** (1951) 193: 265-275.

241. Birnboim, H.C. and Dolly, J. *A rapid alkaline extraction procedure for screening recombinant plasmid DNA.* **Nucleic Acids Res.** (1979) 7: 1513-1523.

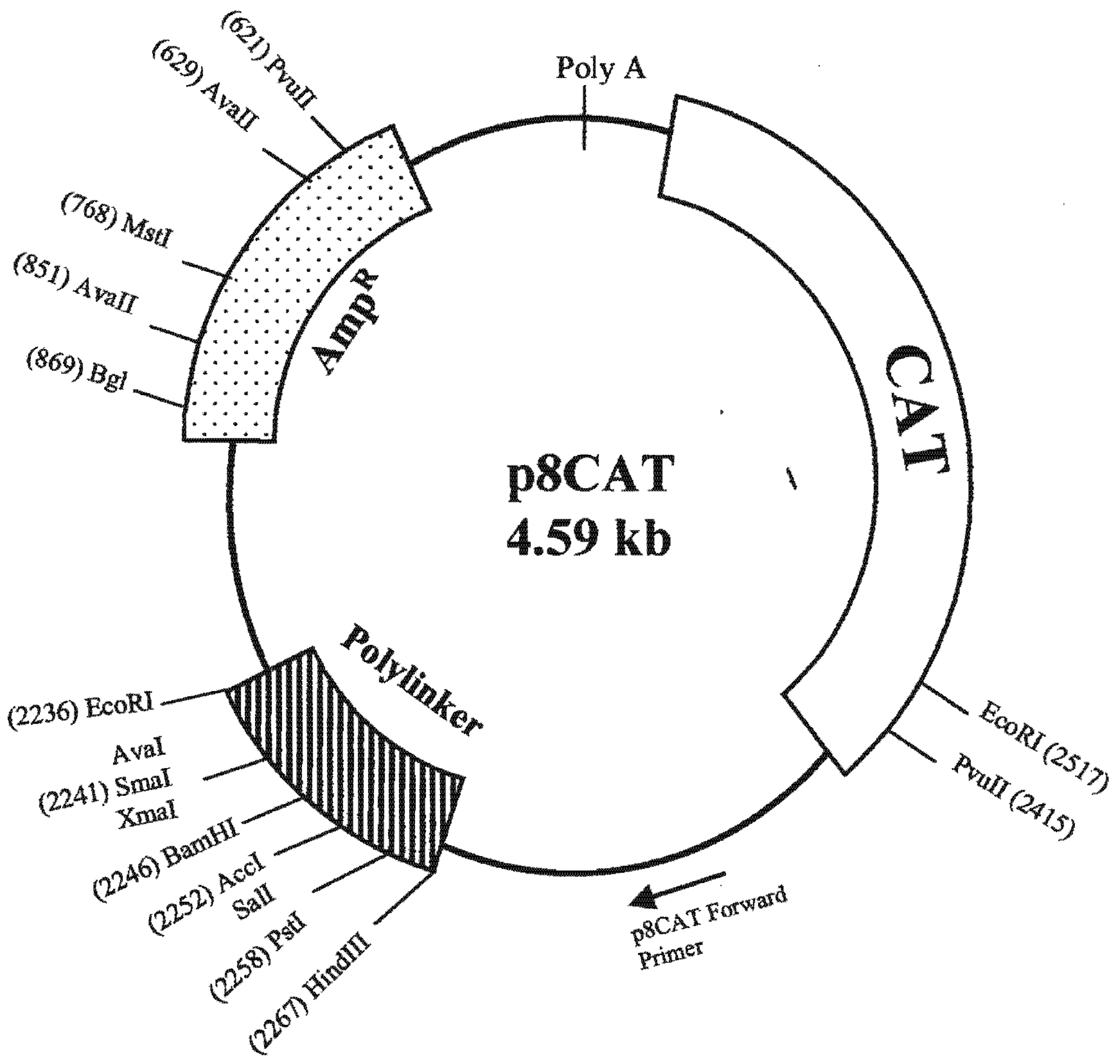
242. Sanger, F., Nicklen, S. and Coulson, A.R. *DNA sequencing with chain-terminating inhibitors.* **Proc. Natl. Acad. Sci. USA** (1977) 74: 5463-5467.

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

244. Sambrook, J., Fritsch, E.F. and Maniatis, T. **Molecular cloning, A laboratory manual** 2nd Edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989).

245. Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. *Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei.* **Nucleic Acids Res.** (1983) 11:1475-1489.

246. Merril, C.R., Goldman, D. and van Keuren, M.L. *Gel protein stains: silver stain.* **Methods in Enzymol.** (1984) 104: 441-447.



Appendix: p8CAT expression vector circle map. The p8CAT vector is a promoterless vector and it contains a CAT reporter gene which codes for the chloroamphenicol acetyl transferase upstream of the polylinker region. The vector also contains an ampicillin resistance gene (Amp^R) which is useful for selective growth in bacteria. The promoter DNA of interest can be cloned into the polylinker region. The p8CAT forward primer described in Table 2.2 is indicated by an arrow.