Dissection Analysis of the Negative Regulation of the Human Urokinase-type Plasminogen Activator (uPA) Gene

Thesis

How to cite:


© 1998 Roberta Benfante

Version: Version of Record
Dissection Analysis of the Negative Regulation of the Human Urokinase-type Plasminogen Activator (uPA) Gene

ROBERTA BENFANTE

A thesis submitted in partial fulfilment of the requirements of the Open University for the degree of Doctor of Philosophy in Molecular and Cellular Biology

July 1998

Department of Biological and Technical Research (DIBIT)
Milan, Italy
To my father
Abstract

The urokinase-type plasminogen activator (uPA) is a serine protease involved in processes such as cell migration and invasion through the activation of the extracellular protease plasmin responsible of the degradation of the proteins of the extracellular matrix. The ability of the cells to migrate and to invade the surrounding tissues plays a fundamental role in many normal and pathological processes such as fibrinolysis, wound healing, angiogenesis, embryogenesis, gametogenesis, ovulation, mammary gland involution and tumor metastasis. Due to its high destructive potential, localized activation of plasminogen is regulated by a complex network of molecular interactions involving both specific inhibitors (PAI-1 and PAI-2), and cell bound receptor (uPAR) and the synthesis of every component of the plasminogen activator system is tightly regulated by a number of factors like hormones, growth factors and cytokines.

On the other hand, the restricted expression of uPA in the organism to a very few cell types (kidney and lung), its inducibility by different stimuli and its overexpression in tumours and several transformed and tumoral cell lines, indicate that uPA regulation occurs also at the level of gene expression.

The gene encoding uPA and its 5' flanking region have been sequenced and characterized in human, mouse and pig. Both in vivo and in vitro studies of progressive 5'deletions of the regulatory region of uPA, have revealed the presence of positive and negative cis-acting sequences and of specific contributions of the proximal regulatory regions to cell-type specific expression of the gene, suggesting the presence of multiple array of cis-acting sequences specific for different transcription factors that direct uPA expression in a cell-type specific manner.

In the human uPA promoter the enhancer is located between -2100 and -1870 from the transcriptional start site and contains two binding sites for the
Abstract

transcription factor AP1 and PEA3 which are important for both the constitutive and regulated expression of the gene. Furthermore, the cooperation between the two AP1 sites is required for TPA inducibility and is mediated by other proteins called Upstream Enhancer Factors (UEF).

A negative cis-acting element has been localized between -1870 and -1570 and contains a binding site for a multiproteic complex (NF-kB/c-rel) that mediates the inducibility by TPA in HeLa and HepG2 cell lines.

Much evidence, however, indicates that the regulation of uPA is mainly negative.

The studies described in this thesis have led to the identification of negative regulatory elements that may play an important role in the establishment of a silenced phenotype in a cell-type specific manner.

The analysis of promoter activity in cell lines not expressing (HeLa and CV1) and expressing uPA (PC3) allowed the identification of at least three regions that play different roles in the silencing of the uPA gene in different cell lines, named S1 (-1870/-1428), S2 (-787/-537) and S3 (-537/-86).

Of the three only S2 shows cell-type specificity, as it is active only in cells that do not express uPA; S1 and S3, on the other hand, could act as modulators of uPA gene expression in those cells that express uPA. All of them negatively regulate the activity of the minimal promoter, suggesting that they can interfere with the formation of a competent pre-initiation complex at the start site of transcription.

A dissection analysis of silencer S2 detected the presence of multiple silencing units, although the deletion of a single one does not have any effect on the activity of the promoter. Transcription from a heterologous promoter is affected only when more than one copy of a single unit is cloned in front of it, suggesting that in the context of the whole S2 region each unit acts synergistically with the others in the silencing of the gene. DNase I footprinting analysis of S2 showed that this region is extensively protected by
nuclear extract from uPA producing (PC-3) and not producing (HeLa) cell lines, with few differences between them. However, the EMSA analysis of the complexes indicated that their molecular composition may be different in the two cell lines tested.

In particular, in HeLa cell extract, a single strand binding activity seems to bind S2 and it may be responsible for the assembly of a negative-acting complex that would prevent the loading of a competent transcription initiation complex at the start site of transcription.

Furthermore, the presence of two binding sites for proteins belonging to the HMG-box containing protein family, suggested that the assembly of a stereospecific complex is required for the activity of silencer S2, although it is not clear if their activity is required for the silencing of the gene or for the inactivation of the inhibition in those cells that express uPA as in HepG2 cell line. In these cells uPA is expressed at a very low basal level, but its expression is inducible by TPA. In this cell line S2 is not active; on the contrary its presence is required for basal expression of uPA.

The presence in S2 of a sequence matching the transforming growth factor β inhibitory element (TIE), described to be important for mediating the inhibition of TPA induction by TGF-β1 of the stromelysin gene, has led me to investigate the role of this element on S2 activity in HepG2 cells. Transient transfection analysis of uPA promoter deletions showed that, at least in this system, the TIE is not involved in the regulation of uPA expression.

Preliminary results, shown in appendix 2, have suggested another interesting aspect of the negative regulation of uPA expression by the p53 tumor suppressor gene product, although no p53 binding to uPA promoter has been shown so far.

In conclusion, the studies presented in this thesis have showed that a complex array of regulatory sequences, in addition to the enhancer and
Abstract

minimal promoter, appear to regulate uPA transcription, some of which are cell-type specific.
1. Introduction.................................................................1

Prologue.............................................................................1

The class II Basal Promoter Apparatus..............................2

The RNA Pol II holoenzyme ("mediator")...........................9

The TAFs ..........................................................................10

The human general co-factors or Upstream Stimulatory
Activity (USA) ....................................................................17

Regulation of RNA pol II transcription.............................21

Positive control ................................................................23

Negative control ................................................................31

Passive repression ..........................................................31

Active repression ............................................................37

Chromatin mediated silencing ........................................42

Silencing at telomeres and at HML and
HMR .................................................................................42

Pc-G ................................................................................45

DNA methylation mediated repression .........................48

The plasminogen activation system ................................49

The activation of plasminogen ........................................49

The plasminogen activators .............................................50

uPA and its receptor .....................................................52

The biological role of uPA .............................................53

Role of uPA in cancer ....................................................55

Transcriptional regulation of the uPA gene ....................57

Inducers of uPA synthesis .............................................57

The uPA gene promoter ................................................58

The minimal promoter ..................................................59
Index

The uPA enhancer element ........................................ 59

The PEA3/AP1 element ........................................ 60

The downstream AP-1 site .................................... 61

The COM region .................................................. 61

NF-kB sites ......................................................... 62

cAMP responsive elements .................................. 62

Negative regulatory sites ...................................... 63

unknown sites .................................................... 64

Chromatin structure of human uPA gene .................. 64

2. Aim of the project ............................................. 67

3. Materials and Methods ...................................... 70

   Cell culture, DNA transient transfection and CAT assay. 70

DNA plasmid construction ..................................... 70

   5' deletions .................................................. 71

Internal deletions ............................................... 71

SV40 derived plasmids ........................................ 72

S2 internal deletions .......................................... 72

Nuclear extract and SouthWestern analysis .............. 74

DNase I protection analysis and Electrophoretic mobility shift assay (EMSA) ........................................... 75

Circular permutation assay ................................... 76

4. Results ............................................................ 78

   Cell type specificity of different portions of the uPA gene regulatory region ........................................ 78

   Multiple silencing activities in the 5' flanking region of the human uPA gene ..................................... 79

   S2 has the properties of a silencer .......................... 82
Index

Silencer S2 is made up of multiple silencing units........... 84
DNasel footprinting analysis of S2................................. 87
EMSA analysis of the complexes present in HeLa and PC3 nuclear extracts ......................................................... 90
  Footprint D/L ................................................................ 90
  Footprint F+M ............................................................ 93
  Footprint N+G ........................................................... 94
  Footprint N ................................................................ 95
  Footprint C+H ........................................................... 96
  Footprint I ..................................................................101

Functional analysis of the TIE element ..........................109
A single-strand DNA-binding activity (footprint H+I) .......111

5. Discussion ..................................................................116
A single-strand DNA binding protein in the huPA promoter ................................................................. 126
The TIE element is involved in S2 silencing activity....... 131
A hierarchy of protein-DNA interaction in uPA promoter S2 region ......................................................... 132

6. Appendix 1: Negative regulation of uPA by TGF-β1 ................................................................. 136
Introduction .................................................................. 136
Materials and Methods .................................................. 140
  DNA manipulations ..................................................... 140
  Cell culture, RNA extraction and Northern blot analysis 141
  Transient transfection analysis ..................................... 142
Results ......................................................................... 143
  Effect of TGFβ1 on uPA expression in PC3 cells .......... 143
## Index

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGF-β1 inhibits TPA induction of uPA mRNA in HepG2</td>
<td>143</td>
</tr>
<tr>
<td>Transient transfection analysis of inhibition of TPA</td>
<td>147</td>
</tr>
<tr>
<td>induced uPA promoter driven transcription</td>
<td></td>
</tr>
<tr>
<td>Discussion</td>
<td>149</td>
</tr>
<tr>
<td>7. Appendix 2: Negative regulation of uPA by p53</td>
<td>151</td>
</tr>
<tr>
<td>Introduction</td>
<td>151</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>152</td>
</tr>
<tr>
<td>Cell culture</td>
<td>152</td>
</tr>
<tr>
<td>Cell cycle studies</td>
<td>152</td>
</tr>
<tr>
<td>RNA isolation and Northern blot analysis</td>
<td>153</td>
</tr>
<tr>
<td>Results</td>
<td>154</td>
</tr>
<tr>
<td>Discussion</td>
<td>156</td>
</tr>
<tr>
<td>8. Conclusions and perspectives</td>
<td>159</td>
</tr>
<tr>
<td>9. References</td>
<td>161</td>
</tr>
</tbody>
</table>
1. Introduction

Prologue

Regulation of gene expression in eukaryotes is involved in a wide variety of tissue-specific (Weissman and Singer, 1991), developmental (Biggin and Tjian, 1989), cell-cycle (Weinberg, 1996) and signal-responsive events (Kerr et al., 1990; O'Brien and Granner, 1991).

The nuclear genes of eukaryotic organisms are transcribed by one of three RNA polymerases. RNA polymerase I (Pol I) is utilised exclusively for the expression of genes encoding the 5.8S, 18S, and 28S ribosomal RNAs (rRNA). The form II enzyme (Pol II) transcribes all genes encoding mRNAs, as well as those that specify certain small nuclear RNAs (snRNAs). All tRNA genes, as well as the genes for 5S rRNA and the remaining snRNAs, are transcribed by RNA polymerase III (Pol III).

All three eukaryotic polymerases maintain considerable sequence similarity in their largest subunit (Allison et al., 1985) and have five subunits in common (Young, 1991). Transcription requires that the polymerases associate with promoter regions and form a stable initiation complex.

Studies on the purified enzymes demonstrated that they are incapable of accurate class-specific promoter recognition, that is accomplished by a set of auxiliary factors distinct for each class of genes.

The promoter structure varies with each gene family showing a different array of DNA sequence elements and configuration. However all promoters have a core sequence element which is recognised by a factor that binds DNA specifically and provides a nucleation site for complex formation.
The class II Basal Promoter Apparatus

Class II genes promoters contain combinations of DNA sequences:

1) a core or basal promoter element defined as "minimal DNA elements that are necessary and sufficient for accurate transcription initiation by RNA pol II in reconstituted cell-free system". The best characterised core-promoter elements are the TATA-box (consensus TATAa/tAa/t; Breathnach and Chambon, 1981), located at about 25 base pairs (bp) upstream of the transcription start site and a pyrimidine-rich (consensus YYANt/aYY) initiator element (Inr) located near the transcription start site (Smale and Baltimore, 1989).

Different promoters may contain both, one or neither of these elements denoted as TATA+Inr+, TATA+Inr−, TATA−Inr+ or TATA−Inr− (Weis and Reinberg, 1992).

2) proximal promoter elements that occur between 50 and 200 bp upstream of cap site and are the site of binding for transcriptional activators able to modulate transcription

3) distal enhancer elements located far from the transcription initiation site in either direction and orientation respect to the start site.

Transcription initiation by Pol II is precisely regulated by transcription factors that interacts with these three classes of DNA targets and also with each other.

Accurate transcription by Pol II is accomplished by the assembly of a multiproteic complex in which the interaction of general transcription factors (GTFs) either with promoter elements or with each other by protein-protein interaction leads to the recruitment of RNA pol II in a competent transcription initiation complex (Weil et al., 1979; Reinberg and Roeder, 1987a,b; Reinberg et al., 1987).

So far 8 general initiation factors (TFII A, TFII B, TFII D, TFII E, TFII F, TFII H, TFII I, TFII J) have been isolated and cloned from human, rat, Drosophila and
Introduction


Studies with purified recombinant components and analysis of intermediates that remain stable to various nuclease-protection and electrophoretic mobility shift assays (Buratowski et al., 1989), coupled to mutagenesis techniques and x-ray crystallography (Burke and Kadonaga, 1996) suggest a stepwise model of pre-initiation complex (PIC) assembly on the AdML promoter in vitro. The pathway can be divided as follows (Figure 1.1):

1) TBP (TATA binding-protein) binds the TATA element through minor groove contacts forming a stable complex that results in a distortion (bend) of the DNA that brings sequences upstream and downstream of the TATA element into a closer apposition (Nikolov and Burley, 1994). Furthermore, as a consequence of this orientation, the upper surface of TBP is exposed and available for protein-protein interaction, with a large number of proteins (Timmers and Sharp, 1991; Meisterernst and Roeder, 1991, Meisterernst et al., 1991). TBP can bind TATA element either alone or in a TFII D multisubunit complex containing TBP-associated factors (TAFs; Tanese et al., 1991). The other factors are then recruited sequentially into the initiation complex in the following order: II A, II B, II F with Pol II, II E, II H and II J. TFIIA, although not essential for formation of a functional PIC, can bind stably to this complex (Buratowski et al., 1989; Maldonado et al., 1990) through direct contacts with TBP and with upstream DNA sequences (Geiger et al., 1996; Tan et al., 1996). TFIIA has also an important role of stabilisation of TBP-DNA interactions in those situation when the affinity of the complex is altered by mutations in the DNA binding of TBP (Imbalzano et al., 1994) or in promoters with weak TATA elements. Furthermore one function of TFIIA is
to remove, by way of its association with TFIID, TBP-bound negative co-factors that associate and negatively regulate TFIID activity and prevent binding of TFIIB to the PIC (see "The human general co-factors"; Dr2; Cortes et al., 1992).

2) Binding of TFIIB, through direct contacts both with TBP and sequences upstream and downstream of the TATA element (Nikolov et al., 1995), has a function in TFIIF-RNA pol II recruitment and in start site selection by RNA pol II (Leuther et al., 1996; Li et al., 1994). Like TFIIA, TFIIB also stabilises TBP-TATA complex (Imbalzano et al., 1994; Kim and Roeder, 1994). Binding of TFIIA and TFIIB to TBP does not show overlapping contacts and no direct contacts with each other.

3) The recruitment of the preformed TFIIF-RNA pol II complex through TFIIB interaction with both TFIIF and RNA pol II, follows the binding of TFIIB (Leuther et al., 1996). TFIIF plays a direct role in promoting targeting of RNA pol II through these interactions, destabilizing the non-specific RNA pol II-DNA interaction to non-specific sites (Flores et al., 1991). TFIIF is also implicated in transcriptional elongation although the subunits involved are different from those implicated in initiation (Kephart et al., 1994; Tanet et al., 1995). The RNA pol II recruited to the PIC is not-phosphorylated on the carboxy-terminal domain (CTD) of its largest subunit (formIIA; Lu et al., 1991). The CTD consists of a heptapeptide (consensus YSPTSPS) repeated 26 times in Yeast, 43 in Drosophila and 52 in man, that becomes extensively phosphorylated during the initiation of transcription (form IIO; Baskaran et al., 1993).

4) Binding of TFIIE, through direct interactions with RNA pol II and TFIIF and TBP (Maxon et al., 1994), just upstream of start site is consistent with the proposed functions in promoter melting (Robert et al., 1996).
Introduction

Fig. 1.1: Model for potential pathway of transcription complex assembly
5) TFIIH entry is the last step of PIC assembly. The direct contacts with TFIIE (Maxon et al., 1994; Okhuma et al., 1995) results in an increased stability of bound TFIIE, important because TFIIE mediates the stimulation of CTD phosphorylation by TFIIH (Lu et al., 1992; Okhuma et al., 1995; Akoulitchev et al., 1995). The region covered by the PIC complex extends to position +30 (van Dyke et al., 1988; Buratowski et al., 1989).

Comparative studies of free versus PIC-bound TFIIB (Nikolov et al., 1995) or DNA-bound TBP suggested conformational changes in factors during PIC assembly. Following the formation of the stable (closed) complex, the PIC is activated to an unstable (open) complex, in the presence of ATP and other ribonucleoside triphosphates (Holstege et al., 1996; Jiang et al., 1996). DNA is melted and transcription initiated, followed by promoter clearance and recycling of PIC components (Zawel et al., 1995). The melting of DNA is sustained by a ~10 bp region just upstream of the start site and is dependent upon TFIIH helicase activity (Drapkin et al., 1994; Serizawa et al., 1993; Roy et al., 1994; Pan and Greenblatt, 1994).

After transcription initiation the melted region in the open complex extends downstream (Holstege et al., 1996). Phosphorylation of RNA pol II CTD is also associated with the transition from initiation to elongation suggesting a role for CTD phosphorylation, by TFIIH kinase activity, in promoter clearance (O’Brien et al., 1994; Yankulov et al., 1996; Goodrich and Tjian, 1994; Payne et al., 1989; Laybourn and Dahmus, 1990).

It has been proposed that upon CTD phosphorylation the initiation complex undergoes a conformational change resulting in reversion of interaction between CTD-TBP (Usheva et al., 1992) and RNA pol II-TFIIE (Maxon et al., 1994) as well as interactions between RNA pol II and cofactors binding CTD, once the accessory functions of these cofactors in activator stimulated PIC assembly or initiation are...
Introduction complete (Koleske and Young, 1995; Barberis et al., 1995; see RNA pol II holoenzyme). At or following termination the CTD is dephosphorylated by a CTD-specific phosphatase. This event is required for the polymerase reentry in the competent form (form IIA) in a new round of PIC assembly (Chambers et al., 1995).

Assembly of PIC on promoters that lack TATA motif requires either the direct or indirect interaction of general transcription factors with other core promoter element, as initiator (Figure 1.2). Several functionally distinct initiator elements are reported to direct transcription from TATA-less promoter (Weis and Reinberg, 1992). In two cases, AdML and the murine terminal deoxynucleotidyl transferase (TdT) promoter initiator element, it has been shown that TFIID is required for the assembly of a functional PIC (Pugh and Tjian, 1991; Zhou et al., 1992). The interaction with initiator elements could be mediated by the binding of TAFs with DNA, with the result to bring TBP to promoter where the assembly of PIC could proceed as described for TATA containing promoter (Kaufmann and Smale, 1994; Purnell et al., 1994). For other initiator element, such as that found in the adeno-associated virus (AAV)P5 promoter, the presence of an initiator binding protein, in addition to TFIID, has been postulated (Beaupain et al., 1990). It has been proposed that YY1, a protein that function both as an activator and repressor, may bind to this initiator. However, transcription in vitro from the AAVP5 promoter requires only YY1, TFIIB and RNA polymerase II, suggesting that YY1 can functionally substitute for TBP in the nucleation of an active transcription complex (Usheva and Shenk, 1994). Other Inr-binding factors have been identified and include TFII I, USF and E2F (Carcamo et al., 1991). Furthermore, in Drosophila TATA-less promoters a downstream promoter element (DPE) has been shown to function synergistically with Inr elements (Burke and Kadonaga, 1996). DPE element mediates Inr-dependent site-specific
binding of TFIID, through TAFII60, to the promoter (Burke and Kadonaga, 1997). In conclusion, PIC assembly pathways might vary depending on specific element and the promoter context, at least with respect to the formation of stable intermediates and general factors requirements, increasing the possibility of selective gene regulation by activators and repressors with core promoter-specific functions (Martinez et al., 1995; Kaufmann et al., 1996).

Fig. 1.2: Assembly of active transcription complex at TATA-less promoters. +1 indicates the start site of transcription; Inr indicates the initiator element and DPE the downstream promoter element.

The failure of \textit{in vitro} reconstituted transcription to respond to activators supports the idea of the existence of intermediary factors. So far three types of intermediary factors have been described:

1) The mediator, which associates with RNA polymerase II to form a holoenzyme;
2) TBP Associated Factors (TAFs), which associate with TATA-binding protein (TBP) to form the general transcription factor TFIID;

3) The Upstream Stimulatory Activity (USA), isolated from human cells, whose characterisation is still in progress.

*The RNA Pol II holoenzyme ("mediator")*

Both genetic and biochemical studies of the *Saccharomyces cerevisiae* RNA polymerase II have led to the identification of a multi-protein complex associated to the carboxy-terminal repeat domain (CTD) of RNA pol II, together with TFIIF, TFIIB and TFIIH, forming what is called the RNA pol II holoenzyme (Kim et al., 1994). This complex, termed the mediator (Hengartner et al., 1995; Koleske and Young, 1994), is required for transcriptional activation *in vitro* as well as *in vivo* and it contains about 20 polypeptides including GAL 11 (Chen S.et al., 1993; Fassler and Winston, 1989), SUG1 (Swaffield et al., 1995) and SRB proteins (suppressors of RNA polymerase B; Nonet and Young, 1989; Thompson et al., 1993). Among the SRB proteins are the product of the essential genes SRB 4, SRB 6 and SRB 7, whose temperature-sensitive mutation leads to an immediate shutdown of all mRNA synthesis at the non-permissive temperature and of two non-essential genes, SRB 2 and SRB 5 that are also important for Pol II transcription as deletion of these genes diminishes transcriptional activity in yeast nuclear extracts.

Studies of genes involved in glucose repression in yeast led also to the identification of proteins in the mediator. Mutations in these genes, termed SSN for suppressors of snf1, are able to suppress growth defects of a mutant lacking the SNF1 protein kinase required to relieve glucose repression of gene expression. Sequence comparison indicated that SSN 5, SSN 2, SSN 3 and SSN 8 are identical to the product of SRB 8, SRB 9, SRB 10 and SRB 11 genes (see Table 1.1). These genes are non-essential and have been suggested to function in repression of many polymerase II promoters. This was shown for two other SSN genes whose
SRB counterpart has not been isolated yet, SSN 4 and SSN 7. They are identical to Sin 4 and Rox 3 which were isolated from screens for mutants affecting repression of the HO and CYC7 promoters (Table 1.1). These, together with other two polypeptides implicated in transcriptional repression, RGRl and GAL11, have been demonstrated to be part of the holoenzyme, suggesting a role for the mediator complex in mediating transcriptional repression as well as activation. The identification of yeast Pol II holoenzyme offers another view for initiation complex formation with the pre-assembled holoenzyme to be recruited to promoters through interaction with the already bound TFIID.

<table>
<thead>
<tr>
<th>SSN GENE</th>
<th>Alternate name</th>
<th>Activity</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSN 1</td>
<td>MIG 1</td>
<td>Transcriptional repressor, maybe target of SNF 1 kinase</td>
<td>Vallier and Carlson, 1994</td>
</tr>
<tr>
<td>SSN 2</td>
<td>SRB 9; SCA 1</td>
<td></td>
<td>Hengartner et al., 1995; Song et al., 1996</td>
</tr>
<tr>
<td>SSN 3</td>
<td>SRB 10; UME 5; cyclin-dependent kinase cyclin</td>
<td>Strich et al., 1989; Surosky et al., 1994; Thompson et al., 1993</td>
<td></td>
</tr>
<tr>
<td>SSN 4</td>
<td>ARE 1</td>
<td>pair to SSN 8</td>
<td>Song et al., 1996; Chen et al., 1993</td>
</tr>
<tr>
<td>SSN 5</td>
<td>SIN 4; TSF 3</td>
<td>role in general transcriptional</td>
<td>Hengartner et al., 1995; Song et al., 1996; Hengartner et al., 1995</td>
</tr>
<tr>
<td>SSN 6</td>
<td>CYC 8</td>
<td>Transcriptional repression of many genes in complex with TUP 2</td>
<td>Schultz et al., 1987</td>
</tr>
<tr>
<td>SSN 7</td>
<td>ROX 3</td>
<td>Essential for viability</td>
<td>Song et al., 1996; Rosenblum-Vos et al., 1991</td>
</tr>
<tr>
<td>SSN 8</td>
<td>SRB 11</td>
<td>Cyclin-dependent kinase cyclin pair to SSN 3</td>
<td>Kuchin et al., 1995; Liao et al., 1995</td>
</tr>
</tbody>
</table>

Table 1.1: The SSN gene family

The TAFs

Different lines of evidence indicate that TFII D plays a central role in mediating activation, functioning as a target for activators, and in directing
nucleation of an initiation-competent complex containing the general factors and RNA polymerase II.

TFII D, in fact, appears to be the only GTF with sequence specific DNA binding activity and there are several lines of evidence suggesting that it is the first component to assemble on the core promoter, directing the recruitment of Pol II and other GTFs (Hernandez, 1993; Pugh, 1996).

TFII D is a multi-protein complex composed by TATA-binding protein (TBP) and a set of associated factors: the Pol II TAFs (TAFIIs).

TBP was first purified from yeast as a single polypeptide of 27 KDa (Hahn et al., 1989). Cloning of the gene allowed the identification of TBP homologues from other eukaryotes and to the discovery of TAFs. The complex purified from human and Drosophila showed that TFII D consists of TBP and at least eight associated TAFs ranging in size from about 18-250 KDa (Dynlacht et al., 1991; Tanese et al., 1991; Chen et al., 1994a).

The human and Drosophila TAFIIs are highly conserved (Chen et al., 1994a; Chiang and Roeder, 1995) and several homologues have been identified in yeast where they are essential for viability (Reese et al., 1994; Poon et al., 1995; Verrijzer et al., 1994). As listed in Figure 1.3 human TFII D contains homologues of seven of the eight Drosophila TAFs (dTAFs); a human homologue of dTAF150 that is missing in hTFII D may correspond to a component of the initiator-element factor (CIF) described by Kaufmann et al (1996). In yeast, TFII D homologue of dTAF110/hTAF130 and the homologue of the human TAF105, found by Dikstein et al (1996a) to be a B-cell specific TAF, are missing.

Although the stoichiometry and the arrangement of TBP-TAFs interaction in TFII D is not known yet, it has been shown that Drosophila TAF250 is critical in the assembly of TFII D (Chen et al., 1994a) and in the stabilisation of TBP-TATA interaction.
Other TAF\(\text{II}\)s interact directly with TBP and with each other allowing the formation of multiple combinations able to mediate the integration of different signals to the basal transcription machinery (Burley and Roeder, 1996).

Different biological activities have been ascribed to TAF\(\text{II}\)s leading to the view of TFII D as a highly versatile complex:

- TBP-TAFs complex is involved in **promoter recognition**. TBP is essential for transcription by all three eukaryotic RNA polymerases, as revealed by studies with mutated TBP (Cormack and Struhl, 1992; Schultz et al., 1992). The specificity for a distinct class of genes is to address to its association with different sets of TAFs, able to discriminate between different core promoter structure. So different TBP-TAF complexes are thought to direct class-specific promoter recruitment of the three RNA polymerases (Hernandez, 1993; Table 1.2).

While Pol I and Pol III require TBP-TAF association for transcription of class I and III genes, TBP alone can sustain basal transcription from TATA-containing class II gene. However Pol II TATA-less promoters require TAF\(\text{II}\)s for transcription. While TBP recognises the TATA-box, other downstream elements are recognised by different TAFs allowing TFII D to bind promoter elements, either individually or in combination, and to activate different array of promoters. Furthermore the association with an Inr binding protein and the presence of other DNA sequences, often located downstream of the transcription start site, as the downstream promoter element (DPE; Burke and Kadonaga, 1996), can modulate basal promoter strength by determining promoter strength and the start site of transcription (Verrijzer et al., 1994; Kaufmann et al., 1996). In the presence of TATA-box and Inr elements an efficient nucleation process requires minimally a TBP-dTAF\(\text{II}250\)-dTAF\(\text{II}150\) complex, whereas in the absence of Inr these TAF\(\text{II}\)s can inhibit the TATA-binding activity of TBP. So, depending on the
promoter architecture, TAFIIIs can modulate the stability of TFIID promoter interactions (Verrijzer et al., 1995; Hansen and Tjian, 1995).

<table>
<thead>
<tr>
<th>Polymerase</th>
<th>Class of genes</th>
<th>Core element</th>
<th>TBP/TAF complex</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA Pol I</td>
<td>rRNA genes</td>
<td>UCE/UBF</td>
<td>SL1</td>
<td>Comai et al., 1992</td>
</tr>
<tr>
<td>RNA Pol II</td>
<td>protein coding genes</td>
<td>TATA/TBP</td>
<td>TF IID TAFs</td>
<td>Hernandez, 1993</td>
</tr>
<tr>
<td>snRNA</td>
<td>PSE</td>
<td>SNAPc</td>
<td></td>
<td>Sadowski et al., 1993</td>
</tr>
<tr>
<td>RNA Pol III</td>
<td>5S RNA</td>
<td>Box C/TFIIIA</td>
<td></td>
<td>Hernandez, 1993</td>
</tr>
<tr>
<td>hU6</td>
<td>PSE</td>
<td>TATA</td>
<td>TFIIIB</td>
<td>Tjian and Maniatis, 1994</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bartholomew et al., 1991</td>
</tr>
</tbody>
</table>

Table 1.2: Promoter recognition properties of different TBP/TAF complexes

- Another function ascribed to TFIID is the alteration of **promoter topology** as revealed by DNsasel protection analysis on the AdML promoter. The analysis of the sequence of dTAFII30α, dTAFII40 and dTAFII60 and of their human homologues revealed that they are related in sequence and structure to the histone proteins H2B, H3 and H4 respectively (Burley and Roeder, 1996) supporting the hypothesis that these TAFs bind DNA in a core-histone manner regulating promoter topology and hence transcription by multiple contacts of TFIID with core promoters, or facilitating basal-factor contacts, or mimicking a nucleosome. This latter hypothesis can explain the observation made by Segil et al. (1996).
<table>
<thead>
<tr>
<th>TFB</th>
<th>D. melanogaster</th>
<th>H. sapiens</th>
<th>S. cerevisiae</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBP</td>
<td></td>
<td></td>
<td></td>
<td>TATA-box binding subunit</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>43</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>250/230</td>
<td>250</td>
<td>130/145</td>
<td>equivalent to CCG1, HMG box, Bromodomains</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td></td>
<td>150</td>
<td>downstream promoter contacts</td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>130</td>
<td></td>
<td>contacts Sp1 (Q-rich activator)</td>
</tr>
<tr>
<td></td>
<td>80/85</td>
<td>100</td>
<td>90</td>
<td>contacts TFIIA</td>
</tr>
<tr>
<td></td>
<td>60/62</td>
<td>80</td>
<td>60</td>
<td>WD-40 repeats</td>
</tr>
<tr>
<td></td>
<td>40/42</td>
<td>31/32</td>
<td>17</td>
<td>histone H4 similarity</td>
</tr>
<tr>
<td></td>
<td>30α/28</td>
<td>20</td>
<td>68/61</td>
<td>contacts TFII E, TFII F</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>15</td>
<td></td>
<td>histone H3 similarity</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>contacts acidic activators (VP16, p53)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>contacts TFII B</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>histone H2B similarity</td>
</tr>
<tr>
<td>TAFIIs</td>
<td></td>
<td></td>
<td></td>
<td>contacts multiple activators</td>
</tr>
<tr>
<td></td>
<td>30β</td>
<td>28</td>
<td>40</td>
<td>contacts estrogen receptor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>68</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>53</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>25/23</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>47</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>19</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>105</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1.3: Evolutionary conservation and properties of TFIIID subunits from human, Drosophila and yeast. Where a particular TAF has been assigned different molecular weights, all molecular weights are listed. dTAFII30a/28 and dTAFII22 and their human homologues arise via alternative splicing (adapted from Burley and Roeder, 1996).
that during mitosis 10% to 20% of TFIID is retained in the condensed chromosome. This group hypothesised that the nucleosome-like structure of TFIID permits its incorporation into condensed chromosome marking promoter DNA for reactivation after mitosis.

- TFII D shows catalytic activity, retaining the capability to acetylate histones and to phosphorylate the basal-factor TFII F. Acetylation of histones changes the packaging of DNA in the nucleosome by weakening histone-DNA interaction, allowing access of activators and general factors to DNA. The histone acetylase activity (HAT) resides in TAFII250 (Mizzen et al., 1996). This activity is conserved from yeast to human, conferring to TFII D an important role in controlling access of nucleosome-bound promoter sequences to the basal machinery in vivo. Furthermore TAFII250 displays a kinase activity (Dikstein et al., 1996b) which can phosphorylate itself and the basal factor TFII F. It is not clear yet if the phosphorylation of TFII F influences activator dependent recruitment of RNA pol II into the PIC or affects the initiation/elongation properties of the polymerase.

- Although TBP can replace TFII D in basal transcription, it fails to support activated transcription leading to the coactivator model, that assumed that at least some of the subunit of TFIID serve functionally to link transcription activation domains with the basal transcription complex (Pugh and Tjian, 1990; Dynlacht et al., 1991). A series of experiments aimed to reconstitute dTFIID from its recombinant subunits, showed that addition of TAFs to TBP restored activation confirming that most if not all transcriptional activators require TAFIIs for activation at least in vitro (Chen et al., 1994a). These experiments led to the notion that activators exert their action through direct interaction with TAFIIs, allowing TFIID to integrate multiple signals from different regulators.

From these studies, aimed to identify TFII D-activators interaction, it came out that different classes of activation domains (such as acidic, Gln-rich, Ile-rich)
bind distinct TAFII s in the TFII D complex. A list of TAFII s/activators interaction is reported in table 1.3.

<table>
<thead>
<tr>
<th>TAFII</th>
<th>Activator</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>dTAFII110</td>
<td>Sp1</td>
<td>Chen et al., 1994a</td>
</tr>
<tr>
<td></td>
<td>Bicoid</td>
<td>Jacq et al., 1994</td>
</tr>
<tr>
<td>dTAFII40</td>
<td>VP16</td>
<td>Sauer et al., 1995</td>
</tr>
<tr>
<td></td>
<td>p53 (acidic)</td>
<td></td>
</tr>
<tr>
<td>dTAFII150</td>
<td>NTF-1 (Ile-rich)</td>
<td>Chen et al., 1994a</td>
</tr>
<tr>
<td>dTAFII60</td>
<td>Hunchback</td>
<td></td>
</tr>
<tr>
<td>hTAFII30</td>
<td>estrogen receptor</td>
<td>Thut et al., 1995</td>
</tr>
<tr>
<td>hTAFII55</td>
<td>CTF; Pro-rich</td>
<td></td>
</tr>
</tbody>
</table>

Table 1.3: TAFII s-activators interaction

Data from *in vitro* studies are supported by *in vivo* experiments on a temperature-sensitive (ts) Hamster cell line (ts-13), carrying an amino acid substitution in TAFII250 (Sekiguchi et al., 1991), where the defect affected only a subset of genes at the non permissive temperature (Wang and Tjian, 1994). These results confirmed the idea that individual TAFII s may be required only by a subset of activators in the cell.

The current model for TAFII s enhancing of transcription is that activators-TAFII s contacts lead to increased recruitment of TFII D to the core promoter (Pugh, 1996). This model would account for the synergistic activation by multiple activators, as demonstrated by experiments on *Drosophila* activators Bicoid and Hunchback (Sauer et al., 1995). Each activator alone gives a modest level of activation. Synergism requires both dTAFII110 (that binds Bicoid) and TAFII60 (binding Hunchback) in the TFIID complex, allowing simultaneous contacts between activators and TFIID. This hypothesis is also supported by *in vivo* studies.
(Colgan and Manley, 1992; Klein and Struhl, 1994), which suggest that TFIIID binding can be rate-limiting in the cell and activators can affect promoter occupancy (Figure 1.4).

\[\text{Fig. 1.4: Model of synergistic transcriptional activation by multiple activators-TAFIIIs contacts}\]

**The human general co-factors or Upstream Stimulatory Activity (USA)**

Studies aimed to determine the minimal requirements for activation *in vitro* suggested the involvement of other cofactors in addition to the TAFs and the general machinery.

Biochemical fractionation of mammalian cellular extracts and reconstitution of a functional activator-driven transcription led to the identification of a set of co-factors, distinct from TAFs, termed the human general co-factors (Meisterernst *et al.*, 1991).

In the beginning these co-factor have been isolated as a crude fraction able to enhance stimulation of transcription by activators and termed USA (*upstream stimulatory activity*). Later it was shown that this fraction contains both positive and negative activities termed PCs and NCs, respectively (positive and negative co-factors).

Members of the PCs are PC1, PC2, Dr2 (D repressor 2)/PC3, ACF (activating co-factors) and PC4, CofA (co-factor A), PC 5, PC 6 and HMG-2 (Meisterernst *et al.*, 1991).
1991; Kretzschmar et al., 1994; Merino et al., 1993; Kretzschmar et al., 1993; Ge and Roeder, 1994a,b; Maldonado and Reinberg, 1995; Zawel et al., 1995; Halle et al., 1995; Shykind et al., 1995; Stelzer et al., 1994; Lieberman, 1994). After the cloning of the gene encoding the co-factor PC4 it was shown that it interacts both with TFIIA and the activator domain of the acidic activator VP16. By interaction with TFIIA a role for PC4 in recruitment of TFIIID to the TATA-box has been ascribed, functioning as a bridge between activator and basal machinery (Figure 1.5a).

On the other hand it has been suggested that the co-factors might stimulate the interaction between the activators and the basal machinery by altering the topology of the promoter (i.e. HMG proteins; Shykind et al., 1995; Stelzer et al., 1994) and affecting the accessibility of DNA in the chromatin. Other PCs are involved in processes following the binding of TFIIID (i.e. PC2 and PC5; Halle et al., 1995).

Among the negative co-factors are Ada/Mot1 (Auble et al., 1994) and the negative co-factors NC1 and NC2 (Meisterernst and Roeder, 1991; Meisterernst et al., 1991; Inostroza et al., 1992). Ada/Mot1 releases TBP from DNA and thereby represses transcription. NC1 binds TBP in a manner that competes with TFIIA and results in repression of basal but not activator-dependent transcription (Meisterernst and Roeder, 1991). NC1 has been identified to be HMG-1 (Ge and Roeder, 1994b).

NC2 binds TBP and represses basal transcription by competing with TFIIB and TFIIIB thus inhibiting the assembly of the initiation complex (Figure 1.5b). The protein consists of two subunit (NC2α and NC2β/Dr1; Inostroza et al., 1992; Goppelt et al., 1996; Mermelstein et al., 1996), also defined as the repressor-co-repressor complex Dr1-DRAP (Mermelstein et al., 1996; Kim et al., 1997). Furthermore NC2 displays homology with the histones fold motif of histones
H2A and H2B and it has been shown to be necessary for the dimerization of the two subunits of NC2 (Goppelt et al., 1996; Mermelstein et al., 1996).

NC2 is conserved between yeast and man suggesting for it an essential role in controlling the overall basal activity in the cell, maybe through mechanisms aimed to modulate the topology of transcription complexes during initiation and elongation of transcription in a chromatin folded DNA.
In some cases, the same cofactor has been identified, by different laboratories, either as positive or negative cofactor (see PC3/Dr2; Merino et al., 1993; Kretzschmar et al., 1993). This discrepancy can be explained by the fact that some negative factors repress basal transcription, but this repression is overcome by activators so that the net fold activation is greater in their presence than in their absence. Furthermore, PCs repress basal transcription at higher concentrations.

The identification of three different groups of cofactors, the human general co-factors, the TAFs and the RNA polymerase II holoenzyme-associated mediator, sharing the same function of activation of transcription, suggested that they might cooperate functionally to support activators function, possibly depending on the organism and the specific gene. It has been proposed that the mediator generally function in yeast to facilitate recruitment of the RNA polymerase holoenzyme, although recent reports describe the identification of a mammalian RNA polymerase II holoenzyme (Kim et al., 1994; Chao et al., 1996; Maldonado et al., 1996). This holoenzyme would cooperate with the human general co-factors by the interaction between these and TAFs (Kretzschmar et al., 1994; Ge and Roeder, 1994a).

The distinction, at least in vitro, between yeast and the higher eukaryotes in the requirement of TAFs for activated transcription made necessary to establish the role of yeast TAFs in vivo.

While analysis of mutant TAF genes revealed that they are essential for cell viability, experiments of TAFs depletion showed that cell growth rapidly stopped but transcriptional activation of actively transcribed genes, induced subsequent to TAF depletion, was not affected. Together these results led to the conclusion that TAFs are not generally required for transcriptional activation in yeast (Walker et al., 1996; Moqtaderi et al., 1996).
It is possible that the mechanism of gene regulation in higher eukaryotes had acquired additional complexity as the result of the larger number of genes that need to be regulated and as the activation in metazoan responds both to developmental and physiological signals. It has been then hypothesised that TAFs evolved to mediate alternative pathways of promoter recognition independent of the TATA motif and also to acquire a coactivator function in response to the increased requirement for expression and regulation of more complex genomes.

**Regulation of RNA pol II transcription**

Eukaryotic genes are expressed *in vivo* at very low levels (or not at all). Regulation of gene expression upon developmental and physiological stimuli is mediated by the action of a large number of regulatory proteins that recognise specific sequences (enhancers and silencers) located upstream or downstream the core promoter. The complexity of the combinatorial interaction of factors involved in the formation of the transcription initiation complex and the multiple steps that can be rate-limiting, provides several targets of regulation of polymerase activity by factors able to activate or repress transcription. Among these are a) the recognition of promoter sequences by TFIID; b) the recruitment of TFIIB, essential for RNA pol II entry and c) the escape of RNA pol II from the promoter, controlled by TFIIH. The specificity of regulation of gene expression by a relatively small number of transcription factors is provided by the modular structure of both regulatory proteins and the enhancers (silencers): a typical transcription factor is infact constituted by a DNA binding-domain, a multimerization domain, allowing the formation of homo- and heterodimers, and a transactivation domain. These domains can be combined in different fashion giving raise to new sets of factors; the enhancers show a combination of different transcription factors binding-sites, whose arrangement provide the
potential to create unique nucleoprotein complexes forming heterodimers within and among families of transcription factors (Tijan and Maniatis, 1994).

Recent studies have shown that the activity of some enhancers (or silencers) requires proteins that function as architectural components suggesting that activation (repression) depends on the assembly of a highly specific three-dimensional nucleoprotein complex, facilitating multiple protein-DNA and protein-protein interactions (reviewed in Tijan and Maniatis, 1994 and Werner and Burley, 1997; Figure 1.6). Examples of such proteins are the members of the high-mobility group (HMG) proteins that have the properties of inducing a sharp bend in the DNA helix (Grosschedl et al., 1994). Some of them binds DNA non-specifically (HMG-1 and UBF; Paull et al., 1993; Jantzen et al., 1990). Others, as LEF-1 (lymphoid enhancer factor 1; Giese et al., 1991) and SRY (sex determining factor; Van de Wetering and Clevers, 1992) bind in a sequence-specific manner their target promoter. In the case of LEF-1, several experiments on the activation of the T-cell receptor α enhancer provided a model for its function as architectural protein. Upon its binding to the TCRα enhancer a sharp bend is induced in the DNA helix allowing the interaction of other activators that bind sequences flanking the LEF-1 binding site on either side (Giese et al., 1995).

The product of HMG-I(Y), which is unrelated to HMG-domain proteins, plays the same role as LEF-1 in the virus-inducible enhancer of the β-interferon gene but, besides altering DNA structure, it also directly contacts both NF-kB and ATF-2, the two activator proteins, inducing conformational changes that augment their interaction with DNA and with each other (Thanos and Maniatis, 1992; Du et al., 1993).

On the other side, YY1 represses c-fos promoter by bending DNA and then preventing interactions between the CREB-activator protein, whose binding site is
located upstream of YY1, and components of the basal transcription machinery (Natesan and Gilman, 1993).

Protein-induced changes in DNA structure may also involve the formation of three-dimensional DNA loops, with proteins that mediate tight wrapping of DNA through extended protein-DNA contacts. The best example is given by the nucleosome and several examples of positioned nucleosomes that help cross-talk between enhancer element and promoters, or juxtaposition of binding site otherwise located far apart, have been described (Perlmann and Wrange, 1988; Thomas and Elgin, 1988; Schild et al., 1993). Furthermore several DNA binding proteins that mimic nucleosome-like structure are involved in the formation of nucleosome-like particles (Putnam et al., 1994).

Finally, the observation that, on some promoters, TFIID or Pol II itself is promoter-engaged before activation implies alternative pathways for transcriptional control (Rougvie and Lis, 1988; Chen et al., 1994b).

Positive control

Activators work enhancing the efficiency with which the basal transcription complex is assembled in an active form by contacting multiple components of the basal transcriptional machinery. As it is already outlined in the "Basal promoter apparatus" section, interactions between activators and the components of the basal pol II machinery can be either direct between transcription factors and the GTFs as well as mediated by coactivators (mediator, TAFs, UAS; Triezenberg, 1995). The regulation can occur at the level of 1- stimulation of basal transcription apparatus, 2- by counteracting chromatin-mediated repression, 3- by affecting the efficiency of elongation process.
Fig. 1.6: Models for architectural roles of proteins in the assembly of higher-order nucleoprotein complexes.

A: The TCRα enhancer complex. B: β-interferon enhancer complex. C and D: wrapping of DNA around a histone core leads to juxtapose two binding sites recognised by the same transcription factor, i.e. the glucocorticoid receptor (GR, panel C) in the long terminal repeat of the MMTV, on the same surface of the histone core or to facilitate interaction between enhancer and promoter elements (panel D; i.e. the vitellogenin gene promoter, Schild, 1993). For further explanations see text. The figure has been adapted from Grosschedl, 1995.

1- Direct physical interactions have been demonstrated between acidic activators and TBP (Inglese et al., 1991), TFIIB (Lin and Greene, 1991; Roberts et al., 1993), and TFIIF (Xiao et al., 1994). Multiple lines of evidence suggest that the targeting of TFIIB by activators is especially relevant (Hahn, 1993). The association of TFIIB is a rate-limiting step in PIC assembly (Lin and Greene, 1991; Ha et al., 1993). This is reasonable since TFIIB interacts with other components of the initiation complex as TBP, RNA pol II and TFIIF (Ha et al., 1993). Binding of activators as VP16 would induce a conformational change in TFIIB so to expose
Furthermore, as TFIIIB is released from the complex when RNA pol II exits the promoter, a possible function of activators could be to increase the rate of TFIIIB reassociation or retain TFIIIB at the promoter, maintaining the promoter in a state accessible to the polymerase and capable of multiple initiation events. Direct interaction between TFIIH and activators can be involved in the modulation of TFIIH enzymatic activities, important for transcription initiation (Xiao et al., 1994; Tong et al., 1995).

In vivo footprinting studies have shown that many transcriptional activators binding sites appear to be occupied simultaneously (Rigaud et al., 1991). This is consistent with the view that two or more activators can exert synergistic effects on transcription through concerted interactions with multiple components of the PIC (Metz et al., 1994; Chi et al., 1995; Ge and Roeder, 1994a,b). As discussed above (see "The TAFs"), it has been recently demonstrated that synergy between two different activators (Bicoid and Hunchback) bound to the same promoter, results, at least in part, from specific interactions with two distinct dTAFIIs that enhance TFIID recruitment (Sauer et al., 1995).

2- Another level of regulation of transcription is the chromatin packaging of DNA in cells. Chromatin is a nucleoproteic complex whose unit is composed of an octamer of four different histones (H2A, H2B, H3 and H4), termed the nucleosome, and 146 base pairs of DNA wrapped around it. Nucleosomes repress transcription by blocking access of transcription factors to DNA (Felsenfeld, 1992; Adams and Workman, 1993). Chromatin structure on promoters has been shown to be dynamic, so that it can be altered during regulatory events in the absence of DNA replication (Schmid et al., 1992; Becker, 1994). It has also been shown that changes in chromatin structure accompany induction of tissue-specific genes during development and differentiation (Gross and Garrard, 1988). In vitro transcription analysis on reconstituted chromatin showed that repression by
nucleosome can be prevented by incubating DNA with TFIID prior to nucleosome assembly (Workman and Roeder, 1987). It has been suggested that activators alleviate chromatin mediated repression by changing the nucleosome structure, allowing binding of GTFs to the core promoter (Workman et al., 1988; Taylor et al., 1991; Workman et al., 1991).

A conserved multiprotein complex of ~ 2 MDa (the SWI/SNF complex; Peterson et al., 1994; Côté et al., 1994) with the function of mediating changes in chromatin structure, has been identified by genetic screens and biochemical analysis in yeast, Drosophila (Tamkun et al., 1992; Elfring et al., 1994) and mammals (Kwon et al., 1994; Wang et al., 1996).

Genetic studies of transcriptional regulation in the yeast Saccharomyces cerevisiae led to the identification of a number of SWI (refer to yeast mating-type switching; Stern et al., 1984; Breeden and Nasmyth, 1987) and SNF (abbreviation for sucrose non fermenting, Neigborn and Carlson, 1984) genes as positive regulators of a subset of promoters. The observation that mutation either in SWI or SNF genes lead to a reduced activation of transcription suggested that both SWI and SNF proteins could be components of the same protein complex, as confirmed by co-purification through multiple chromatographic steps. The purified SWI/SNF complex contains SWI1, SWI2/SNF2, SWI3, SNF5, SNF6 and SNF11 in addition to five polypeptides (p78, p68, p50, p47 and p25; Cairns et al., 1994; Côté et al., 1994).

A relation between SWI/SNF protein function and disruption of chromatin structure has come from two observations:

a) mutations in genes encoding chromatin components (such as the core histones) suppresses the defects in growth and transcription due to swi and snf mutations (Hirschhorn et al., 1992),
b) the absence of either SWI2/SNF2 or SNF5 results in a decreased transcription and alteration of chromatin structure in the promoter region of SUC2 gene and these effects can be overcome by a reduction in the level of the core histones H2A and H2B (Hirschhorn et al., 1992). This observation also suggests that SWI and SNF protein might facilitate the function of transcriptional activators by antagonising the repressive action of chromatin. The SWI/SNF complex contains a DNA-stimulated ATPase activity (SWI2/SNF2 subunit; Laurent et al., 1993). Experiments of binding of GAL4 activator to nucleosome reconstituted DNA in the presence of SWI/SNF purified complex and ATP suggested that SWI/SNF complex has a role in destabilizing histone-DNA interaction increasing the binding of transcription factors to the histone-bound DNA (Côté et al., 1994), although the exact nature of this structural change is not known yet. The SWI/SNF complex, at least concerning its function, is conserved in higher eukaryotes. Functional homologues of SWI2 have been identified in Drosophila (brahma; Tamkun et al., 1992) and humans, where there are two genes closely related to brm and SWI2 (Khavari et al., 1993; Muchardt and Yaniv, 1993; Chiba et al., 1994): BRG1 (brm/SWI2-related gene 1 or hSNF2α) and hbrm (hSNFβ). They are both involved in transcriptional activation in vivo and, using an antibody against BRG1, Kwon et al (Kwon et al., 1994) identified in HeLa cells two different protein complexes, termed hSWI/SNFA and hSWI/SNFB, with chromatin remodelling activity in vitro, as good candidates for the human counterpart of the yeast SWI-SNF complex. Also a human gene (INII1) shows similarity to ySNF5 and it has been shown to associate with BRG1 and co-fractionate with hSWI/SNF (Kalpana et al., 1994). These data suggest an evolutionary conserved function of the SWI-SNF complex in disrupting nucleosome structure in an ATP dependent manner, leading to increase binding of transcription factors. Another question to address concerns the targeting of
SWI/SNF complex to the promoter. There are no clear evidences that activators can exert this function, although interaction between the retinoblastoma gene product Rb and SWI/SNF has been reported (Dunaief et al., 1994). Cofractionation studies, aimed to characterise factors that copurify with the RNA pol II holoenzyme in yeast, have however out lighted that the SWI/SNF complex can be found, although not always, associated to the holoenzyme through interactions with the mediator proteins (Wilson et al., 1996). So SWI/SNF complex could be targeted by associating with the holoenzyme.

Another activity with a function of ATP-dependent chromatin remodelling, leading to increased transcription factors binding, has been purified by Wu et al from Drosophila embryo extracts and termed NURF (nucleosome remodelling factor; Tsukiyama and Wu, 1995). NURF is composed by 4 polypeptides one of which (ISWI) shows a significant identity to SWI2/SNF2 at the level of ATPase domain (Tsukiyama and Wu, 1995). A human gene, hSNF2L, also shows 75% identity to ISWI over its entire length appearing as the human homologue of ISWI (Okabe et al., 1992). A putative yeast homologue has been identified (YB95; Tsukiyama et al., 1995).

From these findings it can be hypothesised that NURF and SWI/SNF complexes represent two distinct classes of ATP-dependent nucleosome reorganising activities that are used in different circumstances and from different subsets of genes (Fig. 1.7).
Fig. 1.7: The SWI/SNF and NURF complexes. The ATP-dependent activities of these complexes enhance the accessibility to DNA of transcription factors (Pugh, 1996).

An alternative mechanism for controlling transcription in a chromatin environment (reviewed in Wade et al., 1997) is provided by enzymes able to induce post-translational modification in the histones. Several transcriptional regulators, besides TAFII250, have been found to act as enzymes that acetylate histones. In *S. cerevisiae* transcription factors with acidic activation domains recruit a trimeric coactivator complex, ADA2-ADA5-GCN5 (Guarente, 1995) which contacts the basal transcription machinery. In higher eukaryotes numerous transcription factors, including steroid receptors, use the coactivator p300/CBP in association with another factor P/CAF. These factors have catalytic function as acetyltransferase and act by modifying the amino-terminal tails of particular histones (Brownell et al., 1996; Ogryzko et al., 1996), with the result of destabilization of nucleosome and increased access of components of basal transcription machinery to the promoter (Lee et al., 1993). The activity of acetyltransferase-associated factors is balanced by the presence of regulators that deacetylate histones, resulting in a repression of transcription (Ayer et al., 1995;
Taunton et al., 1996). If persistence of gene activity requires continuous activity of the acetyltransferase, histone acetylation would represent a mechanism to modulate transcription (Figure 1.8).

![Diagram](image)

Fig. 1.8: Acetylation and deacetylation of histones change nucleosome structure, influencing transcription. Several coactivators contain an histone acetyltransferase activity that facilitate the disruption of nucleosomes and recruitment of the basal transcription machinery (A). The presence of histones deacetylases can reverse the action of these coactivators, leading to a fine regulation of transcriptional activity. Ac indicates acetylation.

Other post-transcriptional modifications of histones can modulate the stability of nucleosome, as phosphorylation, ubiquitination, methylation or ADP-ribosylation (reviewed in Matthewes and Watyerborg, 1985; Bradbury, 1992). Transcriptional regulators containing such activities have been identified.
3) Preliminary studies indicate that activators may influence the role that TFIIH plays during promoter clearance. Complexes containing TF IIIE/IIH pause between +12 and +18, after which a large fraction of them resumes elongation. It has been shown that activators affect the efficiency of elongation by decreasing the extent of abortive transcription and increasing the proportion of complexes that overcome the +12 to +18 stalling (Yankulov et al., 1994).

Negative control

Negative regulators of eukaryotic gene expression can inhibit transcription by interfering with any step in transcription initiation. Several mechanisms of repression have been proposed, some of which are referred as passive (involving or interfering with activator proteins function) and active (involving direct effect of repressor proteins through specific repressive domains on the general transcription complex).

Passive repression

Blocking of activator function is realised by different mechanism (Figure 1.9):

a) Interference with activator nuclear localisation. This is one of the earliest steps at which a factor can interfere with the activity of a transcriptional activator, inhibiting the transport of the activator from the cytoplasm into the nucleus. This is obtained by the formation of a complex between a repressor protein and the positive factors in which the nuclear localisation signal of the activator protein is masked by the binding of the repressor, preventing nuclear import (Beg et al., 1992). One example of such mechanism is described by the Rel/NFkB family of transcription factors (reviewed in Blank et al., 1992). Different inhibitor proteins of the IkB family can interfere with the activities of different sets of Rel dimers (Zabel and Baeuerle, 1990; Geisler et al., 1992; Inoue et al., 1992; Kerr et al., 1992). Upon induction by several stimuli, that lead to NF-kB activation, IkBs are
phosphorylated and degraded by the ubiquitin-proteasome pathway, allowing NF-
kB to translocate to the nucleus and to activate the target genes (reviewed in
Thanos and Maniatis, 1995).

Another example comes from the glucocorticoids receptor (GR) that, in the
absence of its ligand, is also retained in the cytoplasm in a complex which includes
the 90 KDa heat shock protein, hsp 90 (Pratt et al., 1992).

b) Interference with the assembly of multisubunit activators. Some
repressors work by competing for association with one of the activator subunits,
preventing the formation of a functional activator. Many activators bind DNA
either as homo- or hetero-multimeric complexes. Such interactions are mediated
by specialised dimerization domains such as the leucine zipper (Landschulz et al.,
1988; Vinson et al., 1989), the POU domain (reviewed in Wegner et al., 1993), the
basic region helix-loop-helix (Murre et al., 1989a,b), and the ability to form
different complexes within a family of related protein gives an enormous source
of regulatory targets, since different dimeric complexes may differ in DNA binding
and in transactivation properties (see for examples Descombes and Schibler, 1991;
Ron and Habener, 1992). This kind of interference can be exerted in two ways: i) a
negative acting factor can form a complex which fails to bind DNA so that the
repressing activity is due to titration of positive transcription factors. An example
comes from families of bH-L-H factors and their helix-loop-helix negative-acting
coregulators, Id and extramacrochaete (emc), which, although able to dimerize
with other H-L-H factors, lack a DNA binding domain (reviewed in Jan and Jan,
1993).
Fig. 1.9: Models of negative control

A: Interference with activator nuclear localisation; B: Interference with assembly of multisubunit activators; C: Interference with activator DNA binding; D: Interference with activity of DNA-bound activators; E: Interference with access of general transcription machinery to DNA. Among the mechanisms of active repression are the interference with pre-initiation complex assembly, the co-repressor model and the positioning of nucleosomes over the TATA-box with recruitment of silencing proteins, as Pc-G, with the formation of a closed compacted chromatin structure.
Similarly the I-POU domain protein (Inhibitory POU; Treacy et al., 1991) and the CHOP-10 (C/EBP-homologous protein; Ron and Habener, 1992) down-modulate the DNA binding activity of their respective dimerization partners; ii) the repressor can form a complex that is able to bind DNA, but lacks domains required for transcriptional activation. This type of regulation involves two steps: competition for binding site and sequestration of a transcription factor in a non functional complex. An example is given by the proto-oncogene product c-myc that activates transcription when it is complexed with its partner max (myc auxilliary factor; Amati et al., 1993; Littlewood et al., 1992). When present in excess, max is a negative regulator of transcription by forming homodimers that bind DNA but do not activate transcription (Amin et al., 1993; Gu et al., 1993). Other examples have been described including two members of the CREB family (CREM and CREB-2; Foulkes et al., 1991; Karpinski et al., 1992); C/EBP-related factor LIP (Descombes and Schibler, 1991; Ron and Habener, 1992); two members of the AP-1 family Jun B (Chiu et al., 1989) and FosB-s (Nakabeppu and Nathans, 1991). In these latter examples negative regulators are the product of alternative splicing or of alternative use of translation initiation codon, of genes encoding also the activator form (reviewed in Foulkes and Sassone-Corsi, 1992).

c) Interference with activator DNA binding. Some repressors work by competing with a functional activator for binding to the same or overlapping recognition site, preventing the binding of the activator by steric hindrance. Transcription factors belonging to the same family may show similar or identical DNA binding specificity as Drosophila homeodomain-containing proteins; the cyclic-AMP response element binding factor (CREB) and the activator protein 1 (AP1) factors, where the specificity of CREB activity, and hence the antagonism between CREB and AP1 at a non-consensus binding site, is modulated by cAMP dependent phosphorylation of CREB (Masquilier and Sassone-Corsi, 1992; Lamph
et al., 1990); the members of the Sp1 family of activators (Hagen et al., 1994). In these examples it is clear that the negative effector may simply be the weaker of two activators. Another important example comes from the nuclear hormone receptors family. Very many negative regulatory processes involve the nuclear hormone receptors and several act through cross-talk with the Fos and jun proto-oncogene products (AP1 factor) regulatory pathway. The AP1 factor is involved in proliferative state of the cell (Ransone and Verma, 1990) while the steroid/thyroid hormones are found to induce differentiation, or to maintain cells in a quiescent state. An antagonism between the proliferative function of AP1 and the differentiative function of nuclear hormone receptors has been described involving a reciprocal interference with transcriptional activation (Yang-Yen et al., 1990; Schule et al., 1990a; Jonat et al., 1990). Furthermore while binding of the receptor to DNA is not required for its negative regulatory activity, the zinc-finger DNA binding domain is necessary maybe because of an involvement of protein-protein interaction with the components of AP1 complex. On the other hand the leucine zipper domain (the dimerization domain) of c-jun and the N-terminal region of c-Fos are implicated in negative regulation of GR function (Schule et al., 1990a). Interactions between receptor and AP1 factors have been demonstrated both in vitro (Touray et al., 1991) and in vivo (Yang-Yen et al., 1990) and several mechanisms have been proposed: i) the interaction mutually inhibits DNA binding of the two classes of proteins (this is true in vitro [Yang-Yen et al., 1990] but it is not known whether it occurs also in vivo [Jonat et al., 1990]); ii) the activity of AP1 complex may be modulated by the formation of a ternary complex of AP1/TRE/GR (Jonat et al., 1990); iii) binding to a GRE (Glucocorticoid responsive element) is required. This kind of binding interference has been showed for the osteocalcin promoter (a RARE overlaps an AP1 binding site; Schule et al., 1990b) and the α-fetoprotein promoter where a GRE overlaps an AP1
consensus site (Zhang et al., 1991; Guertin et al., 1988); iv) occupation of a composite response element. This mechanism has been described for the proliferin gene, where GR is active when c-Jun/cFos heterodimer occupies the AP1 site but not when c-Jun homodimer binds the element (Miner and Yamamoto, 1991; Diamond et al., 1990; Yoshinaga and Yamamoto, 1991).

d) Interference with the activity of DNA bound activators. In many promoters repressor binding sites are adjacent to, although not overlapping, binding sites for transcriptional activators. A repressor could block activator function by directly contacting the adjacent DNA-bound activator and masking the protein domain responsible for transcriptional stimulation. Binding to DNA is not strictly required as for GAL 80, a regulator of galactose metabolism in yeast, that acts by masking the activation domain of GAL4 (Ma and Ptashne, 1988). Similarly in mammals the retinoblastoma gene product Rb modulates the activity of DNA binding factors with which it interacts, blocking the contact between the activation domain and the polymerase complex (Weintraub et al., 1992).

e) Activation occurs through the action of intermediary factors or cofactors (Ptashne and Gann, 1990; Pugh and Tjian, 1990; Martin et al., 1990). A large excess of transcription factor, in vitro, can lead to repression of activated transcription by a squelching mechanism, in which cofactors are titrated off (Gill and Ptashne, 1988). Squelching has been observed also in vivo without artificially high concentrations of the negatively acting factor (Meyer et al., 1989), suggesting that the cofactors might be limiting for some transcription activation processes, making of squelching a physiologically significant mechanism of regulating gene expression. An example is given by the adenovirus E1a gene product. Besides activating transcription of viral genes it is also able to repress transactivation by several transcription factor. A model for E1a action is that it interferes with the transmission of the activating signals by interaction with cofactors, or with
polymerase complex (Rochette-Egly et al., 1990). Furthermore the interaction of E1a with TFIID has been demonstrated allowing for both positive and negative effect on transcription (Horikoshi et al., 1991).

**Active repression**

An efficient mode to repress genes controlled by multiple activators is to interfere directly with the general transcription machinery, inhibiting transcription initiation by a disruption of the polymerase complex at the promoter or by impairing the catalytic activity of the complex. Repressor proteins could interfere with the assembly of GTFs and sterically block the addition of subsequent proteins; interact with the general transcription machinery and prevent disassembly step (it will freeze the assembly); recruit another factor into the general factor assembly that could act through either of the previous mechanism.

Although they share mechanism of interaction with the basal machinery, while the general repressors of basal transcription do not show any specific DNA binding activity (Dr1, Dr2, NC1, NC2, see also "The human general cofactors"), specific repressor proteins function is mediated by the recognition of specific cis-regulatory elements in the promoters. DNA sequences recognised by repressor proteins have been termed either negatively acting sequences or silencers, although these terms are not completely interchangeable. Silencer, by analogy with enhancer, is referred to a sequence that act at a great distance and independently of orientation, although some have exhibited a varying degree of dependence upon position and orientation. Like the enhancers, they are composite in structure or show increased effects on transcription with increased copy number (Lee et al., 1991; Savagner et al., 1990; Nishiyama et al., 1994). Many negative cis-regulatory sequence, with the properties of a silencer, have been identified in genes encoding growth-hormone (Larsen et al., 1986), insulin (Nir et al., 1986), renin (Burt et al., 1989; Barrett et al., 1992), interleukin-2 receptor α chain
Introduction

(Smith and Greene, 1989), immunoglobulin Kappa light chain (Pierce et al., 1991); immunoglobulin heavy chain (Imler et al., 1987), T cell-receptor α chain (Winoto and Baltimore, 1989), α-fetoprotein (Muglia and Rothman-Denes, 1986), vimentin (Farrel et al., 1990), collagen II (Savagner et al., 1990) and ε-globin (Cao et al., 1989).

Repressors, like activators, have been shown to consist of modular structure, containing separable DNA binding and repression domains as reported for the Wilm's tumor gene product WT1 (Madden et al., 1993), the Drosophila developmental regulators even-skipped (eve; Han and Manley, 1993), engrailed (en; Jaynes and O'Farrell, 1991) and Kruppel (Kr; Licht et al., 1990), the thyroid hormone receptor gene c-erbA and its viral counterpart v-ErbA (Banahmad et al., 1992), although no clear aminoacid sequence similarities exist between the repressing domains of different active repressors. However, they show a common richness in alanine, glutammine and /or proline residues as in eve, WT1, Kr, en, or domains rich in charged aminoacids (v-ErbA). Another motif is the carboxy-terminal WRPW (Trp,Arg,Pro,Trp) domain of Drosophila Hairy related bHLH repressor protein, which interacts with the Groucho family proteins of corepressor (Fisher et al., 1996). Schnabel and Abate-Shen (Schanbel and Abate-Shen, 1996) demonstrated that the transcriptional repressing activity of HoxA7 resides in its homeodomain, suggesting for homeodomains an additional role in transcriptional control beside their function in DNA binding. Furthermore they showed that residues in the N-terminal-arm of the homeodomain play a role in the differential repressing activity of various Hox proteins, leading to the hypothesis that these residues are necessary for distinguishing the functional action of homeodomain proteins in vivo.

By analogy with transcriptional activation that involves communication with every component of the transcriptional machinery through direct and indirect contacts, mechanisms of repression involving promoter occlusion or
interaction with the basal transcription complex have been described (Hansen et al., 1981; DeLuca and Schaffer, 1988; Okhuma et al., 1990). Many activators show also repressive function depending on the arrangement of their binding sites with respect to the promoter. Dissection of the mechanism by which they operate as transcriptional repressor led to models involving either interaction with corepressor molecules or directly with components of the basal transcription apparatus. The yeast Leu3p protein, involved in the regulation of genes required for branched-chain aminoacid biosynthesis and nitrogen metabolism, represses gene expression by forming a complex with Mot1p/TFIID. It is not clear yet if Mot1p mediates Leu3p function by destabilization of TBP-DNA interaction or whether a stable TBP-Mot1p-DNA complex, promoted by Leu3p binding at the promoter, may prohibit the subsequent association of downstream general transcription factors (Wade and Jaehning, 1996). Similarly ICP4 activates most HSV genes while strongly repressing activated transcription from other promoters containing strong ICP4 binding sites near the start site of transcription. As a general inhibitor of activated transcription it has been postulated that repression by ICP4 must involve a global alteration of the general transcription complex that blocks activation but still enables basal transcription. Gu et al (1995) showed that ICP4 is complexed with TFIID/TFIIB and suggested that binding to DNA of this tripartite complex causes a conformational change in the transcription complex that is unable to respond to positive factors, allowing only basal transcription. Alternatively, it has been proposed that the presence of ICP4 might interfere with the access of other general cofactors as PC4 and p15 (Ge and Roeder, 1994a; Kretzschmar et al., 1994).

Other repressors act directly on basal transcription. The fly even-skipped protein (eve; Um et al., 1995), the murine homeodomain protein MSX-1 (Catron et al., 1995; Zhang et al., 1996), the unliganded human thyroid receptor (v-Erb A
protooncogene gene product; Fondell et al., 1993) and the adenovirus E1B-55K protein (Yew et al., 1994) repress transcription by direct interaction. The even-skipped repressor protein interacts directly with the TATA binding protein (Um et al., 1995), while the product of v-Erb A gene can bind TFIIB (Fondell et al., 1993). MSX-1 mediates repression via alanine-plus-proline rich regions and interacts with basal factor complexes in gel shift assays (may be TBP; Catron et al., 1995).

Regarding eve, two model for transcriptional repression have been proposed: i) binding to TBP (or TFIID) would interfere with the binding of other GTFs or ii) interaction eve/TBP destabilizes DNA-TFIID binding. Some of these interactions may be mediated by corepressors such as the Groucho protein in D.melanogaster (Paroush et al., 1994).

The recruitment of a corepressor adds another important level of regulation, because the interaction between the DNA binding protein and the corepressor can be regulated.

In yeast, SSN6/TUP1 repressor is involved in transcriptional repression of several diverse sets of gene including a-specific, haploid-specific, and glucose-repressible genes. This complex interacts with a variety of specific DNA binding-proteins (Keleher et al., 1992), making of this complex a general transcriptional corepressor whose specificity is dictated by the DNA binding specificity of protein partners. A target for SSN6/TUP1 repressor is the cyclin-kinase complex associated to the holoenzyme (the product of SRB 10, the kinase, and SRB 11, the cyclin). It is not clear yet whether this repressor acts by modulating the activity of the kinase or uses the interaction for tethering to the transcription machinery apparatus (Kuchin et al., 1995; Liao et al., 1995).

Studies of transcription in vitro support the idea that histone proteins can act as transcriptional repressors of eukaryotic genes, probably by preventing TFIID access to the DNA (Knezetic and Luse, 1986; Wasylyk and Chambon, 1979).
Incubation of TFIID with the template DNA prior to nucleosome assembly prevents the nucleosomal inhibition of transcription, which suggests that it is TFIID binding that is inhibited by the presence of histone complexes (Workman and Roeder, 1987; Knezetic et al., 1988). Specifically positioned nucleosomes have been observed at both repressed and activated promoters and have been proposed to be involved in transcriptional regulation by placing regulatory sites in an inaccessible place on the nucleosome in the case of repression or enhancing expression by juxtaposing activator binding sites in the case of activation (reviewed in Lu et al., 1994). Some repressor proteins might block transcription from a target promoter by directing the formation of a positioned nucleosome over the TATA-box (Roth et al., 1990). Studies of a glucose-repressed gene also show a correlation between SSN6/TUP1 repression and the presence of a positioned nucleosome in the initiation region (Perez-Ortin et al., 1987; Matallana et al., 1992). Thus occlusion of promoter DNA by a positioned nucleosome might contribute to transcriptional repression by SSN6/TUP1. It is not clear yet what determines nucleosome positioning but it has been proposed that either DNA sequences (Thoma and Simpson, 1985; Simpson, 1991) or DNA binding protein that create a boundary in chromatin (Fedor et al., 1988) or proteins that actively position nucleosomes at adjacent sequences (Roth et al., 1990) might contribute to this.

As already mentioned in the "positive control" section, there is a general correlation between core histone acetylation and gene activity. Conversely, it has been thought that core histone deacetylation leads to transcriptional repression. This is consistent with the finding of a mammalian histone deacetylase (HDAC1; Taunton et al., 1996) that is related to the yeast Rpd3 protein which is required for full repression as well as full activation of gene expression (Vidal and Gaber, 1991). Several recent papers suggest that transcriptional repression by a sequence-
specific DNA-binding protein can be mediated by the recruitment of a deacetylase (Rpd3, HDAC1, HDAC2) to the promoter region. The recruitment seems to be mediated by other molecules working as corepressors (Sin3, N-Cor/SMART; reviewed in Wolffe, 1997).

Chromatin mediated silencing

Silencing at telomeres and at HML and HMR

Silencing may involve changes in chromatin structure by establishing specific structural domains. This is sustained by the finding of factors that through binding to silencer are required either for nuclear scaffold attachment or transcriptional repression (Hofmann et al., 1989; Shore, 1994).

One of the best studied mechanism of repression occurs at the yeast Mating Type loci and near telomeres.

The budding yeast exists as haploid cells of a or α mating-type, which can mate to form a/α diploids. The genes involved in determining mating-type reside at three unlinked loci: MAT, HML and HMR. The mating-type of a haploid cell is determined by the allele, MAT a or MAT α, present at the MAT locus. Yeast cells can switch their mating-type by replacing the allele at MAT with an allele of the opposite mating type copied from either HML or HMR. The genes at the HM loci are transcriptionally inactive. The silent mating type loci HML and HMR have flanking sequences called the E and I elements that are required in cis for repression of gene expression (Laurenson and Rine, 1992; Brand et al., 1987; Figure 1.10a). Analysis of the E element revealed three functional sub-elements that contribute to the silencer function, A, B and E, none of which is able to cause silencing on its own. The A element contains an ARS consensus sequence (Campbell and Newlon, 1991), responsible for the binding of ORC (origin recognition complex; Bell and Stillman, 1992) to all four silencers (HML-E, HML-I, HMR-E, HMR-I) in vitro (Bell et al., 1993). The E and B elements contain binding
sites for the transcription factors RAP1 and ABF1 respectively (Shore and Nasmyth, 1987; Diffley and Stillman, 1989). ORC seems to play a critical role in the establishment of silencing at HM loci (Fig. 1.10b) as well as ABF1 and RIF1 (a RAP1 interacting factor; Hardy et al., 1992). It is not clear yet whether the involvement of ORC is caused by a necessary role for replication and/or whether ORC plays a more direct role in establishing a refractory chromatin structure (Newlon, 1993). What is known is that the repressed state can be established only following the passage through S-phase leading to the hypothesis that the establishment might be a consequence of replication associated chromatin assembly (Miller and Nasmyth, 1984).

It appears, however, that repression at these loci, as well as at telomeres, is mediated by a multiproteic complex containing nucleosomes and the product of SIR genes (silent information regulator; Rine and Herskowitz, 1987; Aparicio et al., 1991). While SIR1, probably through interaction with ORC, appears to act exclusively during the establishment of silencing (Chien et al., 1993), SIR2, SIR3 and SIR4 are required for the maintenance of the repressed state (Pillus and Rine, 1989). Recently, SIR3 has been directly implicated in spreading of the repressive state (Fig. 1.10c; Renauld et al., 1993). A direct involvement of histones in silencing was suggested by studies on deletion of the aminoterminal part of histone 3 and histone 4 that leads to a derepression of HM loci and the genes located near telomeres (Kayne et al., 1988). Genetic studies have also showed that SIR3 and SIR4 might interact directly with the histone aminotermini to nucleate a repressive structure (Hecht et al., 1995). Furthermore the packaging of HM loci in an inaccessible chromatin structure has also been determined by its inaccessibility to various enzymatic probes (Loo and Rine, 1994). Interaction between SIR3 and SIR4 with RAP1 has been demonstrated by two-hybrid experiments (Fig. 1.10b; Moretti et al., 1994; Cockell et al., 1995).
The transcriptional state of the silent domain is stably inherited through multiple cell cycles and the intact silencer is required for the efficient inheritance of the repressed chromatin structures following replication. So the silencer has all the information promoting the reformation of the silent chromatin in the progeny (Holmes and Broach, 1996).

---

**Fig.1.10: Model for silencing establishment at HM loci and telomeres in yeast.**

**A:** MAT, HML and HMR are the three loci that contain the regulatory genes for mating-type determination. The three subelements and the protein bound to the HMR-E silencer are shown. **B:** Model for silencing at HM loci. **C:** Model of protein-protein interaction involved in the nucleation and propagation of chromatin-mediated gene repression at telomeres. R is for RAP1. 3 and 4 are for SIR3 and SIR4 respectively (Cockell, 1995).
The studies on silencing in yeast stress the importance of maintaining stably repressed genes, as developmental regulatory genes, in multicellular organisms. The homeotic genes of Drosophila (reviewed in Gehring et al., 1994) provide a key model for studying mechanisms that maintain state of gene expression during development. Homeotic gene expression depends on a set of regulatory factors acting earlier in development and of another set of factors that maintain the patterns during later development. The homeotic genes encode transcription factors required for proper determination of identities of the body segments. Each homeotic gene is expressed in a spatially restricted domain along the anterior-posterior axis (A/P). The proper regulation of these genes is under the control of proteins that repress (gap protein) or activate (pair-rule) their expression by competing for binding to nearby or overlapping sites in the promoters (Shimell, et al., 1994; Qian et al., 1993; Zhang and Schultz, 1992; Muller and Bienz, 1992). Later in development other classes of proteins are required for the correct maintenance of this pattern and genes involved in maintaining the repressive (Polycomb-group, Pc-G) or active (Tritorax-group, trx-G) state have been identified. The current model for the action of these genes is that they "lock" gene expression in an inactive or active state by changing in chromatin structure. Then, Pc-G genes would maintain repressed those homeotic genes that were off during earlier step of development as trx-G genes maintain active genes there were in a on state. Pleiotropic phenotypes of many Pc-G mutations suggested that Pc-G action is not limited to homeotic gene expression but rather they exert a more general control on oogenesis, neural development and cell proliferation (Phillips and Shearn, 1990; Paro, 1990; Adler et al., 1991). So far 13 members of this group have been identified by genetic studies six of which have been characterised at molecular level: Polycomb (Pc), polyhomeotic (ph), Posterior sex combs (Psc), Enhancer of
zeste E(z), Polycomb-like (Pcl) and extra sex combs (esc). Coimmunoprecipitation experiments (Franke et al., 1992) and double-immunofluorescence labelling of polythene chromosomes (Franke et al., 1992; Rastelli et al., 1993; Lonie et al., 1994) indicate that the Pc-G gene products form several large complexes, whose composition may vary in a cell-type specific manner playing different roles in different tissues during development.

A Pc-G response element (PRE) has been identified in the regulatory region of genes of ANT-C and BX-C complexes and it has been demonstrated that it is necessary for the establishment of repression of reporter gene expression in germ line transfection studies, although so far no DNA binding activity has been identified for any of the Pc-G proteins (Simon et al., 1993; Chan et al., 1994; Gindhart and Kaufman, 1995). This leaves an open question about the mechanism by which these sequences are recognised by the Pc-G proteins. Once targeted, the Pc-G would maintain repression of homeotic genes by the creation of a locked chromatin structure not accessible to any activator protein (Fig. 1.11).

![Fig. 1.11: Proposed mechanism for stable repression by Polycomb group protein](image-url)
This model, termed chromatin accessibility model, is sustained by several observations: i) immunoprecipitation of UV-crosslinked Pc proteins showed that they are physically associated with sequences corresponding to the entire length of *Ubx* and *AbdA* genes. Some of these sequences contain high-affinity binding sites as tested by the ability to confer repression on a linked *lacZ* gene (Orlando and Paro, 1993); ii) Pc protein contain a region termed the "chromodomain" that is conserved in the *Drosophila* heterochromatin associated factor (HP1; Paro and Hogness, 1991), suggesting a structural similarity between Pc and HP1 leading to the idea that Pc-G proteins repress transcription by packaging the DNA into an heterochromatin-like structure.; iii) the *Drosophila* gene BRM, whose function has been associated with that of the chromatin remodelling SWI/SNF complex in yeast, has been isolated as suppressor of Pc mutations (Tamkun *et al.*, 1992); iv) the ability of sequence-specific DNA binding protein to interact with Pc-G repressed genes is inhibited in a manner similar to inhibition by stable nucleosomes (g and Paro, 1995; McCall and Bender, 1996). In this model Pc-G gene products would be targeted, by binding to high-affinity binding sites, to nucleate formation of a complex that then would spread to adjacent region creating a stable repressed chromatin structure (Figure 1.11).

It was found that the Pc chromodomain plays an important role in directing Pc-G complexes assembly. Mutations in the chromodomain do not only inhibit the binding of the Pc protein to its target genes but also affect the integrity of the entire complex (Franke *et al.*, 1995). Once again the importance of such regulatory mechanism is outlined by the identification of mammalian homologues of Pc-G gene products. *bmi-1*, identified as oncogenes in mouse, (Haupt *et al.*, 1991; Van Lohuizen *et al.*, 1991) and the human *mel-18* (Ishida *et al.*, 1993; Kanno *et al.*, 1995), show about 43% identity with fly Psc in the region surrounding the RING finger motif; the mouse Rae-28 protein contains several regions of sequence
similarity to the fly ph protein (Nomura et al., 1994); a number of mammalian proteins have been described to contain chromodomains as mouse M33 protein (Pearce et al., 1992; Saunders et al., 1993), that is a good candidate as a functional Pc homologue.

From the examples reported above it is clear that the activity of promoter regions depends on the balance of function of complexes that through their effect on chromatin structure can lead to repression or activation. In particular several line of evidences are consistent with a model in which nucleosome mobility plays a key role in determining promoter activity and hence gene regulation. Increased nucleosome mobility might result from sliding of nucleosome along DNA or from direct transfer of the nucleosome from one to an adjacent segment of DNA. Activators, through ATP-dependent remodelling complexes, would act by increasing mobility of nucleosomes and might help displace Pc-G complexes preventing silencing (Aparicio and Gottschling, 1994; Zink and Paro, 1995), while repressors could function by directing formation of a structure that decrease nucleosome mobility and by blocking the recruitment of factors that increased mobility.

**DNA methylation mediated repression**

DNA methylation has been proposed to have an active role in gene regulation (Li et al., 1993; Li et al., 1992). Transient transfection and *in vitro* transcription studies demonstrated that DNA methylation leads to the repression of transcription by interfering with the binding of transcription factor and by directing the assembly of specific-nucleoprotein complexes formed by methyl-specific DNA binding proteins and histones which potentially block transcription factors access (Murray and Grosveld, 1987; Iguchi-Ariga and Schaffner, 1989; Ball et al., 1983; Boyes and Bird, 1991; Jost and Hofsteenge, 1992). Specialised chromatin structures have also been implicated in maintaining the transcriptionally silenced
state (Graessmann and Graessmann, 1988; Kass et al., 1994; Hsieh, 1994). Recently Kass et al. (1997) have shown that the process of silencing transcription by DNA methylation involves both the inhibition of transcription initiation and the removal of engaged transcriptional machinery from active templates, followed by a time-dependent assembly of a repressive structure that includes a higher-order nucleosomal DNA structure (Kass et al., 1997).

**The plasminogen activation system**

Extracellular targeted proteolysis is essential for cell migration and tissue remodelling. One major system involved in these processes is the activation of plasminogen. The activation of plasminogen into plasmin leads to the degradation of extracellular matrix molecules, as fibronectin and laminin, and this event is at the centre of a variety of biological processes such as fibrinolysis, wound healing, angiogenesis, tumour metastasis, embryogenesis, gametogenesis, ovulation and mammary gland involution (reviewed in Saksela and Rifkin, 1988; Plow et al., 1995).

**The activation of plasminogen**

Plasminogen is found in high levels almost ubiquitiously (Raum et al., 1980; Isseroff and Rifkin, 1983; Miles and Plow, 1985; Maijar et al., 1986; Saksela and Vihko, 1986). Localised activation of plasminogen is regulated by a complex network of molecular interactions in which activators (urokinase-type and tissue-type plasminogen activators), specific inhibitors (PAI-1, PAI-2, PAI-3, protease nexin-I and II, α-2 antiplasmin) and receptors (uPA receptor, plasminogen receptor) are involved (see Figure 1.12). Both plasminogen activators (PAs) as well as their product, plasmin, belong to the family of serine proteases which also include trypsin, chymotrypsin, thrombin, elastase and others. Serine proteases
contain a catalytic triad in their active site, serine, histidine, aspartic, which are in close contact and are important for the proteolytic cleavage (Stryer, 1995).

Plasminogen is activated by the PAs which cleave a single peptide bond between Arg 560 and Val 561. The two resulting polypeptides chains of active plasmin are held together by a disulphide bond.

![Fig. 1.12: The plasminogen activation system](image)

**The plasminogen activators**

Both urokinase-type and tissue-type plasminogen activators (uPA and tPA) catalyse the activation of plasminogen but with distinct roles *in vivo*. In the absence of fibrin, uPA has a substantial plasminogen activator activity, whereas tPA is hardly active. Only binding of tPA to fibrin causes a several fold stimulation of its activity (Andreasen *et al.*, 1990), playing a crucial role in intravascular fibrinolysis (Collen and Lijnen, 1991) and thrombolysis (Nilsson *et al.*, 1985). uPA has a more general role in the blood stream primarily in the cell-mediated proteolysis during macrophages invasion, wound-healing,
Introduction

embryogenesis, invasiveness and metastasis (Dano et al., 1985; Fazioli and Blasi, 1994; Sappino et al., 1989; Moscatelli and Rifkin, 1988; Kirchheimer and Remold, 1989; Baker et al., 1990). However, uPA and tPA knock-out experiments in mice have shown that many of the functions of uPA and tPA are shared and that only the knock-out of both genes leads to severe effects in the animals (Carmeliet and Collen, 1994; 1995). tPA deficient mice show an impaired thrombolysis and defective long term potentiation, a process in the brain that is related to learning and memory (Frey et al., 1996); on the other hand, inactivation of the uPA gene results in occasional spontaneous fibrin deposition, deficiency in wound-healing and abolishment of cellular recruitment in response to pulmonary inflammation (Gyetko et al., 1996). Combined uPA− and tPA− mice show an extensive spontaneous fibrin deposition in various organs and totally lack thrombolysis after endotoxin injection.

The two activators are product of different genes, the human gene for tPA is on chromosome 8, whereas that for human uPA is located on chromosome 10. The human tPA gene is 29-32.7 Kb long, consists of 14 exons (Ny et al., 1984) and codes for a 70 KDa protein via a 2.7 Kb mRNA (Fisher et al., 1985; Pennica et al., 1983). The human 6.4 Kb uPA gene contains 11 exons (Riccio et al., 1985) and encodes for a 53 KDa protein via a 2.5 Kb mRNA (Verde et al., 1984). The difference in gene size is due to long introns in the tPA gene, which contain several open reading frames without a known function (Degen et al., 1986).

The two PAs are highly similar in their basic structures indicative of a close evolutionary relationship, but share less than 40% identity at the aminoacid level (Degen et al., 1986). Both uPA and tPA have a similar catalytic domain located in the carboxy-terminal part of both enzymes, a growth factor-like domain, which in the case of uPA is responsible for the binding to its specific receptor, and kringle domains (one in uPA and two in tPA). The additional kringle domain and the
amino-terminal finger-like domain present in tPA seem to be responsible for its binding to fibrin (van Zonneveld et al., 1986).

**uPA and its receptor**

uPA is produced from cells as a single-chain pro-enzyme (pro-uPA) and converted by limited proteolysis into an active two-chain molecule.

Binding of uPA to its specific cellular receptor (uPAR) (Stoppelli et al., 1985; Vassalli et al., 1985) regulates uPA activity by enhancing the rate of conversion of pro-uPA into the active two chain molecule (Ellis et al., 1989; Manchanda and Schwartz, 1991), localising the proteolytic activity to the cell surface, or inhibiting the active enzyme through specific inhibitors (Cubellis et al., 1989; Ellis et al., 1990). The initiating event in the conversion of pro-uPA to active uPA in vivo is still not known. However, in vitro experiments show that plasmin is able to activate pro-uPA in solution (Wun et al., 1982), but a twenty-fold increase in uPA activation rate is achieved when pro-uPA is on the cell surface (Ellis et al., 1989). Proteases other than plasmin, involved in different physiological processes, can activate pro-uPA. Among these plasma kallikrein (Ichinose et al., 1986) trypsin (Koivunen et al., 1989), thermolysin (Marcotte and Henkin, 1993), cathepsin G (Learmonth et al., 1992), cathepsin B (Kobayashi et al., 1991) and L 6 (Goretzki et al., 1992), nerve growth factor γ (Wolf et al., 1993) and thrombin (Nauland and Rijken, 1994).

The urokinase receptor is a heavily glycosylated 50-55 KDa protein made up of three homologous repeats of ~90 residues each, of which the amino-terminal domain is involved in ligand binding (Behrendt et al., 1990; 1991, Roldan et al., 1990), and is anchored to the plasma membrane by a glycosyl-phosphatidylinositol (GPI) moiety (Ploug et al., 1991). Glycosylation is necessary for the intracellular transport and maturation of the receptor but can also affect the affinity of the receptor for uPA (Picone et al., 1989; Møller et al., 1993).
Although most receptors mediate rapid endocytosis of their extracellular ligand, the uPA receptor does not internalise bound active uPA (Stoppelli et al., 1986a). Cell-surface uPA activity is regulated by the availability of extracellular PAIs. Once the uPAR-uPA-PAI-1 complex has been formed at the cell surface it is rapidly internalised and degraded (Cubellis et al., 1990). This has been also shown with uPA-PAI-2 (Estreicher et al., 1990) and uPA-PN-1 complexes (Conese et al., 1994). This internalisation event occurs by the interaction of these complexes with the alpha2-macroglobulin receptor/low density lipoprotein receptor-related protein (α2MR/LRP; Nykjær et al., 1992). The uPA receptor appears to be transiently internalised with the uPA-PAI-1 complex, and rapidly recycled back to the surface (Conese and Blasi, 1995; Nykjaer et al., 1997). The ability of inhibitors to block surface uPA activity and of uPA receptors to recycle might result in continuous modulation of migration-competent areas of the cell membrane by modifying the location of the proteolytic activity on the cell surface (Fazioli and Blasi, 1994). For instance, in non-migrating cells, uPAR is distributed on the cell surface at focal contacts or on the apical side (Limongi et al., 1995) whereas in actively migrating monocytes uPAR rapidly redistributes at the leading edge of the cells (Estreicher et al., 1990).

**The biological role of uPA**

Cells are structurally and functionally integrated through the interactions with the extracellular matrix and surrounding cells. Interactions are mediated by specific receptors, such as integrins and cadherins, with intracellular domains of these receptors connected to the cytoskeleton at focal adhesion, and by a number of molecules as fibronectin, laminin, thrombospondin, tenascin and collagen in the ECM, and vinculin, actin, α-actinin and paxilin in the cell (reviewed in Ridley, 1994; Carraway and Carraway, 1995).
The important role of uPA in cell migration and invasiveness has been so far explained by the ability of uPA to activate plasminogen to plasmin at the cell surface promoting degradation of fibrin (Reich, 1978) and extracellular matrix (ECM) and basement membrane glycoproteins (Quigley et al., 1987). Plasmin, in turn, activates other matrix-degrading proteases belonging to the metalloproteases family such as procollagenase type IV to collagenase (Saksela and Rifkin, 1988) and stromelysin (Matrisian, 1990). The activation of the metalloproteases family by means of plasminogen activators can therefore result in a synergistic and effective degradation of most, if not all, matrix components (Fig. 1.12).

uPA mediated extracellular proteolysis is involved in many physiological processes sustaining cell migration as i) the migration of keratinocytes in wound healing processes (Grøndal-Hansen et al., 1988) which can be stimulated by the epidermal growth factor (EGF) (Rørth et al., 1990; Jensen and Rodeck, 1993); ii) the trophoblast cell invasion into the uterine wall during the process of mouse embryo implantation (Sappino et al., 1989; Teesalu et al., 1996), where each component of the system displays a regional expression in the apical pole of the implanting zone. In particular, uPA expression is confined to the inner layer (extraembryonic throphoblasts), uPAR synthesis is pronounced in the external, decidual layer and PAI-1 is confined to the intermediate layer; iii) the migration of myoblasts during myogenesis and muscle regeneration (Quax et al., 1992; Munoz-Canoves and Felez, 1993). In response to muscle damage satellite muscle cells are activated and undergo proliferation, differentiation and fusion, leading to the formation of long-multinucleated myotubes (Carlson, 1973). To a degenerative period that leads to tissue breakdown follows a remodelling period characterised by tissue regeneration (Allbrook, 1981). Plasminogen activation seems to be involved in these processes as suggested also by in vitro studies of muscle differentiation showing that uPA gene expression is modulated during the
differentiation of myoblasts into myotubes. The uPA level is high during satellite cells proliferation and migration and decrease when myoblasts fuse to myotubes (Ibañez-Tallon, 1997; Munoz-Canoves et al., 1997).

A multitude of studies in a variety of different systems have shown that the uPA-uPAR interaction may affect cell migration through multiple mechanisms (see figure 1.13). An example is given by the process of angiogenesis. Endothelial cell migration is positively regulated by the basic fibroblast growth factor (bFGF) (Gualandris and Presta, 1995) and negatively by the transforming growth factor β (TGF-β). It results in a regulatory circuit with bFGF stimulating uPA and collagenase synthesis, uPA in turn activating TGF-β, which eventually decreases uPA activity by inducing PAI-1 synthesis (Fazioli and Blasi, 1994). Binding of uPA can induce a chemotactic response as shown for human monocytic leukaemia cell line THP-1 cells, independently of the catalytic activity (Resnati et al., 1996).

In recent years, a number of evidence indicate the possibility of a direct intracellular signalling as an additional way by which the uPA-uPAR system can influence cell migration (Busso et al., 1994; Sitrin et al., 1996; Resnati et al., 1996). The recent discovery of substrates of uPA other than plasminogen, such as the pro-transforming growth factor (pro-TGF-β) (Odekon et al., 1994), and the precursor from hepatocyte growth factor (pro-HGF) (Naldini et al., 1992) which is a cell-migration stimulating protein, provides a role for uPA as modulator of cell migration by regulating the activity of other protein factors.

**Role of uPA in cancer**

Due to their destructive potential, extracellular proteolytic enzymes have to be precisely and tightly controlled in their spatial and temporal expression. A pathological situation where these control mechanisms appear to fail dramatically is observed in malignant tumour cells, which exhibits progressive invasion of surrounding normal tissues.
The are commonly found to be expressed by malignant tumours (reviewed in Pollanen et al., 1991). Tumour tissue extracts from colon, breast, lung and prostate carcinomas have higher uPA activity than their normal non-transformed tissues.

Fig. 1.13: The different mechanisms by which the uPA-uPAR interaction may affect cell migration
uPA is represented by the three linked domains: the growth factor domain (GFD), the kringle domain (Kr) and the protease (Pr) domain; ECM, extracellular matrix; pro-HGF, pro-hepatocyte growth factor; pro-MSP, pro-macrophage stimulating protein; TGF-β, Transforming growth factor β. Solid lines indicate stimulation and broken lines indicate inhibition (Fazioli and Blasi, 1994).

In addition, PAs are found to be highly expressed in cells which have been infected with oncogenic viruses (Zhang and Schultz, 1992), correlating with the
appearance of certain features of malignant cells such as anchorage independent
growth, increased motility, reduced cytoskeletal organisation and morphological
changes. Several data strongly suggest that the presence of high levels of uPA,
uPAR and PAI-1 in human breast cancer may be used as an independent
prognostic marker for the early relapses of breast cancer patients (Jänicke et al.,
1990; Jänicke et al., 1991; Grøndal-Hansen et al., 1993) as well as for bladder and
lung cancer (Hasui et al., 1992; Pedersen et al., 1994a, b). Therefore, several
approaches have been made to use uPA and its receptor as anti-metastatic targets
(Fazioli and Blasi, 1994). It has been shown in a number of studies that tumour
invasiveness and metastasis can be blocked or stimulated in several model
systems by a direct decrease or increase of uPA activity (Danø et al., 1985; Ossowski
and Reich, 1983; Hearing et al., 1988; Axelrod et al., 1989; Ossowski et al., 1991;
Mignatti and Rifkin, 1993). Moreover, recent studies show that uPA-deficient
mice are more resistant to melanoma induction, again confirming the
contribution of uPA to malignant progression (Shapiro et al., 1996).

Transcriptional regulation of the uPA gene

Due to the critical consequences of its activation, uPA expression requires a
tight regulation. Many of the changes in the levels of uPA synthesis appear to be
due to changes in the rate of transcription, but also post-transcriptional
(Henderson et al., 1992; Nagamine et al., 1995) and post-translational mechanisms
regulate its expression.

Inducers of uPA synthesis

The distribution of uPA in the organism is under normal circumstances very
restricted. The presence of uPA in normal adult mouse is only detected in very
few cell types such as the fibroblast-like cells in the gastrointestinal tract, the
tubular cells in kidney and the pneumocytes in lung (Larsson et al., 1984).
However, many cell types, often derived from non-uPA expressing tissues, produce easily detectable amounts of uPA when transferred to cell culture. The increased uPA synthesis most likely reflects the altered growth conditions imposed on the cells in culture leading to a highly proliferative state.

The study of uPA synthesis and regulation in cell culture is, however, very useful when trying to uncover the regulatory elements responsible for the expression of uPA and the mechanisms by which they are activated. In this way, uPA synthesis has been shown to be modulated by a large number of stimuli in many different cell lines. Reflecting the link between cancer and uPA gene expression, many of the uPA inducing agents are involved in growth control regulation. Growth factors (such as fibroblast growth factor, transforming growth factor-beta, epidermal growth factor), phorbol esters, polypeptidic (FSH, LH, prolactin, calcitonin) and steroid hormones (estrogens, progesterone, dexamethasone and androgens), cAMP, antimitotic drugs (as colchicine and cytochalasin B), inflammatory cytokines (like tumour necrosis factor, interleukins and interferon γ), oncogenes and UV light, induce uPA gene expression (reviewed in Danø et al., 1985; Blasi and Verde, 1990; Ibañez-Tallon, 1993; Nagamine et al., 1995; Besser et al., 1996).

The transcriptional activation has been shown to be regulated independently through several signal transduction pathways involving protein kinase A (PKA), protein kinase C (PKC), cytoskeletal reorganisation and okadaic acid-sensitive protein phosphatase (Besser et al., 1996).

The uPA gene promoter

The genes encoding uPA and 5' flanking regions have been sequenced and extensively characterised in human (Riccio et al., 1985; Verde et al., 1988; Fig. 1.14), mouse (Degen et al., 1987) and pig (Nagamine et al., 1984). The length of the promoter sequences published so far are different between the three mammalian
promoters, 2.3 kb, 7 kb and 4.7 kb for the human, murine and porcine genes respectively (Riccio et al., 1985; Cassady et al., 1991). *In vitro* and *in vivo* studies have revealed the presence of a number of regulatory regions, some of them identical between all three mammalian promoters, but surprisingly, the far upstream region of the promoters share higher homology than the regions closer to the transcription start site.

**The minimal promoter**

The minimal promoter (MP) contains a TATA-box characteristic of many regulated genes, which is located 25 bp upstream of the transcription initiation site in all uPA promoters, but lacks the typical CAAT box. There are several copies of GGGCGG sequences, which are potential binding sites for Spl (four in the human, two in the murine and five in the porcine promoter) (Verde et al., 1988; Rossi et al., 1990).

**The uPA enhancer element**

The enhancer element is located at 2 kb (human and porcine promoters) and at 2.4 kb (murine promoter) upstream of the transcription initiation site (Verde, 1988 *et al.*; Rorth *et al.*, 1990; Cassady *et al.*, 1991). Two binding sites for the transcription factor AP-1, one of which is combined with the binding site for
PEA3, are present within the 120 bp-long enhancer and are crucial for its activity (PEA3/AP1a and AP1b sites) (Nerlov et al., 1991). In addition, a cooperation mediating (COM) element is positioned between the combined PEA3/AP1 and the AP1 sites (Nerlov et al., 1992). The COM element does not appear to have transactivation-mediating activity by itself, but it is important for synergistic activation of both AP-1 sites (De Cesare et al., 1996).

A similar enhancer modular structure has been described in the promoters of the stromelysin, interleukin-3, and cytokine LD78 genes (Sirum-Connolly and Brinckerhoff, 1991; Taylor et al., 1996; Nomiyama et al., 1993). These promoters are also regulated through AP-1 sites, and seem to have sequence homologies to the uPA enhancer not only in the transactivating elements, but also in the COM region.

The PEA3/AP1 element

The PEA3/AP1 element is conserved in the human, murine and porcine uPA promoters.

The GAGGAAA sequence has been shown to bind proteins of the Ets family, while the TGAAGTCA is an octameric AP-1 site, which can be imperfectly aligned with the octameric TREs of the c-jun promoter and with the CRE of the somatostatin promoter. The uPA AP1a sites of the murine and porcine enhancers are also formed by a non-canonical octameric binding sequence. Similar elements with a PEA3 site adjacent to an AP-1 site have been found in several promoters such as the polyoma A virus early enhancer and human collagenase enhancers. Maximal induction of these genes by growth stimuli requires both the AP-1 and the PEA3 binding sites.

It has been found that this PEA3/AP1 element is responsible for uPA gene induction through PKC activation by TPA (Rorth et al., 1990; Nerlov et al., 1991), protein phosphatase 2A (PP2A) inhibition by okadaic acid (Nagamine and Ziegler,
cytoskeletal reorganisation by colchicine or cytochalasins (Botteri et al., 1990), as well as for the response to several growth factors (Besser et al., 1995a) and oncogenes (Lengyel et al., 1995a,b; Nagamine et al., 1995). Moreover, the PEA3/AP1 element has also recently been identified as the target of the uPA transcriptional activation by middle-T antigen (Besser et al., 1995b; Urich et al., 1995), GM-CSF (Granulocyte-macrophage colony stimulation factor; Stacey et al., 1995) and interleukin-1 (Besser et al., 1996).

Recent studies have shown that in vitro the PEA3 site interacts with ets-2, a member of the PEA3/ets family, and in vivo mediates the induction of uPA transcription by colony-stimulating factor 1 (CSF-1; Stacey et al., 1995).

De Cesare et al. (1995) have reported that the uPA octameric upstream AP1a site interacts with the dimer formed by c-Jun and ATF-2 and that these heterodimers show particular regulatory features distinct from canonical AP-1 dimers: c-Fos can antagonise the transcriptional induction via c-Jun/ATF-2 since c-jun/c-fos heterodimers have a much lower affinity for that site.

The downstream AP-1 site

The downstream AP-1b site is an heptameric non canonical site which differs from the TRE canonical sequence because of the central adenine. This element is identical in all three uPA enhancers. The two AP-1 sites of the uPA enhancer are not functionally equivalent: mutations of the downstream site AP-1b have a relatively small effect, reducing inducibility by about 75%, while the upstream site AP-1a has a more profound influence (Nerlov et al., 1992). c-Jun/c-Fos heterodimers seem to bind the AP-1b element (Verde P., personal communication).

The COM region

DNaseI footprinting and site-directed mutagenesis have identified two areas of the COM region (u-COM and d-COM) important for the function of the
enhancer (Nerlov et al., 1991). Four different factors, uPA enhancer factors (UEF) 1 to 4, appear to bind to these regions. Mutations within the COM element, affecting the binding of the different UEFs impair both inducibility (Nerlov et al., 1992) as well as basal level activity (De Cesare et al., 1996) of the enhancer. Transfection and mutagenesis analysis have shown that the COM activity is position-dependent but orientation-independent, and that its function appears to be structural since it does not have transactivation properties (De Cesare et al., 1996).

**NF-kB sites**

Two NF-kB elements at -1580 and -1865 bp have been identified in the human uPA promoter. The site at -1580 mediates the induction by phorbol esters and tumour necrosis factor in HepG2 cells and HT1080 cells (Hansen et al., 1992), while the site at -1865 is involved in the TPA response in HeLa cells (Novak et al., 1991). Two different dimers have been shown to bind the site at -1580, NF-kB (p50 and p65 subunits) and a dimer of p65 and c-rel (Hansen et al., 1992). It is not known which of these dimers is responsible for the induction. However, the induction via this site may be repressed by a negative regulatory element upstream of the NF-kB site (Verde et al., 1988; Hansen et al., 1992).

**cAMP responsive elements**

Although it has been shown that PKA activation by cAMP agonists can induce uPA from all three mammalian species studied, the responsive elements have only been identified in the human and pig promoters.

Two putative AP2 binding sites in the human uPA promoter, close to the transcription start site have been shown to be responsible for PKA-dependent induction of uPA gene expression in murine Sertoli cells (Rossi et al., 1990) and the nuclear factors binding to them have been identified (Grimaldi et al., 1993a,b).

On the other hand, studies of the porcine uPA promoter in LLC-PK1 cells (Degen et al., 1985; Nagamine et al., 1983), indicate that PKA-dependent uPA
induction seems to be conferred by different elements identified in a DNase I-hypersensitivity study (Lee et al., 1994). A cAMP-responsive region at -3.4 kb of the porcine uPA promoter was studied in more detail. It is composed of three elements, termed A, B, and C, all required for full cAMP responsiveness, suggesting that the proteins binding to these sites interact with each other (Menoud et al., 1993; Marksitzer et al., 1995). Menoud et al. (1993) identified the protein binding to the C element as the porcine homologue of mouse liver factor B (LFB3) and human and rat variant hepatocyte nuclear factor 1 (vHNF1). The A and B elements contain CREB-like sequences but the binding proteins for these sites are not known. A high homology between the porcine uPA promoter A,B,C elements and human upstream sequences has been found, indicative of a similar PKA-dependent regulation of the human uPA gene (Ibañez-Tallon, 1997).

**Negative regulatory sites**

Several reports indicate that in cells that do not produce uPA the main regulatory mechanism may be a negative one. The first indirect evidence is that the level of uPA mRNA is increased upon cycloheximide treatment (Stoppelli et al., 1986b; Grimaldi et al., 1986; Altus et al., 1987) and that this effect is in part due to increased rate of transcription, although an increased stability of uPA mRNA cannot be excluded. Secondly, in LLC-PK1 cells, in which uPA expression is dependent on calcitonin or cyclic AMP treatment (Nagamine et al., 1983), recessive calcitonin-independent mutants have been isolated (Hofstetter et al., 1987). Deletion analysis of the 5' flanking region of human and porcine gene promoter identified a negative regulatory element located in the human promoter between -1824 and -1572 bp (Verde et al., 1988), and in the porcine between -2.7 and -1.6 kb (von der Ahe et al., 1990). Furthermore, a repressor acting through the NF-kB site at -1865 bp of the human promoter has been described (Novak et al., 1991).
unknown sites

The induction of the uPA promoter can be modulated by other signalling pathways. It has been reported that Ca$^{2+}$ can enhance the induction by cAMP, TPA and okadaic acid in LLC-PK1 cells although it does not induce the uPA gene by itself (Nagamine and Ziegler, 1991). Negative effects on uPA induction by cAMP, TPA and interleukin-1α/β and all-trans-retinoic acid have been reported for anti-inflammatory glucocorticoids. This probably occurs by an indirect mechanism, because no putative binding site for the glucocorticoid receptor is present in the uPA promoter, and the repression can be blocked by protein synthesis inhibitors (Vitti and Hamilton, 1988; Pearson et al., 1987). The mechanisms regulating Ca$^{2+}$ and glucocorticoid modulation of uPA gene expression are not well understood (Nagamine et al., 1995).

Chromatin structure of human uPA gene

Recently, in our laboratory, the regulation of the human urokinase gene expression has been related to the chromatin structure of the gene (Ibañez-Tallon, 1997). The analysis has been carried out in cancer cells which constitutively express urokinase (PC3), in cells which do not produce uPA (HeLa) and in HepG2 cells in which urokinase transcription is induced with TPA (Nerlov et al., 1992), relating the chromatin structure to the different functional states of the gene.

The DNase I hypersensitive sites (HS) analysis have revealed the presence of 8 HS sites in the active gene (see Figure 1.15). Three sites (HS1, HS2 and HS3) are located in the most 5' part of the promoter (also characterised in our lab.). The HS4 site maps in the enhancer element; HS5 corresponds to a putative Z-DNA structure; HS6 is localised between -0.7 and -0.55 Kb; HS7 spans the start site of transcription and the first intron and HS8 is 3' of the polyA site. Important differences between cell types were observed suggesting that a general loosening of
the chromatin structure over the proximal promoter region (PC3 versus HeLa) and local rearrangements (HepG2 -/+ TPA) correlate with uPA transcriptional activation. In cell lines not expressing uPA the minimal promoter (revealed as HS7) and the proximal regulatory region (identified by HS6) show resistance to DNase I, indicative of a compact chromatin organisation and possibly inaccessibility to the basal transcription machinery.

Furthermore the chromatin studies on the mouse urokinase gene during in vitro myoblast differentiation revealed a pattern of HS in which HS5, 6 and 7 correspond respectively to the HS4, 6 and 7 of the human gene and are present at the myoblast and early myotube stages. An overall decrease in DNaseI hypersensitivity over the distal promoter region correlates with differentiation.

Due to the quite high level of species conservation between the human and murine uPA promoters is important to underline that nuclease hypersensitivity associated to the minimal promoter and start site of transcription seems to be a necessary condition for both genes to be active. Human HeLa and mouse LB6 cells do not express detectable levels of uPA and do not have HS at the start site of transcription (see Figure 1.16).
Introduction

Fig. 1.15: Schematic representation of DNase I hypersensitive sites of huPA gene in different cell lines.
Both the promoter region (-11.8/+1) and the gene (+1/+9000) of human uPA are represented (Ibanez-Tallon, 1997).

Fig. 1.16: Schematic representation of homologous cis-acting regulative elements, associated to DNase I hypersensitivity, in the human and murine uPA promoters (Ibanez-Tallon, 1997).
2. Aim of the project

As described in the introduction, uPA is involved in normal and pathological processes by activating proteolytic reactions at the cell surface and is evident that due to its high destructive potential, uPA expression needs to be tightly regulated.

The restricted uPA expression in the organisms, very few cell types, and its overexpression in tumours and several transformed cell lines, makes of major importance to study and characterise the element(s) involved in the negative regulation of uPA, in order to understand the mechanisms that lead to high expression of the gene in tumour cells.

Progressive 5' deletions of the regulatory region of human uPA, from -2350 with respect to the transcription start site, fused to the CAT reporter gene, have already revealed the presence of positive (the enhancer, up to -1824) and negative cis-acting sequences (-1824/ -1572) and of a specific contribution to uPA expression of the region -1572 to -72 in the kidney derived A1251 cells, but not in HFS10 fibroblasts (Verde et al.,1988), suggesting the presence of multiple array of cis-acting sequences specific for different transcription factors that direct uPA expression in a cell-type specific manner.

Furthermore Cannio et al. (1991) described the presence of another negative regulatory region (-660/-86) in the proximal region of human uPA promoter, the activity of which was dependent on the presence of the enhancer. In particular, he mapped the silencing activity in the region -660/-537 and the enhancer dependence in the region -537/-301.

At the beginning of this project I decided to further characterise the silencer region described by Cannio et al. but I found out that part of the constructs used in
that analysis contained several important cloning mistakes so to invalidate all the results got that far.

I then decided to characterised again the 5' flanking region of the human uPA gene to find out elements that were involved in the negative regulation of the gene in cells not producing uPA compared to uPA producing cells (Figure 2.1).

As shown in the figure, the human HT1080 cell line and the human prostatic derived adenocarcinoma cell line PC-3 express uPA at high levels; the cervical carcinoma derived HeLa cells and the simian fibroblast-like CV1 cell line express little if not all uPA, while the human hepatocyte cell line HepG2, expresses uPA at a very low basal level inducible upon treatment with phorbol ester TPA.

All the constructs used in the transient transfection analysis are derived from the -2212/+30 region of the human uPA promoter driven the expression of CAT gene, in order to avoid the presence of great part of the 5'Alu sequence (-2350/-2140). On the other hand Verde et al (1988) already showed that the region up to -2.2 Kb contains the sequences needed to direct a cell-type specific expression of uPA, in culture.

In the second part of the thesis I started the characterisation of the factors that might be involved in the establishment of the silenced state of the uPA gene.
Other aspects of negative regulation of uPA, in different systems (repression by TGFβ and control of the expression of uPA during the cell cycle), that have raised my interested, are treated in the two appendices, as the results described are preliminar and they would need further studies.
3. Materials and Methods

*Cell culture, DNA transient transfection and CAT assay*

The cervical carcinoma HeLa, the prostatic derived adenocarcinoma PC-3 and the simian fibroblast-like CV-1 cells were grown in Dulbecco’s modified Eagle’s medium (DMEM; Gibco, UK) supplemented with 10% foetal calf serum, penicillin (100 U/ml), streptomycin (100 μg/ml) and glutamine (2 mg/ml).

Cells were transfected by the calcium phosphate precipitation technique (Graham and van der Eb, 1973). Briefly, cells were plated at a density of 7x10^5 per 100 mm plate 24 hours prior to transfection. A calcium phosphate DNA precipitate containing 10 μg of reporter plasmid, 1 μg of EF1α-β galactosidase expression vector and carrier DNA (fish sperm DNA, Boehringer Mannheim) up to 20 μg (DNA mix), was formed by adding dropwise an equal volume of DNA mix to a 2x HBS buffer (1.4 mM Na2HPO4, 12 mM glucose, 27.4 mM NaCl, 10 mM KCl, 40 mM Hepes) and added to cells for 16 hours. After removal of the precipitate cell extracts were prepared 48 hours after transfection by the three freeze and thawing cycles of the cells previously resuspended in 100 μl of Tris-HCl 0.25 M, pH 7.8 (Sambrook et al., 1989).

The CAT protein produced was quantitated employing a non radioactive CAT-ELISA kit (Boehringer Mannheim) and the amount was normalised to β-galactosidase activity (Sambrook et al., 1989).

*DNA plasmid construction*

All DNA manipulations were carried out by standard techniques (Sambrook et al., 1989) and the constructs verified by restriction analysis and DNA sequencing.

The -2212 to +30 region of the human uPA promoter (EcoRI/SmaI in p-uPA-CAT-2212; Verde et al., 1988) was cloned in the EcoRI/SmaI site of pGEM7Zf(+)...
vector (Promega; p2212). The Hind III/BamHI CAT cassette from pEMBL8-CAT (Colantuoni et al., 1987) was cloned in the Hind III/BamHI sites of p huPA2212 plasmid. The derived plasmid (p2212CAT) is used as the basic construct for all the following manipulations.

5' deletions

1870CAT, 660CAT, 787CAT: The fragments EcoRI/EcoRV (-1870/-537) and EcoRI/Smal (-660/+30) from p-uPA-CAT-1870 and p-uPA-CAT-660, respectively, (Verde et al., 1988) are blunt ended and cloned in the blunt XhoI and EcoRV (1870CAT) or Smal (660CAT) sites of p2212CAT digested plasmid.

For 787CAT the cloned region is Hpal/EcoRV (-787/-537) from p2212.

537CAT and 86CAT (MPCAT): p2212CAT is digested respectively with XhoI/EcoRV and XhoI/Eco47III, blunt ended and religated.

Internal deletions

E537CAT (ES3MPCAT), EMPCAT, EΔS3CAT: They are obtained by digestion and religation of p2212CAT vector with the following restriction enzymes: Bsu36I/EcoRV (E537CAT); Bsu36I/Eco47III (EMPCAT); EcoRV/Eco47III (p2212ΔS3CAT).

E1428CAT, E787CAT, E660CAT: The -1428/-537 (StuI/EcoRV), -787/-537 (Hpal/EcoRV) and -660/-537 (AluI/EcoRV) fragments from p2212 were cloned in Bsu36I/EcoRV linearized p2212CAT, respectively.

EΔS2CAT: The p2212 plasmid is digested Hpal/EcoRV and religated (p2212ΔS2). The StuI/Eco47III from p2212ΔS2 is then cloned in the StuI/Eco47III linearized p2212CAT (EΔS2CAT).

ES1MPCAT, S1MPCAT: They are obtained by StuI and Eco47III digestion and religation of p2212CAT and pl870CAT, respectively.

ES2MPCAT, S2MPCAT: pE787CAT and p787CAT are digested with EcoRV and Eco47III and religated to give respectively, pES2MPCAT and pS2MPCAT.
Materials and Methods

SV40 derived plasmids

The -787/-537 S2 region (HpaI/EcoRV in the huPA promoter) is cloned in both orientation in the blunt ended Bgl II site (pCATsil and pCATsil inv) and in the XbaI site (pCATsil 3') of pCAT control vector (Promega) which contains the SV40 promoter and enhancer.

S2 internal deletions

The deletions of the regions internal to the S2 silencer have been obtained by recombinant PCR according to Vallette et al. (1989).

The primers used for primary amplification were the following:

$\Delta 1$: 5' TTGGTGCTTCTCTTTTGGA 3' (#Δ1cs) and 5'AGCTCTGCCTTCTCTGTA 3' (#U3BE; -1500/-1474 in the uPA promoter), 5'TCCCAAAAGAGAGACAACAA 3' (#Δ1) and 5'AGTGGGAGTCCCCCAGATATCA 3' (#5 low; -520/-540); $\Delta 2$: 5'ATGATGGTTACGCATAAAACTG3' (#Δ2cs) and #U3BE, 5'CAGTTTATGCGTAACATCAT 3' (#Δ2) and #5 low; $\Delta 3$: 5'ATTACTGAAACAGGGCTTGA 3' (#Δ3cs) and #U3BE, 5'TCAAGCCCTTGTTTCAGTAAT 3' (#Δ3) and #5 low; $\Delta 4$: 5'TTIGATGGTGAAGTAAAATT 3' (#Δ4cs) and #U3BE, 5'AAATTTTCAGCAGAGATCAA 3' (#Δ4) and #5 low; $\Delta 5$: 5'ATGGAACAAAGCCGGCTTA 3' (#Δ4cs) and #U3BE, 5'ATGGAACAAAGCCGGCTTA 3' (#Δ5 low; -557/-576) and #U3BE.

After mixing together the primary PCR products, the second amplification has been carried out using as external primers #U3BE and #5 low (pΔ1 to pΔ4). The final products are digested Stu I/ EcoRV and cloned in StuI/EcoRV linearized p2212CAT (p2212S2Δ1 to Δ4).

To obtain Δ5 construct the primary product is directly digested with Stu I or HpaI, and cloned in the StuI/ EcoRV digested p2212CAT (p2212S2Δ5) or in the blunt Bgl II site of pCAT control (psilΔ5), respectively.

For the other S2-SV40 derived constructs the psilΔ2 to psilΔ4 plasmids were derived by digestion of p2212S2Δ2 to Δ4 with HpaI/EcoRV and cloned in the blunt
ended Bgl II site of pCAT control. psilΔ1 was obtained by PCR using as template p2212S2Δ1CAT construct and the following primers: 5’ TCCCAAAAGAGAAGGACACCAA 3’ (#Δ1) and #5 low. The fragment is then digested HpaI/EcoRV and cloned in the blunted Bgl II site of pCAT control.

$p\Delta 1\Delta 2$, $p\Delta 3\Delta 4$, $p3’$ half, $p5’$ half: These constructs are obtained by the same strategy. In the case of $p\Delta 1\Delta 2$ and $p\Delta 3\Delta 4$ the DNA templates were psilΔ1 and psilΔ3, respectively. For $p$ 3’-half and $p$ 5’-half the template was p2212CAT.

In the first step of amplification the primers are 5’ AAGTTGGGTAACGCCAGGGT 3’ (#pCAT up, 341/420 in pCAT control vector) and #Δ2cs, 5’ AGCTTCCTTAGCTCCTGAAA 3’ (#EC13; 605/638 in pCAT control) and #Δ2 (pΔ1Δ2); #pCAT up and 5’ GTGACCTGAAACAGGGCTTGA 3’ (#Δ3Δ4cs), 5’ TCAAGCCCTGTTCAGTCAC 3’ (#Δ3Δ4) and #EC13 (pΔ3Δ4). The second amplification is performed using as external primers #pCAT up and #EC13. The products are then digested Sau3AI and cloned in Bgl II site of pCAT control.

The $p$ 5’-half and $p$ 3’-half constructs are derived by a single amplification step with the following primers 5’ CAGTAATCTGGCCTTCTTCC 3’ and #5 low (p3’-half), 5’ ACTGAAAATTTCAGGTCTGG 3’ and 5’ TCCCAAGAGCCGCTTAACAC 3’ (#HpaI) for $p$ 5’-half.

The fragments are then digested EcoRV (p 3’-half) and Hpa I (p 5’-half) and cloned in blunted Bgl II site of pCAT control.


**Nuclear extract and SouthWestern analysis**

Nuclear extracts were obtained with a modified method of Dignam et al., 1983 (Ausubel et al., 1994).

Briefly, HeLa and PC-3 cells were grown in 150 mm plates to 80% confluence in DMEM supplemented with 10% foetal calf serum, penicillin (100 U/ml), streptomycin (100 μg/ml) and glutamine (2 mg/ml).

After harvesting, the cells are resuspended in a volume equal to 3 packed cell volume (pcv) of hypotonic buffer (10 mM HEPES pH 7.9 at 4°C, 1.5 mM MgCl₂, 10 mM KCl, 0.2 mM PMSE, 0.5 mM DTT), swollen on ice for 10 min. and homogenised with 10 strokes (type B pestle). The nuclei are pelleted by centrifugation at 3300xg for 15 min. and resuspended in 1/2 packed nuclei volume (pnv) of low-salt buffer (20 mM HEPES pH 7.9 at 4°C; 25% glycerol, 1.5 mM MgCl₂, 20 mM KCl, 0.2 mM EDTA, 0.2 mM PMSE, 0.5 mM DTT).

In a dropwise fashion a volume equal to 1/2 pnv of low-salt buffer containing 1.6 M KCl (high-salt buffer) was added (0.4 M KCl final concentration) and the nuclei were extracted for 30 minutes at 4°C with continuous gentle mixing.

After centrifugation the nuclear extract is dialysed against 50 volumes of a buffer containing 20 mM HEPES pH 7.9 at 4°C; 20% glycerol; 100 mM KCl; 0.2 mM EDTA; 0.5 mM DTT; 0.5 mM PMSF for 4-16 hours. After centrifugation, to remove the precipitated protein, the extract is aliquoted, frozen in liquid nitrogen and store at -80°C. The protein concentration is determined by the Bradford method (BioRad; Bradford M., 1976).

For column chromatography fractionation, purified HeLa cells nuclei, obtained by 5x10⁹ growing exponential cells (Ghent, Belgium), are extracted according to the method of Dignam et al. (1983). Briefly, the nuclei pellet (10 ml) is resuspended in 2.5 volume of buffer C (20 mM Hepes pH 7.9, 25% Glycerol, 0.42 M
NaCl, 1.5 mM MgCl₂, 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF, 1x Aprotinin, 1 μg/ml Leupeptin, 1μg/ml Pepstatina A, 1 mM Sodium bisulphite) and stirred on ice for 60 min. Nuclear extract is then clarified by ultracentrifugation at 40000 rpm in the 70.1 Ti rotor (Beckman) at 4°C for 1 hour and dialysed in 50 volume of buffer D.

29 OD₂₈₀ are loaded on a 25 ml Heparin Sepharose column (Pharmacia) pre-equilibrated in Buffer D. The column has been washed with 2 volume of the same buffer containing 200 mM KCl (Fraction W; 3.55 OD₂₈₀) and eluted with 10 volume of 200 mM - 1 M KCl linear gradient made in buffer D.

50 μg of fraction W are sized fractionated on a 10% SDS-polyacrylamide gel and the protein transferred to a nitrocellulose filter (250 mA, 4 hrs in transfer buffer: 191 mM glycine, 25 mM Tris, 20% methanol). The proteins are then renatured by incubating the filter in 100 ml of renaturation buffer (10mM Tris-HCl pH 7.5, 10 mM Magnesium Acetate, 10 mM KCl, 10 mM 2-mercaptoethanol, 0.1 mM EDTA, 10% glycerol, 1x Denhardt's solution) for 2 hours with 3 changes.

Binding is carried out with 5 ml of a mixture containing 1x binding buffer (see above) and 5x10⁵ cpm of labelled oligonucleotide (~100 fmoles) as described in the legend of the figure. Competitor DNA is added as fold molar excess.

**DNase I protection analysis and Electrophoretic mobility shift assay (EMSA)**

40 to 80 μg of nuclear extract are used in DNase I footprinting experiment. The reaction is performed in 50 μl of a mixture containing 25 μl of 2X binding buffer (40 mM HEPES pH7.9; 4 mM MgCl₂; 100 mM KCl; 1 mM EDTA; 2 mM DTT; 20% glycerol) and 2 μg of poly dI-C (Pharmacia) and incubated on ice for 30 min. After adding 20.000 cpm of ³²P-S2 labelled fragment the incubation proceeds for 30 min at r.t.

DNase I treatment is performed adding to each sample 50 μl of MgCl₂ 10 mM/ CaCl₂ 5mM and DNase I (1ng/μg of extract) freshly diluted in the same
buffer. After 2 min. at r.t. digestion was stopped by adding 100 μl stop solution (20 mM Tris pH 7.5; 200 mM NaCl; 20 mM EDTA; 1% SDS; 100 μg/ml proteinase K; 50 μg/ml tRNA) and incubating the samples 30 min. at 60°C. Samples are then phenol/ chloroform extracted and DNA is precipitated with 1/10 vol of sodium acetate pH 7.0 and 2.5 vol. of cold EtOH.

DNA was then separated on 6% or 8% denaturing polyacrylamide gel, using standard Maxam-Gilbert G+A sequencing reaction of the fragment as a marker (Sambrook et al., 1989).

For EMSA analysis, 3-5 μg of nuclear extract are preincubated on ice for 30 minutes in 25 μl final volume of a mixture containing 20 mM HEPES pH 7.9; 2 mM MgCl2; 50 mM KCl; 0.5 mM EDTA; 1 mM DTT; 10% glycerol and 2 μg of poly dI-C (Pharmacia). 10 fmols of labelled oligonucleotide DNA probes and molar excess of cold competitors (see legend of figures) are added simultaneously and incubated for 30 minutes at room temperature. Nucleoproteic complexes are then analysed on 5% polyacrylamide gels in 0.5x TBE run at 7 watts for 90 minutes. Prior to loading, gels are prerun at 300 volts for 30'. The gels were dried and autoradiographed using intensifying screens at -80°C for 24-48 hours.

**Circular permutation assay**

The HpaI/EcoRV (-787/-537) S2 region and the annealed synthetic oligonucleotide #H (5' TAAGCCCGGCTTTGTTCCAT 3') are cloned in the XbaI site of pBend2 plasmid (Kim et al., 1989). Circular permuted fragments are obtained by restriction with different restriction enzymes in the tandemly repeated polylinker of pBend2. Each fragment is then labelled by Klenow filling in using α32P-dCTP or by Polynucleotide kinase and γ32P-ATP (Sambrook et al., 1989) and incubated in the presence of titrated amount of human SRY-HMG box (kindly provided by M. Bianchi, DIBIT, Milan; Ferrari et al., 1992), human SOX-4 HMG box and human
TCF-1 HMG box provided by M. van de Wetering (University of Utrecht, The Netherlands; van de Wetering and Clevers, 1992; van de Wetering et al., 1993). The binding is carried out in 10 µl (final volume) of a buffer containing 8% Ficoll, 100 mM NaCl, 10 mM Hepes pH 7.9 for 10 min on ice. Electrophoresis is carried out as previously described.

The mobility of protein-DNA complexes was normalised to the mobility of free DNA (R_{bound}/R_{free}, vertical axis in the graphs); the distance between the 5' end of the probe was normalised to the total length of the probe (flexure displacement, horizontal axis in the graphs). The points in the graphs were interpolated with a second-order equation by means of least squares algorithm (Cricket Graph application on a Macintosh computer).
4. Results

Cell type specificity of different portions of the uPA gene regulatory region.

DNA constructs carrying different portions of the uPA regulatory region (from -2212 to +30) fused to the CAT reporter gene have been transfected in cells either not producing (CV1 and HeLa cells) or producing high constitutive level of uPA (PC-3).

As shown in Table 4.1, the transcriptional activity of the minimal promoter (-86 to +30; pMPCAT) was unexpectedly higher (about 2.5 fold) in CV1 than in PC-3 and HeLa cells. When I used the full length construct p2212CAT, which includes the enhancer, the minimal promoter and the sequences spanning the two regions, an increase of only 3 fold over the minimal promoter activity was observed in PC-3 cells, while in CV1 and HeLa cells the transcriptional activity dropped by about three and five fold (Table 4.1). When a fragment containing the enhancer (-2212 to -1870) was added upstream of the minimal promoter (pEMPCAT), the activity was increased in all three cell lines tested although with different efficiency: seven and six fold in the case of CV1 and HeLa and 25 fold in PC-3 cells (Table 4.1).

These data showed that the activity of the minimal promoter, cell-type dependent, is in all cases increased by the enhancer and that the sequences between the enhancer and the minimal promoter modulate enhancer activation in a cell-specific way by downregulating transcription.
Table 4.1: Cell-type specificity of uPA promoter activity
Plasmid constructs with the CAT gene under the control of the indicated uPA subunit promoter regions were tested, in transient transfection, for CAT activity in cell lines either not producing (CV-1 and HeLa) or producing high constitutive level of uPA (PC-3). CAT values are normalised to β-galactosidase activity and they are the means ± standard deviation (SD) of a number of independent experiments as indicated in parenthesis.

Multiple silencing activities in the 5' flanking region of the human uPA gene.

In order to investigate in greater detail the mechanism of negative regulation of transcription in the uPA gene, 5' deletions of the p2212 CAT construct were tested in CV1 and in PC-3 cells (Figure 4.1). Deletion of the enhancer region reduced the activity of the construct (compare p2212 to p1870); further deletions (up to -537) restored the activity to that of the p2212 construct in CV-1, but did not have a significant effect in PC-3 cells. Only the deletion of the region spanning from -537 to -86 increased transcription in both cell lines (Figure 4.1).
Next, I prepared internal deletions leaving the enhancer and the minimal promoter (-86) intact, in order to analyse the effects of the negative region(s) on the enhancer activity (Figure 4.2). In PC-3 cells, deletions of the region -1820 to -537 minimally (two-fold) increased transcription. However, deletion of the -537 to -86 region (called S3) caused a stronger increase in CAT activity. I then conclude that in this cell line the only two functionally most important regions are the enhancer and S3 (Fig. 4.1 and 4.2).

In CV1 and HeLa cells, however, regions -1820 to -1428 (S1), -787 to -537 (S2) appeared to contribute to the down regulation of transcription with an additive effect to that of -537 to -86 (S3). Deletion of S1 increased transcription about 2.5 fold in CV1 and five-fold in HeLa. Deletions of both S1 and S2 (-1820 to -537) had an about seven fold effect, and deletion of all three increased transcription 20 to 25 fold. Deletion of either S2 or S3 region alone confirmed their role as cis-acting
negative regulatory elements with S1 and S2 showing greater cell type specific activity (compare CV1 and HeLa versus PC-3) than S3.

Constructs (E787CAT, E660CAT, E537CAT, and EMPCAT) where the enhancer is located downstream the CAT gene showed the same activity, excluding that the increased CAT expression was due to a position effect of the enhancer closer and closer to the transcription start site (data not shown).

Furthermore, the deletion analysis suggested that S2 activity depends on the enhancer, as its activity was not detected in the 5' nested deletions of figure 4.1. S1 and S3, instead, appear to modulate the activity of the minimal promoter.
In order to test this possibility I cloned the S1, S2 and S3 regions in front of the uPA minimal promoter in the presence or in the absence of the enhancer. As it is shown in figure 4.3, in CV1 cells all three regions S1, S2 and S3 repress the activity of the constructs by the same level either in the presence or in the absence of the enhancer. From these data it is possible to conclude that these silencing regions suppress the activity of the minimal promoter.

![Diagram showing the effect of S1, S2, and S3 on the uPA minimal promoter and enhancer activity in CV1 cells.](image)

**Fig. 4.3: S1, S2 and S3 effect on the uPA minimal promoter and enhancer activity in CV1 cells.**

The negative regulatory elements S1, S2 and S3 are cloned in front of the minimal promoter in the presence or absence of huPA enhancer and their activity is tested in CV1 cell line. The values are expressed as fold repression (pEMPCAT/pESMPCAT and pMPCAT/pSMPCAT, respectively) ± SD. In parenthesis the number of independent experiments is indicated.

**S2 has the properties of a silencer**

Of the three negative regulatory regions only S2 showed to act in a cell-type specific manner, being active only in those cells that do not produce uPA. I then decided to further characterise the activity of the S2 region cloning the -787 to -537
region upstream of the SV40 enhancer/promoter driven CAT gene (pCATsil). As shown in Figure 4.4, the S2 sequence represses the SV40 enhancer-promoter in CV1 and HeLa cells (about 4 fold) but not in PC-3 cells and its function is orientation-independent (compare pCAT sil and pCAT sil inv). Similar results have been obtained with the S2 region cloned downstream the reporter gene, confirming that the increased activity does not depend on the relative distance with respect to the transcription start site. On the contrary, the S2 sequence appears to have the properties of a silencer, acting also outside of the uPA gene context and in an orientation- and position independent way.

![Diagram](image)

**Fig. 4.4: Effect of S2 on SV40 promoter/enhancer driven transcription of CAT gene in CV1, HeLa and PC-3 cell lines.**

Region S2 has been cloned in either orientation and 5' or 3' of SV40 promoter. The constructs are tested by transient transfection in HeLa (yellow bar), CV1 (red bar) and PC-3 (green bar) cell lines. The CAT activity is expressed as relative to the pCAT control construct for each cell line ± standard deviation. The data are the mean of 4 independent experiments.
**Silencer S2 is made up of multiple silencing units**

Comparison of the S2 sequence of human versus murine and porcine uPA 5' flanking regions, shows several interspersed regions of high homology (Fig. 4.5), identifying putative murine and porcine S2 elements.

**Fig. 4.5: Comparison between human uPA S2 and murine and porcine uPA promoters.**

Identical residues between human, mouse and pig uPA promoter are indicated by a bar, while a rhomboid indicates the same residues between human and porcine uPA promoters. 1 to 6 defines the six regions of high homology. Below, the genomic organisation of region 1 to 6 in human, mouse and pig uPA 5' flanking region is drawn.
In the mouse, however, these sequences span a wider region (-1007 to -603) in contrast to the porcine and uPA promoter that encompass 285 and 250 base pairs respectively. Regions of non-homology separate the different modules with a degree of similarity higher between human and pig S2 than between mouse and the two other species.

The putative murine S2 is characterised by long insertion with respect to pig and human S2 element. However the highest homology resides in the most 5' part of the element with module 1, 2 and 3 showing an about 80% homology between the three species. Modules 4 and 5, although highly homologous between mouse and man, are less conserved in the porcine promoter, showing the presence of inserted and deleted residues. Module 6 is present downstream module 5 both in mouse and porcine uPA promoters.

The conserved sequences may represent functionally active regions within the S2 silencer. Therefore, I deleted such sequences from the human uPA regulatory region (deletions Δ1 to Δ5, see Fig. 4.5) and analysed the effect in the context of both the p2212CAT and S2-SV40 CAT constructs by transfection in CV1 cells. As shown in Figure 4.6, deletion of any single homology domain did not significantly affect transcription in the natural context of p2212CAT, with the possible exception of Δ5 which moderately decreased the silencing activity.

However, when the same deletions were analysed in the S2-SV40CAT construct, where the absolute activity of CAT is one order of magnitude higher than with the uPA promoter, significant effects were observed.

Indeed, deletion of anyone of the homology regions led to a decrease in silencing activity of at least 50% (pΔ1 to pΔ5 versus pCAT sil). The combined deletion Δ1+Δ2 showed a reduction in silencing activity of about 70%, while Δ3+Δ4 was able to totally abolish the S2 negative effect.
Fig. 4.6: Characterisation of silencer S2
The functional analysis of region 1 to 5 of silencer S2 is made in CV1 cells. The activity is indicated both as picograms of CAT/unit of β-galactosidase and fold repression. p2212 CAT is for the deleted constructs in the context of huPA2212CAT. pS2-SV40 is for the constructs analysed in the context of SV40 enhancer/promoter driven transcription. In parenthesis the number of independent experiments is indicated. (-) is for not tested.

In order to test whether sequences other than those conserved between man and mouse were also involved in silencing, I deleted the entire 5' or 3' half of S2. As shown in Figure 4.6, in both cases the activity of the silencer was totally lost (p5'-half and p3'-half).
These data indicate that the conserved as well as non conserved regions have a functional significance, but also show that multiple regions contribute to the activity and that some of them are redundant.

**DNasel footprinting analysis of S2**

In order to investigate whether the silencing activity was modulated by the interaction of nuclear proteins to the S2 sequence, DNase I footprinting analysis was carried out with increasing amount of nuclear extract from uPA producing (PC-3) and not producing (HeLa) cells. Both strands of S2 region were subjected to the analysis, that revealed a pattern of protected regions covering almost the entire 250 bp long S2 fragment either with HeLa or PC3 nuclear extract already at low concentration of nuclear extract (Fig. 4.7). The protected areas encompass both the 5 regions homologous between the promoter of human, mouse and pig uPA gene (open red boxes) and some of the sequences in between (Fig. 4.8).

The differences between the PC3 and HeLa extracts reside in the regions -627/-600 and -711/-704, where protection is observed only in the presence of HeLa nuclear extract, more and more evident at higher concentration of nuclear extract, and in some DNase I hyper- or hypo-sensitive sites, with the -772 hypersensitive site, in the lower strand, specific for HeLa nuclear extract. Region 3 does not show DNase I protection. However, in the absence of nuclear extract, this sequence is cleaved by DNase I only at position -666 and -657 in the upper strand, and shows a few cuts in the lower strand at position -665, -664, -660, -659 and -656 (Fig. 4.7). Therefore I cannot exclude that factor(s) bind this region (see also Discussion).
Fig. 4.7: DNase I footprinting analysis of S2 region.

40, 60 and 80 µg of nuclear extract from HeLa or PC-3 cells are incubated with either the upper or lower strand labelled S2 region and subjected to digestion by DNase I. The red lines indicate the region of homology with the murine uPA promoter; the blue boxes the region protected by either nuclear extracts, whereas the orange boxes are for the areas protected by HeLa nuclear extract only. The arrows indicate the hypersensitive or hyposensitive sites.
A close look of the sequence of region S2 shows the presence of several direct and inverted repeats (Fig. 4.9). In particular, the sequence between -783/-775 is repeated at -762/-754 with the difference in two residues (green type); the one spanning region -746/-735 (red type) is present at position -570/-560, with the only insertion of an A residue in the former, and -596/-604 in the lower strand; the sequence TGATA, that matches with the consensus for the GATA-1 binding factor, is present in the lower strand at position -688/-692 and -712/-716 and in the upper strand between -618/-614 (pink type); brown types show a direct repeat of the sequence CAATCCT at position -732/-726 and -699/-693; the CCCCCT motif is present at -625/-620 and -598/-593 (orange type); the GAAAG sequence (light blue type) is in the upper strand between -585/-581 and in the lower strand at -626/-629. Furthermore, this sequence can be seen as a part of a longer motif that include also the red-typed sequences (-746 GGACTTTGTCC -735; -672 GGACTTTAACC -661; -632 TGCCCTTTCCCC -621; -581 CTTCGCGC -588 in the lower strand, and -570 GGCTTTGTGCC -560).
The CAAT motif is repeated four times along the entire S2 region and precisely at position -774, -732, -699 and -580, while the sequence ATGCCTA (-754/-747) is present other two times at position -688/-682 and -576/-571 with a mismatch (pink line). The TTTAAA motif, present in box 3, is repeated in the lower strand at the border of the same area (-648/-653). Finally, the sequence CCATC is present twice (-610/-606 and -561/-557), while sequence protected in footprint I is also present in the lower strand at position -605/-614 and -759/-769 (blue line).

The presence of the same sequences at different locations along the whole S2 fragment would explain why the fragment is entirely protected. Furthermore, because of this redundancy of sequence motifs, I did not study in detailed all the footprints but I focused my attention on those protections that are specifically present in HeLa nuclear extract and on those containing known sequence motifs that could be important in elucidating the way S2 silences transcription.

Fig. 4.9: Direct and inverted repeats in S2 silencer region
The regions protected by HeLa and PC-3 nuclear extracts are in orange and blue boxes, respectively. The red boxes indicate the conserved regions between mouse and human uPA. The direct and inverted sequences present in huPA S2 are identified by the same colour (as font or line).
EMSAs analysis of the complexes present in HeLa and PC3 nuclear extracts

**Footprint D/L**

Footprint D/L, encompassing region -726/-705, is composed by a part that is protected both in HeLa and PC-3 extracts and a part specific for HeLa cells. In order to analyse the nature of the complexes protecting these sequences, I performed a mobility shift assay (EMSA) with an oligonucleotide encompassing the region -737/-705 (oligonucleotide #1; Fig. 4.10, panel A). Two retarded bands are detected in the presence of both HeLa and PC-3 nuclear extracts, but they show a different mobility (1, 1*, 2, 3). The binding of both complexes is specific as they are competed by an excess of 50-fold unlabelled oligonucleotide #1, although the affinity seems to be higher for the complexes detected by HeLa extract than by PC-3 (50x vs 500x). A careful analysis of the sequences present in the region of footprint D/L showed a similarity with the combined binding site for the *Drosophila HOX gene Antennapedia (Antp)* and factors belonging to the PBX family of transcriptional cofactors (PBC-HOX, Fig. 4.10 panel B; Mann and Chan, 1996). The same sequence is present also in yeast in promoters (i.e. HSG-operator) that are silenced by the Mata1p/Mata2p complex (Mata1p is 65% identical to the homeodomain EXD). When the oligonucleotides OIB and H+P (Fig. 4.10, panel B), which are similar to the PBC+HOX site (Berthelsen *et al.*, 1988b), are used as competitor, either of them is not able to compete the complexes formed by oligonucleotide #1 in HeLa and PC-3 nuclear extracts.

These results tend to exclude the possibility that such factors bind the S2 silencer. However, the competition analysis should be performed in the presence of other oligonucleotides resembling more the element present in the HSG-operator, where the two half sites of the element are separated by 12 base pairs.
Fig. 4.10: EMSA analysis of footprint D+L

A: HeLa and PC-3 nuclear extracts are incubated in the presence of labelled oligonucleotide #1 as probe (+). Competition analysis has been carried out with 5, 50 and 500 fold excess of unlabelled oligonucleotides #1 and #OIB and with 5 and 50-fold excess for #H+P. B: sequences of oligonucleotides used in this study. The haploid-specific gene operator (HSG operator) is the binding site for the Mata1p-Mata2p complex; PBC-HOX site is the consensus for PBX-HOX complex where the two N are variable nucleotides and are proposed to contribute to HOX specificity; #1 is the oligonucleotide used as probe (-736/-705 in huPA promoter); # OIB and #H+P contain, respectively, the uPA promoter binding site for Prep1-Pbx complex and the binding site for Pbx-Hox complex as described in Berthelsen et al (1988a,b).

The sequence present in oligonucleotide #1 revealed also a 10/13 homology with the consensus sequence of the yeast mating type silencer to which the autonomous replicating sequence binding factor 1 (ABF-1) binds (Fig. 4.11b). This factor has been shown to bind also the silencer of the e-globin gene (Peters et al., 1993). The same sequence matches also with other protected regions in the huPA promoter (-613/-601 and -538/-550, Fig. 4.11c). However a competition analysis of
Results

the complexes formed by oligonucleotide #1 and either HeLa or PC-3 nuclear extracts showed that not even 500-fold excess of unlabelled #ABF-1 consensus oligonucleotide is able to compete the retarded bands detected by oligonucleotide #1 in EMSA (Fig. 4.11a).

Fig. 4.11: Competition analysis of footprint D+L with ABF-1 binding site
A: EMSA analysis of oligonucleotide #1 in the presence of HeLa and PC-3 nuclear extracts. Competition with 5, 50, 500-fold excess of unlabelled oligonucleotide #1 and #ABF-1 consensus. B: Comparison between sequences present in huPA promoter, oligonucleotide #1, e-globin silencer and ABF-1 consensus. C: Putative binding sites for GATA-1 (red-type) and ABF-1 (black-type) transcription factors in huPA S2 promoter region (-787/-538). The footprint M and N are shown in square brackets (yellow-type and green-type respectively). The lines above text are for sequence present in the upper strand, below is for lower strand. The repeated sequence in the huPA S2 region are underline in blue.

The most 5’ABF-1 site in uPA promoter differs from the ABF-1 consensus in the last three residues at the 3’ of the sequence. It would be interesting to test whether the recombinant ABF-1 protein would bind the uPA sequence and if the
missing competition is due to a different affinity of the protein for non canonical binding sites.

**Footprint F+M**

An oligonucleotide encompassing the region -656/-614 has been used to study the complexes binding this area. Two major complexes are detected in HeLa nuclear extract (Fig. 4.12a). Only complex 1 seems to be specific as it is competed already by a 30-fold excess of unlabelled oligonucleotide. The specificity of the complexes has been also tested by competition with a series of oligonucleotides, showed in fig.4.12b, covering the region of footprint N+G (oligo #3), H+I (oligo #4), N (oligo #N) and I (#TIE).

**Fig. 4.12: EMSA analysis of footprint F+M**

A: Oligo#2, encompassing the region of footprint F+M, is used as probe. Competition by an excess of specific unlabelled oligonucleotide is present in lane 2 and 3 (30- and 60-fold). Other competitors are oligo #4 (lane 4), oligo #3 (lane 5), oligo #AP-1 (lane 6), oligo #N (lane 9), oligo #5 (lane 10), oligo #TIE (lane 11). Lane 8: 50x oligo #2. Lanes 1 and 7: no competitor. B: Oligonucleotides used in EMSA analysis. S2 region: -786/-538.

Of all oligonucleotides tested only #N seems to compete complex 1. The lower intensity of the complexes in the presence of #TIE-oligo as competitor (Fig.
4.12a, lane 11) is due to a lower $^{32}$P-oligo #2 loading rather than to a competitive effect. The competition by oligo #N might be due to the presence of the CCCCCTT motif, although in the context of S2 the region covered by #N is not protected by HeLa nuclear extract. The same is not true with oligo #3 that is not able to compete any of the complexes although another motif (GAAAG) is present in both oligonucleotides.

**Footprint N+G**

An oligonucleotide spanning the region of footprint N+G (oligo #3; -613/-570) detects in HeLa nuclear extract 5 major and 3 minor complexes (Fig. 4.13, lane 1), 3 of which (4, 7 and 8) seems to be non specific as they are not competed by an excess (30- and 60-fold) unlabelled oligonucleotide (Fig. 4.13, lanes 2 and 3).

![EMSA analysis of footprinting N+G in HeLa nuclear extract](image)

**Fig. 4.13: EMSA analysis of footprinting N+G in HeLa nuclear extract**

A: An oligonucleotide encompassing the region -613/-570 (oligo #3) has been used as probe to detect complexes in the presence of HeLa nuclear extract. The specificity has been assessed by competition analysis with 30 and 60-fold excess specific oligonucleotide (lanes 2 and 3), 50x oligonucleotide #4 (lane 4), 50x oligonucleotide #2 (lane 5), 50x oligonucleotide #AP-1 (lane 6 and 11), 50x oligonucleotide oligo #TIE (lane 7), 50x oligonucleotide #H (lane 8), 50x oligonucleotide #N (lane 9), 50x oligonucleotide #5 (lane 10). Lane 1: no competitor. B: Oligonucleotides used in EMSA analysis. S2 region: -786/-538.
Oligo #2 (lane 5), as well as oligo #4 (lane 4), is not able to compete any of the complexes. Also an AP-1 consensus oligonucleotide (lanes 6 and 10) shows no ability to compete. Oligonucleotide #N, covering part of oligo #3, seems to compete two of the non specific complexes (4 and 8; lane 9).

Finally, in the presence of oligonucleotide #H the intensity of complex 8 increases, suggesting an increased stability for this complex.

Footprint N

In order to better investigate the footprint specifically observed with HeLa nuclear extract, the oligo #N, spanning region -621/-593, was used. 5 major complexes are detected in HeLa nuclear extract (Fig. 4.14).

![Fig. 4.14: EMSA analysis of region N](image)

The oligonucleotide #N is used as probe to detect complexes in the presence of HeLa nuclear extract. The following competitors are used as 50-fold excess oligo #N: #4 (lane 2), #2 (lane 3), #3 (lane 4), #TIE (lane 5), #H (lane 6), #N (lane 7), #5 (lane 8), #AP-1 (lane 9). Lane 1: no competitors.

Complex 3 is not specific since none of the oligonucleotide tested compete for this binding. Complex 1 and 2 are specific and are competed also by oligo #3 (lane 4). Complex 4 is competed by oligo #4, #3, #TIE and #AP-1. In the presence of the
Results

oligo encompassing the region of footprint H (-576/-557) complex 4 disappears, substituted by a faster migrating complex (lane 6).

Footprint C+H

Two potential HMG-box containing protein binding sites are found within the human S2 region: at -782/-771 and at -570/-560 (Fig. 4.15). In mouse and pig uPA promoters, the latter sequence is found downstream module 5, while it precedes it in man (Fig. 4.5, region 6).

Fig. 4.15: The HMG-box consensus sequence in the human urokinase promoter
The two consensus are shown in blue. In red the TIE element

Proteins belonging to this family of transcription factors are listed in Table 4.2. For some of them (Lef-1) an architectural role, beside the ability of transactivate transcription, has been outlined (Giese et al., 1995). For others, like SOX-4, a clear role as activators of transcription has been demonstrated (van de Wetering et al., 1993). To this family belongs the sex-determining factor SRY (Sinclair et al., 1990), whose potential target genes are starting to be identified (Cohen et al., 1994). Though these factors can bind the same sequence, some specificity has been shown in an in vitro analysis by EMSA (see Table 4.2).
The High Mobility Group box (HMG-box) transcription factors

<table>
<thead>
<tr>
<th>Protein</th>
<th>HMG Domains</th>
<th>Sequence recognition and/or DNAse I footprinting</th>
<th>Biochemical or genetic function</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEF-1/TCF-la</td>
<td>1</td>
<td>5' CCTTTGAAGCT 3'</td>
<td>• Lymphoid transcription (preB;proB;T cell)</td>
</tr>
<tr>
<td>TCF-1</td>
<td>1</td>
<td>5' CCTTTGAAGCT 3'</td>
<td>• Lymphoid transcription (T cell lineage)</td>
</tr>
<tr>
<td>SRY</td>
<td>1</td>
<td>5' CCTTTGTGCT 3' 5' CCTTTGTCT 3' 5' CCTTGAA 3'</td>
<td>• Sex determination (testis)</td>
</tr>
<tr>
<td>SOX-4</td>
<td>1</td>
<td>5' CTTGTG 3'</td>
<td>• Lymphoid transcription (T cell; pre-B)</td>
</tr>
<tr>
<td>SOX-5</td>
<td>1</td>
<td>5' ATTGTT 3' 5' CTTGTG 3'</td>
<td>• mRNA in gonads adult mice, lymphonodes, lung, heart</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Post-meiotic cells in testis (round spermatids)</td>
</tr>
</tbody>
</table>

Table 4.2: The HMG-box family of transcription factors

Fig. 4.16: EMSA analysis of footprint H

The oligonucleotide spanning the region of footprint H (HMG-oligo) has been used to analyse the presence of an HMG-box containing protein in HeLa, PC-3 and CV1 nuclear extract. Lanes 2, 4 and 6: the binding is made up in the presence of 30-fold excess unlabelled oligonucleotide. Lanes 1, 3 and 5: no competitor.
These elements, in the uPA promoter, are both protected either by HeLa or PC-3 nuclear extract. The specificity of such binding has been tested by EMSA analysis where an oligonucleotide covering the region present in footprint H detects only 1 complex with either HeLa, PC-3 or CV-1 nuclear extracts (Fig. 4.16).

In order to better characterise this complex I wanted to test whether the same oligo was able to bind purified SOX-4, SRY and TCF-1 proteins. At this purpose the purified SOX-4 HMG-box and TCF-1 HMG-box were used in band-shift analysis (Van de Wetering and Clevers, 1992). Only SOX-4 HMG-box was able to recognise uPA oligonucleotide (Fig. 4.17).

![Fig. 4.17: Binding of HMG-box containing proteins to the huPA promoter](image)

*The HMG-oligo has been cloned in the tandemly repeated poly linker of pBend2 (see Materials and methods). The Eco RV fragment has been used as probe to detect binding of the purified human SOX-4 and TCF-1 HMG-boxes. Increasing amount of proteins have been tested. The open arrow indicate the retarded complex. The closed arrow indicate the free probe.*

As these proteins have the ability to induce a sharp bend in the DNA double helix, the same oligonucleotide was tested in a circular permutation analysis in the presence of SOX-4, SRY and TCF-1 HMG-boxes. As shown in Fig. 4.18 (panel
B), again TCF-1 does not recognise the uPA sequence, while both SRY and SOX-4 HMG-box do.

![Circular permutation analysis of the -576/-557 region of the huPA promoter](image)

**Fig. 4.18:** Circular permutation analysis of the -576/-557 region of the huPA promoter

The HMG-oligo is cloned in the tandemly repeated polylinker of pBend2 plasmid. By using different restriction enzymes, circular permuted fragments of 146 bp have been generated (panel A). Panel B: Electrophoretic mobility of the circular permuted fragments in the presence of purified hSOX-4, hSRY and hTCF-1 HMG-boxes. The arrow indicates the free probe. Panel C: Analysis of the bending parameters. The mobility of the protein-DNA complexes (R bound) was normalised to the mobility of the corresponding free DNA (R free). The distance of the centre of the oligonucleotide HMG from the 5' end of the probe was divided by the total length of the probe (flexure displacement). The plotted points are interpolated with a quadratic function (see Materials and Methods).

Furthermore these proteins induce, upon binding, a bend in the DNA as shown by the different mobility of the complex depending on the different location of the oligonucleotide with respect to the ends of the DNA fragment (Fig. 4.18, panel A). No differences in the mobility of the free DNA probes are observed,
indicating that the sequence is not intrinsically curved. The binding of hSOX-4 HMG-box, however, caused a deviation of the axis of DNA of about 43°, while hSRY HMG-box bends DNA of about 74°C. This results is in agreement with the published values for SRY (Ferrari et al., 1992, Giese et al., 1992).

A circular permutation analysis has been also performed with the entire S2 region. A gel electrophoresis analysis of circular permuted fragment at 20 °C and 4 °C revealed that S2 is not intrinsically curved (Fig. 4.19).

![Circular permutation assay of S2](https://example.com/circular_permutation_assay)

**Fig. 4.19: Circular permutation assay of S2**
The whole S2 sequence is cloned in the polylinker of pBend2 plasmid. The different fragments are generated by the restriction enzymes indicated in the figure. Bg=Bgl II, C=Clal I, X=Xho I, E=Eco RV, P=Pvu II, S=Sma I, St=Stu I, B=Bam HI. The draw is not in scale. The length of the fragment is the 370 bp (249 bp S2 + 121 bp polylinker). The run is performed at 20 °C and 4°C with the same result.

However, binding of SRY-box to these fragments revealed a different mobility typical of bent DNA, with an angle of deviation of about 46° (Fig. 4.20).
Fig. 4.20: Circular permutation analysis of S2-hSRY complex
The whole S2 region has been cloned in the tandemly repeated poly linker of pBend 2 plasmid. The autoradiograms show the complexes formed by fragments generated by different restriction enzymes and hSRY purified HMG-box (open arrow). B=Bgl II; S=Spe I; E= Eco RV; St= Stu I; Ba= Bam HI.

All these results suggest that a protein belonging to the SRY-SOX family might recognise and bind the sequence present in the urokinase promoter, and induce a bending not only locally but of all fragment, although there is no direct evidence that the factor present in HeLa and PC-3 nuclear extract is indeed belonging to the HMG-box family of transcription factors.

Footprint I
At -553/-538 there is a conserved TIE element (TGFß1 Inhibitory Element), tandemly repeated in the opposite orientation, which in the mouse is found at -666. The TIE element has been shown to be involved in silencing of several genes (Kerr et al., 1990; Pietenpol et al., 1991; see Appendix 1). Furthermore two other TIE-like sequence are present in the lower strand at position -769/-760 and -613/-604 (Table 4.3). Of these elements only one appears to be protected (-547/-538) in DNase I footprinting by either HeLa or PC-3 extract (footprint I), although this protection is not so clear in the lower strand for PC-3 extract. The most 5' part does
not show any protection, while the -613/-604 is protected only with extracts from HeLa cells (footprint N).

| TIE consensus       | 5' GNNTTGGtGa 3' |
| huPA                | 5' TGGTTGTGTTGGTGA 3' - 553/-538 |
|                     | 5' TGATGGTGC 3' - 605/-613 |
|                     | 5' TGTTGGTGC 3' - 759/-767 |

Table 4.3: Sequences of TIE-like element present in huPA promoter

In order to analyse the role of the TIE element, an EMSA analysis, in the presence of HeLa nuclear extract, has been performed with two oligonucleotides, #TIE and #4, spanning regions -555/-531 and -566/-537 respectively. Both oligonucleotides contain the TIE element. As shown in Fig. 4.21 (panel A and B) three complexes (1a, 2a, 3a and 1b, 2b, 3b respectively) are formed with either oligonucleotides, although they behave differently to competition analysis. Band 1a is not completely competed by the unlabelled TIE oligonucleotide (lane 5, panel A), showing partial lack of specificity, while 50 fold excess of unlabelled oligo #4 completely competes complex 1b (lane 2, panel B). Complexes 2a and 3a, formed by #TIE, on the other hand, are self-competed and, hence, specific, while only the faster migrating complex detected by #4 (complex 3b) is competed out by the homologous unlabelled oligonucleotide. The specificity of such retarded bands is further assessed by the lack of competition by oligonucleotides covering the most 3'-half of S2 (panel C). Complex 3a is competed by #3, #N and #4 (lanes 4, 7 and 2 respectively), while complex 3b shows only a partial competition by #3 (lane 4). In the presence of #N and #TIE (lane 5 and 7, panel B) the mobility of the complex 3b changes, becoming faster, with a greater effect in the presence of #TIE than #N as
well as competition by #H of complex 3a (lane 6, panel A). Complexes 3a and b are not competed by #2 (lanes 3, panel A and B).

Fig. 4.21: Competition analysis of TIE and #4 oligonucleotides
32p - #TIE (panel A) and 32p - #4 (panel B) are incubated in the presence of HeLa nuclear extract. 
A: Lane 1: no competitor; lane 2: 50-fold excess #4; lane 3: 50-fold excess #2; lane 4: 50-fold excess #3; lane 5: #TIE; lane 6: #H; lane 7: #N; lane 8: #AP-1; lane 9: #5. B: Lane 1: no competitor. Lane 2: 50-fold excess specific competitor. Oligo #2, #3, #N, #H, #TIE, #AP1 are used as competitors in lanes 3-8, respectively. C: Map of the oligonucleotides, used as competitors, in S2 region. D: The sequences of the oligonucleotides used are shown.

The specificity of complexes detected by #TIE is further assessed by point mutation analysis, of those residues that, according to the consensus (see table 4.3, capital letters, and table 4.4) may be important for TIE function (Fig. 4.22, panel A).
Table 4.4: Oligonucleotides used in the study of footprint I

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
<th>Length</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>#TIE</td>
<td>5' ACTGGTTGTGTTGGTGATATCTGGG 3'</td>
<td>531</td>
<td>-555/-531</td>
</tr>
<tr>
<td>mut 3</td>
<td>5' C-A-C-A-C-A 3'</td>
<td>3'</td>
<td></td>
</tr>
<tr>
<td>mut 4</td>
<td>5' C-C-C-C-C-C 3'</td>
<td>3'</td>
<td></td>
</tr>
<tr>
<td>#TIE rev</td>
<td>5' TCCACTGGTTGTG 3'</td>
<td>537</td>
<td>Δ543/537</td>
</tr>
<tr>
<td>#TIE dir</td>
<td>5' CCACTGGTTGTAATCT 3'</td>
<td>548</td>
<td>Δ552/548</td>
</tr>
<tr>
<td>#5</td>
<td>5' TGATATCTGGGGACTGCCACT 3'</td>
<td>521</td>
<td>-541/-521</td>
</tr>
</tbody>
</table>

Fig. 4.22: EMSA analysis of region I (TIE region)

The oligonucleotide termed TIE, encompassing the region of footprint 1, has been used as probe to detect complexes in the presence of HeLa nuclear extract. Panel A: the specificity of binding has been tested by competition analysis using 50 fold excess of the following oligonucleotides: wt TIE (lane 2), mutant 1 (lane 3), mutant 2 (lane 4), mutant 3 (lane 5), mutant 4 (lane 6), TIE dir (lane 7), TIE inv (lane 8), AP-1 (lane 9) and oligo #5 (lane 10). Lane 1: no competitor. Panel B: The affinity of the complexes for the wt oligonucleotide has been evaluated by competition with 5, 50 and 500-fold excess of unlabelled oligonucleotide as indicated in the figure.
Mutant 1, 2, 3 were unable to compete for binding of bands 1 and 2 but not for band 3, while mutant 4 was able to compete all bands (lanes 3-6, panel A). The affinity was tested by competition of 5, 50 and 500 fold excess oligonucleotide (Fig. 4.22, panel B). Furthermore oligonucleotides containing only half of the entire TIE sequence (#TIE dir and #TIE rev) are not able to compete any of the bands, as well as an oligonucleotide (#5) spanning region -541/-521. An AP-1 consensus oligonucleotide competes complex 3a.

These data are confirmed by direct binding of the mutant oligonucleotides to the HeLa nuclear extract, showing the increased affinity of the three complexes for mutant 4 (Fig. 4.23).

Fig. 4.23: EMSA analysis of TIE mutant oligonucleotides at different extract concentration
The TIE wt and the four mutants have been used as probes in EMSA in the presence of increasing amount of HeLa nuclear extract. (+): no extract.
When the same oligonucleotide (#TIE) is used with nuclear extract from PC-3 (Fig. 4.24), besides complex 1 and 2, an additional band, with intermediate mobility is present (complex 4). As for HeLa, band 1 is not completely competed by the unlabelled oligonucleotide, unlike the lower and the intermediate bands. Band 4 behaves as bands 1 and 2 in point mutation analysis as it is not competed by mutant 1, 2 and 3 and is completely competed by mutant 4. Furthermore PC-3 nuclear extract does not show the fastest band detected in HeLa (complex 3).

Fig. 4.24: EMSA analysis of TIE-binding protein in PC-3 nuclear extract
The TIE oligonucleotide is used as probe in the presence of an equal amount of PC-3 and HeLa nuclear extract. The specificity of the retarded complex with PC-3 NE has been tested by competition with 50-fold excess of unlabelled homologous oligonucleotide and TIE mutants. The retarded complexes present in both extracts are indicated by the black full arrows. The PC-3 specific and HeLa specific complexes are indicated by the white open arrow (4) and the black thin arrow (3) respectively.
Recently it has been reported that Sp1 is a member of a multigene family whose members have similar if not identical DNA binding activities (Hagen et al., 1992; Kingsley and Winoto, 1992). Furthermore, in vivo transfection experiments show that Sp1-mediated transcriptional activation of Sp1 responsive promoters are repressed by Sp3 (Hagen et al., 1994). Sp2 and Sp3 have been shown to bind to both GC- and GT-box sequences (Kingsley and Winoto, 1992; Hamann et al., 1994).

The TIE element resembles the binding site for GT-box binding proteins. Therefore, an oligonucleotide specific for Sp3 binding (Hagen et al., 1994) has been used in EMSA analysis to test whether it was able to compete for TIE binding and viceversa. As shown in Fig. 4.25 (panel A) the GT-oligo is not able to compete the binding to TIE oligonucleotide in neither HeLa or PC-3 extracts, and binding of GT-box oligonucleotide by HeLa nuclear extract is not competed by #TIE (Fig. 4.25, panel B).

**Fig. 4.25: GT-box competition analysis**

Panel A: #TIE oligo is used as probe. Lanes 1 and 4: no competitor. Lanes 2 and 5: 50-fold excess unlabelled #TIE; lanes 3 and 6: 50-fold excess #GT-box unlabelled oligo. Panel B: The binding is made using #GT-box oligo as probe. Lane 7: no competitor; lane 8 and 9: 50-fold excess of #GT-box and #TIE, respectively. Panel C: #GT-box oligonucleotide sequence (uteroglobin promoter; Hagen et al., 1994).
Since the HMG-like consensus sequence is recognised by both HeLa and PC-3 nuclear extract and because of the structural role ascribed to these factors, I wanted to investigate the role of HMG in TIE binding in both extracts. For this purpose an oligonucleotide encompassing both sequences (-571/-538) is used in EMSA (#TIE-HMG). Again, in the presence of HeLa nuclear extract two complexes are detected, while in PC-3 an additional, intermediate band is present (Fig. 4.26, lanes 1).

![Fig. 4.26: Comparison between #TIE and #TIE-HMG](image)

The #TIE (lanes 1-4) and the composite #TIE-HMG (lanes 5-8) oligonucleotides are used as probes in the presence of HeLa or PC-3 nuclear extracts. Lanes 1 and 4 do not contain competitor DNA. lanes 2 and 6: 50-fold excess unlabelled #TIE. Lanes 3 and 5: 50-fold excess unlabelled #TIE-HMG; lanes 4: 50-fold excess oligo #GT-box; lanes 8: 50-fold excess #HMG-oligo. The black arrows indicate the complexes formed by #TIE, the clear arrows those formed by #TIE-HMG.

The mobility of these complexes is however slightly faster compared to the TIE oligonucleotide, although the length of the oligonucleotide is longer than that of TIE. The specificity of the complexes is tested by competition with an excess of
unlabelled oligonucleotide as indicated in Fig. 4.26. It must be noted that HMG oligonucleotide (lanes 8) does not compete at all any of the complexes indicating that either proteins recognising the HMG consensus do not contribute to the binding, or that their affinity for the short, HMG-specific, oligonucleotide is much lower. The faster migration of the complexes might, in fact, suggest the formation of a structurally different DNA conformation, such as bent DNA, that migrates faster in a polyacrylamide gel.

**Functional analysis of the TIE element**

When the entire TIE element (Δ5; Fig. 4.6) was deleted either from the context of the full length 2212uPACAT construct or pCATsil, the silencing activity of S2 was reduced by 40-50% (Fig. 4.6). On the other hand, mutations 1, 2, 3 and 4 in TIE sequence, when analysed by EMSA, showed different binding properties in the presence of HeLa nuclear extract (Fig. 4.22 and Fig. 4.23). Therefore I decided to investigate the function of the TIE element by inserting the four mutations in the context of both the full size uPA CAT construct (Fig. 4.27, panel A) and of the construct containing only the enhancer and the -787/+1 region (Fig. 4.27, panel B), and transfecting them in CV-1 cells.

Unexpectedly, none of the mutants had lost its silencing activity (Fig. 4.27); actually, they had gained a better repression activity in the full size constructs. Only mutant 3, in the E787CAT context, shows a slight increase in CAT transcription (Fig. 4.27, panel B).

I have then tested the TIE-like sequence of uPA in the SV40 CAT system in CV1 cells: as shown in Fig. 4.28, this element (pTIE CAT) acted as a silencer in both orientations and in a copy number-dependent manner (p2xTIE and p4xTIE). The presence of the adjacent HMG-box proteins binding site marginally increased the silencing activity (pTIE/HMG CAT).
Results

Fig. 4.27: Transient transfection analysis of TIE mutants in CV-1 cells
Point site mutations have been inserted in the TIE element in the context of the full length promoter 2212uPACAT construct (panel A) and of the E787CAT construct (panel B). The different plasmids have been transfected in CV-1 cells. The data are the mean +/- SD of at least two independent experiments.

Fig. 4.28: Effect of TIE and HMG-box consensus regions on SV40 CAT driven transcription in CV 1 cell line
The oligonucleotide TIE, encompassing the region -558/-537, has been cloned in front of SV40 promoter in pCAT control plasmid, in single (p TIE CAT) or multiple copies with different orientation respect to the direction of transcription (p 2xTIE and p 4xTIE). The p TIE/HMG CAT construct contains the oligonucleotide spanning the region -568/-529. The data are the mean of three independent experiments ± standard deviation.
A single-strand DNA-binding activity (footprint H+I)

In some preparations of labelled oligo #4 DNA the fast migrating complex 3b (Fig. 4.21, panel B) was not detected. Although the double strand probe was always gel purified after labelling, a contamination by single strand oligonucleotide cannot be excluded. I then tested whether the presence of the fast migrating complex was due to an activity recognising the single strand rather than double strand DNA. For this purpose I used, as source of proteins, a Heparin Sepharose column enriched fraction, eluting at 200 mM KCl (fraction W; see Materials and Methods), in which only one activity, comigrating with complex 3b, was detected by double strand oligo #4 (Fig. 4.29, compare lanes 1 and 3).

Fig. 4.29: EMSA analysis of HeLa fraction W
Oligo #4 double strand form is used as probe. Lanes 1 and 2: binding in the presence of HeLa nuclear extract. Lanes 3-6: binding in the presence of fraction W (see methods). Complexes are competed by 30-fold excess oligo #4 double strand (lane 2 and 4), #4 upper strand (lane 5), and #4 lower strand (lane 6). Lanes 1 and 3: no competitor is added.
Surprisingly, the binding of proteins present in fraction W was not competed by an excess of unlabelled double strand oligo #4 (Fig. 4.29, lane 4). On the contrary either the upper or lower strand of oligo #4, used separately as competitors, were able to compete the binding of the complex to the double strand, labelled oligonucleotide (Fig. 4.29, lanes 5 and 6).

In order to address the specificity of such binding I extended the analysis using as probes, in EMSA experiments, either the labelled upper or lower strand of oligo #4. As shown in Fig. 4.30 the labelled upper strand (panel A) detects 4 major complexes (lane 1), all of which are competed by a 30-fold excess of unlabelled homologous oligonucleotide, as well as by the lower, complementary strand (lanes 2 and 3). The competition by the upper strand of #TIE oligo changed the mobility of complex 3, showing the same behaviour of the double strand form when used as competitor of the double strand labelled oligo #4 probe (see Fig. 4.21 panel B, lane 7). The lower strand of #TIE competes all complexes (Fig. 4.30, lane 5). Both the double strand forms of oligo #4 and #TIE do not change the binding capability of the proteins to the single strand DNA (lane 6 and 7).

The specificity is tested by point mutation analysis where a mutation in the TIE sequence partially competes the binding to the oligo wt (lane 8; #4 mut: 5' CTTTGTTCCATCCACTGGTTatatTTGGTGAT 3'). The mutation in the corresponding residues of the lower strand does not change the capability of competing (lane 9). Neither the upper or double strand form of oligo #H competes any of the complexes (lanes 10 and 12); only the lower strand of oligo #H totally competes complexes 1, 3 and 4 and partially complex 2 (lane 11). Finally, the double strand #4 does not bind any protein in this fraction (lane 13).

When the lower strand is used as probe (panel B) two complexes are detected (lane 14). Both are specific as they are competed by an excess of the same unlabelled oligonucleotide (lane 16) as well as by the upper strand of oligo #4
(lane 15). Surprisingly the lower strand of #TIE oligo does not compete any of the complexes (lane 18), while the upper one induces a slight shift of complex 1 (lane 17). Again oligo #4 and #TIE oligonucleotide, in their double strand form, were not able to compete the binding to the lower strand of oligo #4 (lanes 19 and 20). The mutation in the TIE element (#4 mut) completely impaired the ability of specific competition, while the mutation in the corresponding position of the upper strand did not (lanes 22 and 21, respectively). Oligo #H upper strand and double strand form compete only complex 2 (lanes 23 and 25), while #H lower strand does not compete any of the complexes (lane 24).

![Fig. 4.30: EMSA analysis with single-strand DNA probes](image)

The upper and lower strand of oligo #4 are labelled and used as probes in the presence of Fraction W (see Methods; lanes 1 and 14). Competition analysis is performed with 50-fold excess of the following unlabelled oligonucleotides: oligo #4 upper strand (lanes 2 and 15), #4 lower strand (lanes 3 and 16), #TIE upper strand (lanes 4 and 17), #TIE lower strand (lanes 5 and 18), #4 double strand (lanes 6 and 19), #TIE double strand (lanes 7 and 20), #4 upper strand mutant (lanes 8 and 21), #4 lower strand mutant (lanes 9 and 22), #H upper strand (lanes 10 and 23), #H lower strand (lanes 11 and 24), #H double strand (lanes 12 and 25). Lanes 13 and 26: the probe used is #4 double strand (DS). f is for free DNA.
To better characterise this activity I decide to perform a southwestern experiment using as probes the upper and lower strand of oligo #4 (Fig. 4.31). In the presence of the upper strand, two proteins of about 45 and 30 KDa are detected (Fig. 4.31, panel A), specifically competed by a 30-fold excess of unlabelled oligonucleotide (Fig. 4.31, panel B). The labelled lower strand detected only the 30 KDa factor. The double strand form did not detect any protein.

Fig. 4.31: Southwestern analysis of proteins binding oligo #4
After separation on a 10% SDS-PAGE the proteins of fraction W are blotted on a nitrocellulose filter and renatured (see methods). The binding is performed with the upper, lower or double strand (DS) form of oligo #4 (panels A) in the presence of 30-fold excess of the same unlabelled oligonucleotide (panels B).

The competition by either strand could be explained either by the formation of double strand so that the labelled strand is sequestered in the double strand form, or by the real binding of the factor to the opposite strand. It seems that both the upper and lower strand bind the 30 KDa protein, while only the upper strand recognise the 45 KDa factor. The competition delimits the binding to the area comprised between footprint H and I. The reason why binding is not detected in
footprinting experiments could be due to the instability of single strand region when DNA is in its linear form.
5. Discussion

This thesis has focused on the mechanisms that regulate the expression of the urokinase type plasminogen activator (uPA) and in particular about the negative mechanisms involved. The evidences accumulated in the past years about the existence of a negative control of uPA expression (Grimaldi et al., 1986, Hofstetter et al., 1987) led me to try to dissect and characterise the elements involved in this process. A previous analysis of the huPA promoter activity in two cell lines expressing uPA (HT1080 and A-1251; Verde et al., 1988) already indicated the existence of elements involved both in the activation and in the negative modulation of uPA expression. I decided to extend this analysis to cell lines that do not express uPA in order to identify cell-specific cis-acting regulatory elements. For this purpose three cell lines, two not producing (CV1 and HeLa) and one producing high constitutive levels of uPA (PC-3) have been used to analyse constructs in which the transcription of the reporter gene CAT is under the control either of the uPA minimal promoter (MP), in the presence or in the absence of the enhancer (EMPCAT and MPCAT respectively) or the full length promoter (2212uPACAT). The result shown in Table 4.1 confirms that the sequences between the enhancer and the MP contain cell-specific modulatory element.

The 5' -nested deletions analysis and the transient transfections with constructs carrying internal deletions of the uPA 2212 long regulatory region identified at least three regions playing different roles in the silencing of uPA gene in different cell lines (Fig. 4.1 and Fig. 4.2), named S1, S2 and S3. The presence of region S1, located between -1870 and -1428, was already described by Verde et al. (1988) that mapped the negative regulatory element between -1824 and -1572 bp from the start site of transcription. This region comprises the two NF-kB elements
(−1580 and −1865) that mediate the induction of uPA by phorbol esters and TNF in HepG2 cells and HT 1080 (Hansen et al., 1992) and by TPA in HeLa cells (−1865; Novak et al., 1991). Furthermore, it has been described that a repressor acting through the NF-κB site at −1865 bp could counteract the induction (Novak et al., 1991). In the 5' nested deletions analysis, as well as in the internal deletions analysis, this element behaved differently, being active in CV1 and HeLa cells but not in PC-3 (Fig. 4.2). However the activity of this negative regulatory element is not restricted to cell lines that do not produce uPA, as it is shown by the analysis of deletion in HT1080 and A1251 (Verde et al., 1988). From this we can conclude that S1 does not act in a cell-type specific manner (Fig. 5.1).

S3 region, that has never been described before, also lacks specificity being active in all cell lines tested.

The presence of region S2 was hypothesised in the work by Cannio et al. (1991), that mapped the region between −660/−537. Furthermore the author claimed that the activity of the silencer was dependent on the presence of the enhancer and he mapped this activity between −537/−301. However, as already mentioned, the constructs used in this analysis contained several cloning mistakes, that challenged the results. The subject of this thesis allowed me to characterise the region responsible of a negative modulation of uPA gene transcription, mapping it to a wider region of 250 bp, between −787/−537. The cell type specificity is confirmed as S2 is active in cell lines that do not express uPA (see Fig. 5.1).

The cell-type specificity of this element has also been confirmed by transient transfection analysis in HepG2 cell line, where uPA is expressed at a low basal level, but in which its expression can be induced upon treatment with phorbol esters. In this case S2, that in HeLa and CV-1 cells works as a specific negative
Discussion

regulatory element, here may act as a positive element (Fig. 6.7, appendix 1, and Fig. 5.1).

![Diagram of regulatory elements and cell line activity](image)

**Fig. 5.1: huPA promoter contains multiple silencing activities**

S1, S2 and S3 indicate the different negative regulatory regions identified in the huPA promoter. The activity of the silencers is indicated by black (+). In parenthesis is indicated the fold of repression evaluated as CAT activity of the full length construct 2212 CAT over the value of the deleted constructs. In red (+) is the level of uPA produced in each cell line (indicated in blue). N.D. is for not tested.

As already mentioned, the activity of constructs with the enhancer cloned downstream the CAT gene would exclude that the increased CAT expression was due to a position effect of the enhancer closer and closer to the transcriptional start site.

The other effect described in the paper of Cannio et al. was the enhancer dependence of the silencer activity. Since the effect of region S2 on transcription was more evident when studied as internal deletion (Fig. 4.2) than in the 5' nested deletions analysis (Fig. 4.1), I wanted to test whether S2 was working only in the presence of the enhancer, i.e. S2 is able to repress the enhancer activity. However the experiment shown in Fig. 4.3 seems to exclude this possibility as S2 is able to
suppress the transcription in the presence or absence of the enhancer by the same level, as well as S1 and S3. The difference can be explained considering that the enhancer is not totally silent in CV1 cells, in transient transfection assay, (the -2212 construct activity is 2.5 fold higher than -1870 construct, Fig. 4.1), such that in the context of the full length construct (Fig. 4.2) the absolute level of CAT is higher, giving the possibility to detect little differences in CAT expression that are not detected, at least in a significative manner, in the absence of the enhancer. We can then conclude that S2 acts directly on minimal promoter, probably influencing the assembly of a competent pre-initiation complex. This hypothesis is sustained by studies carried out in my laboratory by Ibanez and Crippa on the chromatin structure of uPA 5’ regulatory region (Ibanez, 1997). According to them the DNaseI and Micrococcal nuclease I (MNaseI) analysis indicated that the chromatin structure of uPA promoter depends on the transcriptional state of the gene. In cells not producing uPA (as HeLa) almost all the 5’ flanking region results inaccessible to nuclease digestion indicating the presence of DNA tightly packed in a higher order nucleosomal structure. In PC-3, on the other hand, the regulatory region is more accessible to DNaseI and MNaseI, indicating a general loosening of the chromatin structure. In particular hypersensitive sites have been mapped within the enhancer (HS4), the silencer S2 region (H6) and the MP (HS7). In HeLa HS6 and HS7 are not present, suggesting that the MP is in a compacted chromatin structure inaccessible to the basal transcription machinery (Fig. 5.2).
Since S2 showed a cell-type specific activity I wanted to further characterise it by testing whether the negative regulatory element could act as silencer. The analysis of the effect of S2 on the SV40 heterologous promoter activity led me to conclude that S2 was effectively a silencer as it works in an orientation and position-independent manner (Fig. 4.4).

Due to the high conservation of uPA regulatory elements among different species (see human, mouse and pig promoters; Rørth et al., 1990, Cassady et al., 1991) I searched for the presence of S2 in the mouse and porcine 5'-flanking region, to confirm an important role for S2 in the regulation of the expression of uPA in different species. The finding of highly homologous sequences between mouse, pig and human uPA promoters indicated that S2 can be considered as constituted by 6 modules, highly conserved between the three species, and separated by sequences that diverge (Fig. 4.5). The genomic organisation of these modules is more similar between pig and man with the difference of module 6 that in man precedes module 5 while in pig, as well as in mouse uPA promoter, is located downstream it. Furthermore, in mouse, the modules span a wider region than in man and pig (~400 bp vs 250 bp of huPA S2 and 285 bp of pig uPA promoter; Fig. 5.3).
The DNaseI protection data gave an important clue in the understanding how S2 could work. The finding that almost the entire 250 bp long region is protected by HeLa nuclear extract (Fig. 4.7 and 4.8) might explain why the deletion of any single module does not impair completely the silencing function of S2 (Fig. 4.6). This suggests a redundancy of elements present in this silencer, that may have evolved by a series of duplications of smaller modules. This hypothesis is sustained by the presence of direct and inverted repeats along the whole S2 (Fig. 4.8) and here summarised in Fig. 5.4. Some of these repeated sequences appeared to be always protected in any or almost any location they are present: CCATC in footprint N and G; TGTTGGTGAT in footprint I and N; AAGCCC footprint B and E and at the border of footprint G; CTTTGTTCC sequence is in footprint C and H; CAATCCT at positions -730 and -700 at the border of footprint D and E, and are at least present once within the regions homologous to the murine and porcine uPA promoter (Fig. 5.4).

Of these, only region 3 appears not to be protected at all and furthermore the naked DNA is cut by DNase I at a few position in the lower strand and in the upper strand (Fig. 4.7).
The presence of stretches of A and T residues suggests that the helical conformation of this part of the silencer is different from that of the rest of DNA of S2, although the electrophoretic mobility experiments of S2 circular permuted fragments excluded that this modification results in intrinsic curved DNA (Fig. 4.19). It has been demonstrated that this kind of sequence array is not efficiently cut by DNaseI (Drew and Travers, 1985; Drew, 1991), but it shows the peculiarity to be bendable by external forces like interaction with the histone octamer or with proteins that induce a curvature in the DNA. The high conservation of this module between mouse, pig and human, then, would argue for an important role of the local DNA conformation for S2 activity, suggesting that the entire conformation of S2, rather than one particular sequence, is important for S2 silencing activity, a conformation that can be induced by the binding of factors that stabilise this structure.

The comparison of the complexes detected by different oligonucleotides, some of which overlapping with each other or containing the same sequence (Fig. 5.5), and the analysis of the competition also argue for the hypothesis of
redundant elements in S2. Table 5.1 summarises the results of EMSA and competition analysis from experiments showed in Fig. 4.12, 4.13, 4.14 and 4.21.

![Diagram of oligonucleotides](image)

**Fig. 5.5:** Schematic representation of the oligonucleotides used in the EMSA analysis of the nucleoproteic complexes detected by DNase I footprinting.

<table>
<thead>
<tr>
<th>prob</th>
<th>#2</th>
<th>#3</th>
<th>#N</th>
<th>#TIE</th>
<th>#4</th>
</tr>
</thead>
<tbody>
<tr>
<td>comp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c1</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>c2</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>c3</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>c4</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>c5</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>c6</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>c7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>c8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

**Table 5.1:** Cross-competition analysis.
The arrows indicate a down-shift in the mobility of the complex. (+) indicates competition; (-) no competition; (+/-) partial competition; (++) increased intensity. c = complex; comp = competitor.

Each oligonucleotide (oligo #3, #4, #TIE and #N) detects in HeLa nuclear extract many complexes, some of which are specific, as shown by homologous competition. Some of these oligonucleotides are able to compete the binding of one or more complexes to each other. The analysis of the sequences (table 5.2) showed that oligo #3, #4, #TIE, #N and #H share the sequence CCATC that in
oligo #TIE and #4 is found in the lower strand with a mismatch (CCAAAC instead of CCATC). Oligo #4 contains a second one in the upper strand.

<table>
<thead>
<tr>
<th>Oligo</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>#N</td>
<td>5' CTCTGATAGCA CCATCAAACAAACCCCCT 3'</td>
</tr>
<tr>
<td>#TIE</td>
<td>5' ACTGGTTGT GTTGG TGATATCTGGGG 3'</td>
</tr>
<tr>
<td>#4</td>
<td>5' CTTTGTCCATCCACTGGTTGT GTTGG TGAT 3'</td>
</tr>
<tr>
<td>#H</td>
<td>5' GGCTTTGT CCATCCACTGGTTGT 3'</td>
</tr>
<tr>
<td>#3</td>
<td>5' AGCA CCATCAAACAAACCCCCTTACTGCCGAAAGCAATAAGCCC 3'</td>
</tr>
</tbody>
</table>

Table 5.2: Sequence comparison of the oligonucleotides used in this study

The CCATC sequence is shown in bold in either orientation. This sequence is also present at position -764/-760.

When oligo #TIE is used as competitor, it is able to compete complex 4 detected by #N and complex 3 detected by #4 (3b). In the latter case the competition gave rise to a faster mobility of complex 3. This could be explained by the titration of one of the two putative factors bound to their cognate sequence in oligo #4, or by the binding of some other factors that is mutually exclusive with that bound by the CCATC sequence. Competition by #4 of complex 3 generated by #TIE (3a), that contains only one CCATC motif, would argue for the first hypothesis as this has just one binding site for the CCATC sequence. The increased mobility of complex 3 detected by #TIE (3a) and complex 4 detected by #N in the presence of oligo #H as competitor can be explained only hypothesising that other factors are bound to the oligonucleotides (Fig. 5.6). In particular, oligo #N contains the sequence TTTGTT, in the lower strand, nearby the CCATC motif. It can be hypothesised that the binding to this element (that is the consensus for the HMG-box containing protein) is mutually exclusive with respect to the binding to the CCATC sequence, while for #TIE the sequence that is partially overlapped can be the binding site for
another factor. For #4 again this can be due to the titration of one complex and/or binding of other factors.

![DNA binding protein model](image)

**Fig. 5.6**: Model of the putative DNA binding protein interacting with the region covered by oligonucleotides #3, #N, #4, #TIE and #H.

The blue sphere would recognize the CCATC sequence; the grey one would bind the sequence TTTGTT and the pink one would bind the region in common between oligo #TIE and #H.

Most of the protected areas with HeLa nuclear extract are the same with PC-3 nuclear extract, with the exception of footprints L, M and N (Fig. 4.7). The band-shift analysis with oligonucleotides covering most of S2 region, seems to indicate that the nature of the complexes may indeed differ. Oligo #1 (-726/-705) detects 2 complexes in HeLa and PC-3 but they show different mobility (Fig. 4.10, complex 1 vs 1*). Complex 2 is not present in PC-3 while complex 3 is lacking in HeLa nuclear extract.

Furthermore, when the TIE element, protected by PC-3 and HeLa nuclear extracts in DNase I footprinting, is studied by EMSA analysis, different complexes are detected in the two cases: complex 1 and 2 are common for both nuclear extract, complex 3 is specific for HeLa, while complex 4 is present only in PC-3 (Fig. 4.24). It cannot be excluded that complex 4 is formed through protein-protein interaction that do not alter the binding of other factors as detected by DNase I footprint. Complex 3, on the other hand, that behaves as complex 3 detected by #4
(Fig. 4.21, panel B), could be formed by the factors binding the CCAAC sequence present in the lower strand of #TIE.

The binding to the cognate site for the proteins of the HMG box family in the presence either of HeLa or PC-3 nuclear extracts, as detected by DNase I footprinting (Fig. 4.7) and EMSA analysis (Fig. 4.16), and the evidence that factors belonging to this family can indeed bind the element present in uPA promoter (Fig. 4.18 and 4.20) do not exclude that the binding of these factors, in vivo, occurs only in PC-3.

The faster mobility of the oligo encompassing the HMG+TIE elements compared to the #TIE alone (they show the same pattern of complexes) might be explained by the binding of a factor to the HMG element that, upon binding, bends DNA, giving a structure that runs faster in a polyacrylamide gel (Fig. 4.26). The architectural role evidenced for the HMG box proteins, allowing the assembly of a competent complex able to transactivate transcription, supports this hypothesis. The missing binding by TCF-1 HMG-box, in the experiments of Fig. 4.17 and 4.18, might be explained by the different affinity of this protein for the sequence present in huPA promoter (TCF-1 binds preferentially 5' CCTTTGAAGCT 3' sequence than 5' GGTTCGTCC 3') although I cannot exclude that a problem of degradation of the protein occurred.

**A single-strand DNA binding protein in the huPA promoter**

The EMSA analysis of the complexes formed by HeLa nuclear extract, in the presence of single strand DNA probes, has led to the interesting finding that a protein with a higher affinity for single strand DNA rather than double strand form of S2 region could be present in this cell line. Recently, several papers described the existence of sequence-specific single strand DNA binding activities in different promoters involved in either positive or negative transcriptional regulation. Single strand DNA have been described in c-myc, β-globin, PDGF A,

Table 5.3 summarises the results of the EMSA analysis, in the presence of HeLa nuclear extract, using either the upper or lower strand of #4 as probes (Fig. 4.30).

Both the upper and the lower strand detect more than one complex; however the southwestern analysis (Fig. 4.31) has revealed that the upper strand binds two polypeptides of 40 and 30 KDa, with a greater affinity for the former, while the lower strand binds only the 30 KDa protein. The presence of multiple complexes can be due to the formation of multimers and/or of higher order nucleoproteic complex.

Competition of either strand with their complementary strands (as full length or as part of the entire oligonucleotide sequence) abolishes the binding of all complexes probably due to the formation of double strand form of DNA rather than to a sequence specific competition (Fig. 4.30, lanes 3 and 15). Competition of either strand with an excess of unlabelled homologous oligonucleotide, on the other hand, has showed that neither the 5' part or the 3' part of the oligonucleotides are able to compete any complex, suggesting that the binding requires most of the sequence present in oligo #4 (compare competition by #TIE and #H; Fig. 5.7).
Table 5.3: Cross-competition analysis of single strand probes
The arrows indicate a down-shift in the mobility of the complex. (+) indicates competition; (-) no competition; (+/-) partial competition; (++) increased intensity. c = complex; compet. = competitor.

<table>
<thead>
<tr>
<th>probe</th>
<th>#4 upper</th>
<th>#4 lower</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>c2</td>
<td>c3</td>
</tr>
<tr>
<td>#4 upper</td>
<td>+ + +</td>
<td>+ + +</td>
</tr>
<tr>
<td>#4 lower</td>
<td>+ + +</td>
<td>+ + +</td>
</tr>
<tr>
<td>#TIE upper</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>#TIE lower</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>#4 upp mut</td>
<td>± ± +</td>
<td>± ± +</td>
</tr>
<tr>
<td>#4 low mut</td>
<td>+ + +</td>
<td>+ + +</td>
</tr>
<tr>
<td>#H upper</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>#H lower</td>
<td>+ ± +</td>
<td>+ ± +</td>
</tr>
<tr>
<td>#4 DS</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>#TIE DS</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>#H DS</td>
<td>- - -</td>
<td>- - -</td>
</tr>
</tbody>
</table>

5' CTGTTGCCATCCACTGGTTGTTGAT 3' #4 upper
5' ACTGGTTGTTGTTGATATCTGGGA 3' #TIE upper
5' CTGTTGCCATCCACTGGTTacTGGGAT 3' #4 upp mut
5' GCTGTTGCCATCCACTGGGTG 3' #H upper
5' ATCACAACACACACAGTGATGGGAACAAG 3' #4 lower
5' CCCAGATATCAAAACCAGCAGGATGGGAACAAAG 3' #TIE lower
5' ATCACAATATACACAGTGGATGGAACAAAG 3' #4 low mut
5'CAAACAGTGATGGGAAGGAAGCC 3' #H lower

Fig. 5.7: Sequences of the oligonucleotides used in the EMSA analysis of single strand DNA binding proteins
In particular the 3' part of the oligonucleotide is important for the binding as a mutation of this region (#4 mut), either in the upper or lower strand, partially impaired the competition by oligo #4. Furthermore, the upper strand of oligonucleotide #N is able to compete the binding of complex 2 detected by the upper strand of oligonucleotide #4 (data not shown). Oligo #N contains the sequence CCATC, suggesting the involvement of this element in the recognition of DNA by the 40 KDa single strand DNA binding protein. Since this sequence is present also in the -764/-760 region it is not excluded that more than one region in the S2 silencer can promote the formation of single strand region as shown in the model of fig. 5.8.

The binding of the 40 and 30 KDa factors can promote or simply stabilise single strand regions that would prevent the binding of transacting factors to double strand DNA Fig. 5.8). The binding of these factors to single strand DNA can result in a local torsional stress that can alter the topology of this region of the promoter promoting the formation of higher order nucleoprotein complexes resulting in a compact chromatin structure. The faint protection in the corresponding TIE element in DNase I footprint (Fig. 4.7) could be due to the formation of single strand regions that would become more evident when a supercoiled template is used instead of a linear fragment of DNA.

This model is supported by the recent finding of a family of DNA and RNA binding proteins, the Y box factors, with the protein YB-1 as a member. It has been shown that YB-1 repress interferon-γ-induced transcription of class II human major histocompatibility (MHC) genes (Ting et al., 1994) by binding its cognate sequence in the DRA promoter as single strand templates.

It has been hypothesised that YB-1 can repress transcription by inducing or stabilising single strand regions in the DRA promoter preventing loading and/or function of other DRA-specific transactivating factors.
Fig. 5.8: Proposed model for the binding of single strand DNA binding protein in S2. The blue and green spheres (plain colour) indicate the high affinity binding sites for the 30 and 40 KDa proteins, binding oligo #4, detected in southwestern experiment. The hatched blue and green spheres indicate the low affinity binding sites. The binding of these proteins would prevent the binding of other factors, like the HMG-box containing proteins, important for the activation of transcription in cell lines producing uPA (see text).

The biochemical characterisation of this protein showed that its MW is between 33-34 KDa. Furthermore it has been detected in HeLa nuclear extract, by an anti-YB-1 antisera, a band of approximately 45 KDa consistent with the reported molecular mass for another protein belonging to this family, named dbpB, of 42 KDa (Spitkovsky et al., 1992). The sequence bound by these factors resembles the CCATC sequence of uPA promoter being CAATC. Furthermore, three putative binding sites for YB-1 are present in uPA S2 at position -776/-771, -731/-726, and -698/-693 (Fig. 5.9).

Fig. 5.9: YB-1 and CCATC binding sites in S2
This might suggest that one or more proteins belonging to this family could indeed bind the S2 silencer of huPA promoter.

**The TIE element is involved in S2 silencing activity**

Several line of evidence indicate that the TIE behaves as a functional silencer module. When deleted from the context of the full length promoter the transcriptional activity increases by about 1.5-2 fold (Fig. 4.6). When one copy of TIE is cloned in front of the SV40 promoter it is able to repress SV40 driven transcription by about 1.5 fold (Fig. 4.28, pTIECAT). The increasing copy number represses the transcription up to 3 fold (p2xTIE and p4xTIE) and is independent of the orientation of the single elements. The analysis of four different mutants in TIE sequence (mut 1, mut 2, mut 3 and mut 4) by EMSA (Fig. 4.22b and 4.23), showed that only complex 1 and 2 were affected in mut 1, 2 and 3, while mutant 4 showed an increased affinity for the three complexes. The *in vivo* analysis of these mutants, in the context of the full length promoter and E787CAT constructs (Fig. 4.27), gave an unexpected result as none of the mutations affect the silencing activity of S2 apart from mutant 3 in E787CAT context. Indeed the repressive capability seems to be increased. If this can explain the behaviour of mutant 4, that shows higher affinity for factors present in HeLa nuclear extract, concerning the other mutants the result can be explained by considering complex 3: none of the mutants in fact has lost the ability to compete for the formation of complex 3 formed by the binding to the CCATC sequence, although two of these mutations (mut 1 and mut 2) falls in the CCAAC sequence present in the lower strand of #TIE (Table 5.6). Mutant 1 gives rise to the sequence recognised by YB-1 protein, while mutant 2 restores the T residue at position 4 but changes position 2. The absence of competition by the 3 complexes with #TIE rev and #TIE dir could be due to the lack of CCAAC sequence in TIE rev, and to the absence of flanking
sequences important for the binding of nuclear factors to the CCAAC sequence in the case of TIE dir (see table 4.4).

<table>
<thead>
<tr>
<th>Residues Position</th>
<th>Sequence</th>
<th>Mutant Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>5' CCAACA 3'</td>
<td>#TIE lower strand</td>
<td></td>
</tr>
<tr>
<td>5' CCAATA 3'</td>
<td>#TIE mut1 lower strand</td>
<td></td>
</tr>
<tr>
<td>5' CGATCA 3'</td>
<td>#TIE mut2 lower strand</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.6: Mutants in TIE element

Furthermore, in the context of the whole S2, the other CCATC motifs are present, so it cannot be excluded that in order to see an effect all the sequences should be mutated. Then it would have been important to test these mutations in the context of the TIE element controlling SV40 driven transcription.

A hierarchy of protein-DNA interaction in uPA promoter S2 region

A more detailed analysis of the S2 region (Fig. 4.11) revealed the presence of putative binding-sites for factors belonging to the GATA family of transcription factor and for the yeast ABF-1 factor. The interesting feature of these sites is their organisation, as they resemble that found in the ε-globin silencer (Peters et al., 1993), where the binding of the YY1 factors prevents the binding of the erythroid specific GATA-1 factor (Fig. 5.10).

In the case of the S2 region, the putative binding sites for ABF-1 and GATA-1 overlap or are in the vicinity of the CCATC sequence (see Fig. 11c and Fig. 5.10). An oligonucleotide with the consensus for ABF-1 is however not able to compete the complexes formed by oligonucleotide #1 that encompasses the most 5' ABF-1 site (Fig. 4.11a), but it cannot be excluded that other oligonucleotides with a sequence
derived by the consensus but carrying different residues, can have different affinities for these factors.

![Diagram of the e-globin silencer and huPA S2]

**Fig. 5.10: The e-globin silencer and huPA S2.**
Panel A: The organisation of the e-globin silencer. The black bars indicate the DNaseI footprint of the sense and antisense strand (From Peters et al. - 1993 - JBC: 268, 3430-3437). Panel B: The huPA S2 silencer. The putative binding sites for GATA-1 and ABF-1 are indicated. In green the CCATC or CCAAC motif, and in orange the binding sites for a putative HMG-box containing proteins.

However, the presence of such overlapping putative binding sites for transcription factor might be important for the competition between the formation of an activating or inhibiting complex.

Furthermore the CAATC sequence is the binding site for NF-Y, C/EBP and CTF/NF1 transcription factors. Recently it has been shown that NF-Y binding is important for transactivation of the invariant chain promoter, by promoting the assembly of a stereo specific nucleoprotein complex involving 250 bp of DNA sequence. The presence of the histone fold-like motifs in this factor led to the speculation that a static nucleosome like particle is formed carrying in proximity enhancer and proximal promoter sequences (Linhoff et al., 1997).

If we hypothesise a role for topological constraints in the formation of a correct DNA conformation in the S2 region, it would be interesting to study the effect of the silencer in the context of the genomic DNA, by stable transfection. Right now I can only speculate that S2 might work by freezing particular DNA
conformations that include the stabilisation of single stranded regions by the binding of specific single-strand DNA binding proteins, preventing the binding of transcriptional activators (Fig. 5.11, HeLa). The inaccessibility of DNaseI, as revealed by experiments of chromatin mapping in HeLa and PC-3 cells, might be due either to regions that are rarely cut by DNaseI (poly dA-dT tracts) or by the formation of a higher order structure in which nucleosomes and other non-nucleosomal DNA binding proteins bind and pack the DNA in a close conformation (Fig. 5.11, HeLa).

---

**Fig. 7.11: Model for the assembly of negative or positive nucleoprotein complex at S2**

In **HeLa** cells the binding of single strand DNA binding protein prevent the loading of transactivators to S2 and allows the packing the DNA in a compact nucleosomal structure. Furthermore, in synergism with other regions (i.e. S3) and through the binding of repressor or corepressor molecules, the assembly of a competent pre initiation complex is blocked, thus preventing the transcription of the gene. In **PC-3** cells, the binding of architectural proteins (i.e. HMG-box containing protein) would result in disruption of chromatin driving the assembly of a competent pre-initiation complex and the interaction of transactivators with the basal machinery through looping of intervening sequences.
On the other hand in PC-3 cells the hypersensitivity detected in correspondence of the region of silencer S2 might be interpreted as due to a different nucleoprotein complex in which, for example, HMG family proteins and nucleosomes bind, allowing, through DNA looping, the interaction of sequences located upstream the S2 and the basal promoter elements (Fig. 5.11).
6. Appendix 1: Negative regulation of uPA by TGF-β1

Introduction

Transforming growth factor β1 (TGF-β1) is a hormonally active polypeptide belonging to a large super family of secreted factors that comprises the three TGF-β isoforms (TGF-β1, -β2 and -β3, each encoded by a separate gene), the activins, the bone morphogenetic proteins and many other secreted factors that are all thought to play major roles in differentiation and tissue morphogenesis (Derynck et al., 1988, Massague', 1990). The high degree of evolutionary conservation of the closely related TGF-βs from Drosophila and Caenorhabditis elegans to humans and their widespread expression in embryonic as well as adult tissues suggests an important role for these factors.

A variety of biological activities that affect cell proliferation and differentiation in many cell types have been ascribed to these factors. They stimulate the proliferation of various mesenchymal cell types, but act as a growth inhibitor of many other cell types, including epithelial cells. Furthermore, when present at high concentration, they induce the synthesis of extracellular-matrix proteins (Penttinen et al., 1988), modulate the expression of matrix proteases and protease inhibitors (Lund et al., 1987, Gerwin et al., 1990, Lund et al., 1991), increase integrin expression and thus enhance cell adhesion (Ignotz et al., 1989). TGF-βs also affect mesenchymal differentiation and at low concentration are very potent chemotactic agent for several cell types, especially monocytes and fibroblasts (reviewed in Roberts and Sporn, 1990; Postlethwaite et al., 1987). Several are the observations that in vivo TGF-βs action is largely growth inhibitory (Silberstein and Daniel, 1987) as shown for the early phase of liver
regeneration following partial hepatectomy (Russell et al., 1988), lymphoid and myeloid cell proliferation (Goey et al., 1989, Fontana et al., 1989).

The signalling mechanisms leading to changes in proliferation by TGF-β are likely to be different in cells that are growth stimulated versus those that are growth arrested by TGF-β. Recently the SMAD family of signal transducer proteins has been identified as key components in either activation or inhibition of TGF-βs signal transduction pathways (reviewed in Heldin et al., 1997). Following phosphorylation and activation by receptor-kinases, SMAD complexes translocate into the nucleus where, either directly or in complex with other proteins, affect transcription of specific genes in response to TGF-β.

Furthermore, ligand stimulation of the receptors increases the expression of the inhibitory SMADs, that may have a negative-feedback role in signal transduction.

Most studies of TGF-β induced growth inhibition have focused on its effect on the cell-cycle machinery. Several lines of evidences show that TGF-β1 can inhibit DNA synthesis occurring 12 hr following stimulation of quiescent cells when added at any point prior to the G1/S boundary (Howe et al., 1991, Laiho et al., 1990). One target candidate of TGF-β1, even in late G1, could be the c-myc gene the expression of which has been shown to be necessary for keratinocytes proliferation (Pietenpol et al., 1990b). c-myc gene is induced rapidly upon induction with EGF in many cell types and its expression remains elevated throughout G1 and during early S phase (Coffey et al., 1988). TGF-β1 rapidly reduces c-myc mRNA and protein and cycloheximide treatment has shown that protein synthesis is required for this effect. The block in c-myc expression occurs at the level of transcription initiation (Pietenpol et al., 1990b). Furthermore it has been demonstrated that DNA-tumour virus transformed cells become resistant to the growth-inhibitory effect of TGF-β1 (Pietenpol et al., 1990a) and in particular c-
myc mRNA expression is no longer suppressed. These DNA-tumour virus oncoproteins bind the product of the retinoblastoma gene (pRB) blocking its growth-suppressive activity, and this leads to an aberrant growth control (Munger et al., 1989). It appears then that TGF-β1-induced growth inhibition is a consequence of pRB-mediated suppression (direct or indirect) of c-myc transcription.

TGF-β1 would act maintaining pRB in its underphosphorylated form, necessary for the blocking of cell cycle in G0/G1 phase (Laiho et al., 1990). It has been proposed that the effect on pRB phosphorylation is indirect, involving a direct inhibition of the synthesis of cyclins and associated cyclin-dependent kinases and induction of inhibitors of cyclin-CDK complexes (Geng and Weinberg, 1993, Ewen et al., 1993, Derynck, 1994).

As already mentioned, the block in c-myc expression by TGF-β1 occurs at the level of transcription initiation. A cis-acting regulatory element responsible for the TGF-β1 suppression has been identified (position -100/+71 relative to the P1 transcription start site), suggesting that the synthesis or modification of a protein that binds to this cis-acting element could be the target of the mechanism through which TGF-β1 inhibits cell proliferation. Furthermore, Matrisian and collaborators described a cis-acting element in the promoter of the rat transin/stromelysin gene that mediates the TGF-β1 suppression of its expression (Kerr et al., 1990). This element, termed Transforming growth factor Inhibitory Element (TIE), mediates the repression of EGF-induced expression of the gene (Kerr et al., 1988). The inhibition occurs at transcriptional level through binding of a c-fos containing complex to this element.

The same element is present twice in the human c-myc gene (but one is outside the 5’ region shown to be necessary for TGF-β1 suppression of c-myc transcription while the other is also required for regulation by pRB; Pietenpol et
Appendix 1: Negative regulation of uPA by TGF-β1

as well as in the promoters of other genes known to be suppressed by TGF-β1 as human collagenase (Angel et al., 1987, Edwards et al., 1987), mouse elastase (Stevenson et al., 1986), MRP/proliferin (Connor et al., 1989) and human urokinase (see Table 6.1).

<table>
<thead>
<tr>
<th>Promoter</th>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRANSIN</td>
<td>rat</td>
<td>-709 GAGTTGGTGA</td>
</tr>
<tr>
<td>ELASTASE</td>
<td>mouse</td>
<td>-572 GAGTTGGTGA</td>
</tr>
<tr>
<td>COLLAGENASE</td>
<td>human</td>
<td>-246 GAAATTGGAGA</td>
</tr>
<tr>
<td>MRP/PROLIFERIN</td>
<td>mouse</td>
<td>-1004 GCCTTGGTGT</td>
</tr>
<tr>
<td>c-myc</td>
<td>human</td>
<td>-87 GCCTTGGGCG</td>
</tr>
</tbody>
</table>

Table 6.1: Comparison of promoter region from TGF-β1 inhibited genes

Nucleotides position is relative to the major start site of transcription (P2 in the case of c-myc). N = any nucleotide; capital letters = invariant nucleotides; small letters = preferred nucleotide (From Kerr et al. - 1990 - Cell 61; 267-278)

The TIE element present in the human urokinase promoter lies in the region of silencer S2 (-553/-537) as a partially overlapped and inverted repeat of the TIE consensus (Fig. 6.1).

TGF-β1 induces PAI-1 transcription and modulates the transcription of uPA in a manner that varies with the cell type. The availability of cell lines synthesising different levels of uPA provide a good experimental system to test the effect of TGF-β1 on uPA gene expression. In particular, in order to address whether TIE plays a role in the regulation of uPA gene expression two cell lines, one expressing high constitutive level of uPA (PC-3) and one expressing a low
basal level of uPA but inducible by phorbol-ester treatment (HepG2), have been chosen.

**Fig 6.1: The TIE element in the human urokinase promoter**

Upper: mapping of TIE in the human urokinase promoter. The arrows indicate the partial overlapped and inverted organisation of the element. Lower: comparison of the urokinase sequence with that from transin promoter and consensus. In black outlined is the left and inverted TIE element; in red is the right and direct repeat. The red outlined character represent the overlapped region between the left and the right part of the element.

**Materials and Methods**

**DNA manipulations**

The BamHI- Bgl II fragment from human uPA cDNA (Riccio et al., 1985) and the BamHI fragment from human uPA receptor (uPAR) cDNA (Roldan et al., 1990) were cloned in the BamHI-Bgl II sites and in the BamHI site of pSP73 plasmid (Promega) to give respectively pT7-uPA and pSP6-uPAR constructs (Table 6.2). DNA was purified by caesium chloride gradient and linearized with the appropriate restriction enzyme. After digestion DNA was phenol extracted once and ethanol precipitated and resuspended in an appropriate volume of distilled water.
All the constructs used for transient transfections are described in Materials and Methods (previous section).

**Cell culture, RNA extraction and Northern blot analysis**

HepG2 and PC-3 cells were grown in Dulbecco's modified Eagle medium (DMEM) containing 10% foetal calf serum (FCS) to a confluent state. Cells were serum starved overnight in DMEM containing 0.5% FCS prior to addition of TGF-β1 (Boehringer Mannheim; 5 ng/ml) and PMA (100 ng/ml). RNA was extracted at different times as described in the figures, according to the method described by Chomczynski and Sacchi, 1987. RNA (PC-3: 10 μg per lane, HepG2: 20 μg per lane) was sized fractionated on a 1.2% agarose gel containing 6.6% formaldehyde. The RNA was then transferred to nylon membrane (Costar) by electrotransfer in the following buffer (1x TAE: 10 mM Tris-HCl, 5 mM sodium acetate, 0.5 mM EDTA, pH 7.8) for 4 hours at 40 volts and cross-linked to filter with a Stratagene cross-
linker. Radiolabelled probes were obtained by in vitro transcription of Smal linearized pT7-uPA and of pSP6-uPAR using T7 and SP6 polymerases respectively (Promega). The Pst I fragment of mouse glyceraldehyde-3-phosphate dehydrogenase (GAPDH) cDNA and the Bgl II/SacI fragment from human PAI-1 cDNA (Ginsburg et al., 1986) were labelled by random priming (Rediprime, Amersham) and used for normalisation of sample RNA loading. Filters were then pre-hybridised in mixture containing 50% formamide, 6x SSPE, 5x Denhardt's, 0.5% SDS, 100 mg/ml DNA carrier (MB DNA, Boehringer Mannheim) at 42 °C for 2 hours. Hybridisation with riboprobe was performed in the presence of 10% dextran-sulphate at 42 °C for 12-16 hours. The probe in excess was removed by washing the filters in 3x SSPE/0.1% SDS for 20' at 65 °C followed by 2 washing in 0.3x SSPE/0.1% SDS at 65 °C for 30' each.

The filters were then exposed with intensifying screens at -80 °C for period from 30' (18S) up to 4 days (uPA and p53 probes). mRNA level was quantified by densitometric analysis (Molecular Dynamics, SF, ImageQuant Software).

**Transient transfection analysis**

HepG2 cells were transfected by the calcium-phosphate method. Approximately 10^6 HepG2 cells/reporter construct were seeded in a 10 cm dish the day before transfection in DMEM containing 10% FCS. 4 hours before transfection medium was changed. 15 μg target construct DNA and 1.5 μg of EF1alpha-lacZ plasmid DNA were used per each plate. After 16 hours, the precipitate was removed and the cells were trypsinized and split in three 6 cm plates in DMEM containing 10% FCS and allowed to attach for 6-8 hours. For TGF-β1 and TPA treatment, medium was changed to DMEM containing 0.5% FCS overnight. The day after TPA at 100 ng/ml and TGF-β1 at 5 ng/ml (or otherwise indicated in the legend of the figures) diluted in DMEM were added to the cells (control cells received respectively DMSO and PBS). After 8 hours cells were harvested and
processed for protein extraction (Sambrook et al., 1989). CAT amount was detected by a non radioactive CAT-ELISA kit (Boehringer Mannheim) and the amount was normalised to β-galactosidase activity.

Results

Effect of TGF-β1 on uPA expression in PC3 cells

It has been reported that the prostate derived adenocarcinoma PC3 cells secrete predominantly TGF-β2 isoform but relatively little TGF-β1 protein, although TGF-β1 mRNA level is higher than that of TGF-β2 (Bang et al., 1992, Ikeda et al., 1987). On the other hand TGF-β1 has a certain antiproliferative effect on this cell line (Reyes-Moreno et al., 1995). In order to clarify whether the uPA promoter was sensitive to the inhibitory action of TGF-β1 and hence whether the constitutive high level of uPA was due to a lack of TGF-β1 isoform activation the cells have been treated with 5 ng/ml of TGF-β1 and RNA extracted at different time points has been analysed by Northern analysis. Up to 24 hours treatment did not change the level of uPA mRNA (Fig. 6.2). However hybridisation of the same blot with a probe specific for PAI-1 revealed an increase in the level of mRNA with a peak at 4 hours after addition of TGF-β1 as described for other cell lines (Lund et al., 1987, Gerwin et al., 1990).

TGF-β1 inhibits TPA induction of uPA mRNA in HepG2 cells

TGF-β1 is able to counteract the effect of several mitogens. In order to understand whether the TIE element plays a role in the uPA promoter I tested whether TGF-β1 might be involved in the suppression of TPA induced expression of uPA gene in HepG2 cells. In this cell line uPA gene expression is under the control of the inducible enhancer located 2 Kb upstream of the start site. These cells do not express constitutively high levels of uPA but its expression is induced
by treatment with PMA (phorbol-myristate acetate) reaching a maximum expression level at 8 hours after induction as shown by a northern blot analysis of total RNA prepared at varying times after addition of PMA (Fig. 6.3; Nerlov et al., 1992). When cells were simultaneously treated with TGF-β1 and PMA a reduction in the level of uPA mRNA by about 4 fold was observed (Fig. 6.3), suggesting that TGF-β1 was able to counteract PMA effect on uPA gene expression.

Fig. 6.2: Northern blot analysis of uPA and PAI-1 expression in PC-3 cells
Total RNA was isolated by PC-3 cells treated with 5ng/ml of TGF-β1 for the times indicated. Panel A: autoradiograms of Northern blots. For PAI-1 both mRNAs (3.2 Kb and 2.3 Kb) were detected by the same probe. Panel B: quantification of uPA and PAI-1 mRNAs by densitometric scanning. The values are expressed as relative to GAPDH (glyceraldehyde-3-phosphate dehydrogenase). For PAI-1 values of both mRNAs are summarised.
Appendix 1: Negative regulation of uPA by TGF-β1

Fig. 6.3: Time course analysis of TGF-β1 effect on uPA expression in HepG2 cells
Confluent HepG2 cells were treated either with TPA (100 ng/ml) or TPA (100 ng/ml) and TGF-β1 (5 ng/ml) for the times indicated. RNA was extracted and analysed by Northern blot. Panel A: autoradiograms of hybridisation with human uPA and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) probes. Panel B: scanning densitometric analysis of the film. The values are referred as uPA/GAPDH. Panel C: the values are expressed as TPA/TPA+TGFβ1 (fold repression).

The same result was obtained when cells were pretreated with TGF-β1, 2 and 3 hours before PMA addition (fig. 6.4).

Fig. 6.4: Effect of TGF-β1 pretreatment on TPA induction of uPA mRNA
HepG2 cells were pretreated with 5 ng/ml of TGFβ1 prior to addition of TPA (100 ng/ml). RNA was extracted after 8 hours of TPA induction. Panel A: autoradiograms of uPA and GAPDH hybridisation. Panel B: scanning densitometric analysis of the film. The values are expressed as TPA/TPA+TGFβ1 (fold repression).
Appendix 1: Negative regulation of uPA by TGF-β1

As shown in fig. 6.5, the effect was already obtained at picomolar concentration of TGF-β1 (0.1 ng/ml) and did not increase with increasing amount of inhibitor added (up to 10 ng/ml).

Figure 6.5: Titration of TGFβ1 concentration on TPA-induced uPA expression
HepG2 cells were treated with different amount of TGF-β1 as indicated in the figure together with 100 ng/ml of TPA. RNA was extracted after 8 hours of treatment. Panel A: autoradiograms of uPA and GAPDH hybridisation. Panel B: scanning densitometric analysis of the film. Values are expressed as TPA/TPA+TGF-β1 (fold repression).

TGF-β1 alone did not seem to have any effect on the basal level of uPA mRNA, although I cannot exclude that it was not possible to detect any reduction due to the too low level of uPA mRNA in unstimulated cells (data not shown). Furthermore, TGF-β1 and PMA, when added simultaneously, decreased PMA-stimulated uPAR mRNA level in HepG2 cells (fig. 6.6), confirming that TGF-β plays an important role in the fine regulation of the uPA-uPAR mediated extracellular proteolysis.
Appendix 1: Negative regulation of uPA by TGF-β1

Fig. 6.6: Time course analysis of TGF-β1 effect on uPAR expression in HepG2 cells

HepG2 cells were treated with TPA and TGF-β1 as described in Fig. 6.3. Autoradiograms of hybridisation with human uPAR and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) probes and scanning densitometric analysis of the film are shown.

Transient transfection analysis of inhibition of TPA induced uPA promoter driven transcription

In order to find out the element(s) that mediate the effect of TGF-β1 on TPA induction of uPA, and in particular whether the TIE element is involved, a series of constructs, carrying deletions of the uPA 5' flanking region fused to CAT reporter gene, have been tested in transient transfection of unstimulated and TPA-stimulated, in the presence or absence of TGF-β1, HepG2 cells. As shown in fig. 6.7 the full length construct (2212uPACAT) responded to TPA with a ~ 4-fold increase in CAT level. The simultaneous presence of TGF-β1 decreased this value by about two fold. Deletion of regions S1 (-1820/-1428) and S3 (-537/-86) under unstimulated conditions increased the basal level of transcription of about 3.5 and 4.5-fold respectively over that of the full length construct, in agreement with the presence in these regions of cis-acting negative elements. These constructs were still able to respond to TPA induction although with a lower efficiency than wild-type (2212uPACAT) construct (2.9 and 2 fold respectively), but no inhibition by TGF-β1 of TPA-stimulated CAT activity is observed. On the other hand construct ΔS2, carrying the deletions of the TIE containing region, decreased the basal level
of transcription by about 3-fold, indicating the existence of a positively acting element. This construct was highly inducible by TPA (8.8-fold) and was still repressible by TGF-β1 (~2 fold).

<table>
<thead>
<tr>
<th>RELATIVE CAT ACTIVITY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Basal</td>
</tr>
<tr>
<td>Basal</td>
</tr>
<tr>
<td>ΔS1</td>
</tr>
<tr>
<td>ΔS2</td>
</tr>
<tr>
<td>ΔS3</td>
</tr>
<tr>
<td>EMP</td>
</tr>
<tr>
<td>ES1</td>
</tr>
<tr>
<td>ESC</td>
</tr>
<tr>
<td>ES3</td>
</tr>
<tr>
<td>ES2</td>
</tr>
<tr>
<td>ES3</td>
</tr>
</tbody>
</table>

Fig. 6.7: Transient transfection analysis of uPA promoter activity in HepG2 cells
Plasmid constructs carrying different portions of uPA promoter region were tested in transient transfection in HepG2 cells. The cells were treated either with 100 ng/ml of TPA (blue solid bars) or with 100 ng/ml of TPA and 5 ng/ml of TGF-β1 (green solid bars). Red bars: untreated cells. Values are normalised to β-galactosidase activity and they are the means ± standard deviation of at least 2 independent experiments.

When CAT transcription was driven by uPA enhancer and minimal promoter the basal level was 2.8 fold higher than the full length construct and it responded to TPA induction by increasing the level of CAT of ~6 fold, with no effect by TGF-β1 treatment. Transfection with ES1CAT, ES2CAT and ES3CAT constructs (carrying either region S1, S2 or S3 cloned in front of the minimal promoter in EMP based plasmid) confirmed the result obtained with the deleted constructs, with S1 and S3 repressing transcription by about 4.5 and 3.5 fold respect to EMP construct, and S2 inducing by about 2 fold. All the constructs were TPA inducible (~3 fold) but surprisingly none of them responded to TGF-β1 treatment.
Discussion

The widespread action of the TGF-\(\beta\) family of polypeptides, affecting cell proliferation and differentiation through either stimulatory or inhibitory growth signals, suggest an important role for these factors. This makes the unravelling of the mechanism of their action a major goal in all areas of the research involving cell-cycle control studies, differentiation, cell-cell interaction, cell migration and hence tissue remodelling and inflammation.

The identification of the TIE as cis-acting element involved either in cell-cycle control, by mediating the effect of TGF-\(\beta1\) on \(c-myc\) expression (Pietenpol \textit{et al.}, 1990 a,b) and in the control of cell migration, by interfering with the synthesis of extracellular proteases (Kerr \textit{et al.}, 1990), led me to ask whether the presence of the TIE in the uPA promoter, and in particular in the S2 region, might play a role in the regulation of the expression of uPA gene.

Previous report (Grimaldi \textit{et al.}, 1986) indicated that uPA expression is modulated by factors that influence cell growth in normal murine cells and in particular that uPA mRNA expression increases rapidly following stimulation of quiescent, G0, cells by mitogenic agents. This effect is an early cellular response to growth stimulation (1-3 hours) and is due in part to increased transcriptional activity of the gene.

The observation by the group of Moses (Pietenpol \textit{et al.}, 1990 a,b) that the action of TGF-\(\beta1\) on \(c-myc\) expression is mediated by the product of the retinoblastoma gene (pRB), led me to search for cell lines other than HeLa, where the presence of the papilloma virus oncproteins has a blocking effect on RB growth-suppressive activity.

I then chose two cell lines, one expressing high constitutive level of uPA (PC-3) and one expressing a low but inducible basal level of uPA (HepG2).
The northern blot analysis, however, did not revealed any effect of TGF-β1 on uPA mRNA expression in PC-3, although a certain antiproliferative effect was described (Reyes-Moreno et al., 1995). The missing effect was specific for uPA since TGF-β1 was able to increase the level of PAI-1, as described previously (Fig. 6.2; Lund et al., 1987; Gerwin et al., 1990).

Previous report indicated that TGF-β1 had no effect on proliferation of HepG2 cells (Chapekar et al., 1989); however, it is able to counteract the TPA induction of uPA mRNA, as shown by the Northern blot analysis of Fig. 6.3. Transfection experiments in HepG2 cells with the full length promoter indicated that the -2212 bp 5' flanking region contains all the information necessary to mediate TGF-β1 action on TPA inducibility (Fig. 6.7, 2212uPACAT), but sequences other than TIE, and located in areas different from S2, seem to be involved in this regulation (Fig. 6.7, see ΔS1 and ΔS3 versus ΔS2). The opposite experiment, where S1, S2 and S3 were cloned in front of the minimal promoter and in the presence of the enhancer, seemed to indicate that multiple elements cooperate to TGF-β1 down-regulation of uPA induction by TPA (Fig. 6.7, ES1, ES2, ES3) as none of them, when present alone, responded to TGF-β1 treatment.

This would suggest that sequences lying in regions other than S2 (e.g. S1 and S3) mediate the TGF-β1 effect on TPA induction of uPA gene, and in particular that a cross-talk between these elements occurs. It would be interesting to perform a detailed analysis by nested deletions of these regions in order to find out the sequences responsible for the TGF-β1 inhibitory effect of uPA expression.

It would be also interesting to study the antiproliferative effect of TGF-β1 in other cell lines, in order to ask whether the same mechanisms that regulate c-myc expression can be involved in uPA downregulation in the cell-cycle.
Appendix 2: Negative regulation of uPA by p53

Introduction

The product of the tumour suppressor gene p53 has been implicated in normal cell proliferation and neoplastic transformation (reviewed in Ko and Prives, 1996). Both mutation or absence of expression of p53 lead to a transformed phenotypes in a number of cell systems and pathologies. Rearrangements and deletions of the p53 gene have been described in the human leukaemia cell line HL-60 (Wolf et al., 1985) and in the osteogenic sarcoma cell line SAOS-2 (Masuda et al., 1987) as well as in Friend virus induced murine leukaemia (Chow et al., 1989, Mowat et al., 1985), and in human tumors like colon (Baker et al., 1989) and lung (Takahashi et al., 1989) carcinomas. Introduction of the normal p53 gene into p53-deficient tumour cells suppresses transformation and tumorigenesis (Levine et al., 1991) and inhibits proliferation of many types of cells causing them to arrest at G1/S (Johnson et al., 1993, Livingstone et al., 1992).

p53 plays multiple roles in cells: while dispensable for viability, induced expression of wt p53, for example by DNA-damaging agents, leads to cell cycle arrest and apoptosis, protecting the genome from accumulating excess mutation.

p53 contains a strong transcriptional activation domain and is capable of inducing the expression of genes bearing a p53 binding site consisting of two copies of the 10 bp sequence 5'PuPuPuC(A/T)(T/A)GPyPyPy3' separated by up to 13 bp (Funk et al., 1992, Kern et al., 1991), but also strongly inhibits transcription from many genes lacking p53 binding sites (c-fos, c-jun, IL-6, bcl-2) through a direct interaction with specific (Agoff et al., 1993) or general transcription factors (TBP, TAFs, TFIIH; Mack et al., 1993, Lu and Levine, 1995). Several genes containing p53 binding sites have been already identified. Among these the muscle creatin kinase
Appendix 2: Negative regulation of uPA by p53

(Weintraub et al., 1991), GADD45 (Kastan et al., 1992), MDM2 (Wu et al., 1993), p21/WAF-1 (El-Deiry et al., 1993), cyclin G (Okamoto and Beach, 1994) and Bax (Miyashita and Reed, 1995).

Cellular transformation often results in a dramatic increase in the production of plasminogen activators and in particular of uPA. Previous reports indicated that the expression of uPA is modulated by factors that influence cell growth in normal murine cells, and in particular that uPA mRNA expression increases rapidly following stimulation of quiescent, G0, cells by the competent mitogenic agents (Grimaldi et al., 1986). uPA mRNA increase is an early cellular response to growth stimulation with a peak after 1-3 hours and is due, at least in part, to increased transcriptional activity of the gene, as shown by run-on analysis and does not require new protein synthesis as demonstrated by treatment with cycloheximide. The elevation of uPA mRNA is transient and is restricted to the G0/G1 phase.

I wanted therefore to investigate whether the transient elevation of uPA is regulated by the tumour suppressor gene p53, leading to a linkage between inactivation of p53, increased cell proliferation and synthesis of uPA.

Materials and Methods

Cell culture

T98G and GM47 cells are grown in Earle's minimal essential medium containing 10% foetal calf serum (FCS; GIBCO) at 37 °C. The cell lines were kindly provided by Dr. M. Fiscella (NIH, Bethesda).

Cell cycle studies

Cells are arrested in G0/G1 phase of the cell cycle by allowing them to grow to confluence followed by incubation for 3 days in medium containing 0.5% FCS.
The cells are stimulated to reenter the cell cycle by trypsinizing and replating at a 1:4 split ratio in medium containing 10% FCS with and without the addition of dexamethasone (final concentration 1 µM).

RNA isolation and Northern blot analysis

Total RNA was extracted from cells by guanidine isothiocyanate method (Chomczynski and Sacchi, 1987). RNA (10 µg per lane) was denatured with 6.3% formaldehyde/50% formamide and then size fractionated on a 1.2% agarose gel containing 6.6% formaldehyde. The RNA was then transferred to nylon membrane (Costar) by electrotransfer in the following buffer (1x TAE: 10 mM Tris-HCl, 5 mM sodium acetate, 0.5 mM EDTA, pH 7.8) for 4 hours at 40 volts and cross-linked to filter with a Stratagene cross-linker. Radiolabelled probes were obtained by in vitro transcription of SmaI linearized pT7 RNA uPA, containing the BamHI-Bgl II fragment of human uPA cDNA, and of Hind III linearized pT7 RNA 18S (Ambion) using T7 polymerase (Promega). The p53 cDNA probe was excised by pCMVp53 wt (pC53-SN3) containing the BamHI fragment of human p53 (described in Kern et al., 1992) and labelled by random priming (Rediprime, Amersham). Filters were then pre-hybridized in mixture containing 50% formamide, 6x SSPE, 5x Denhardt's, 0.5% SDS, 100 µg/ml DNA carrier (MB DNA, Boehringer Mannheim) at 42 °C for 2 hours. Hybridization with riboprobe was performed in the presence of 10% dextran-sulphate at 42 °C for 12-16 hours. The probe in excess was removed by washing the filters in 3x SSPE/0.1% SDS for 20' at 65 °C followed by 2 washing in 0.3x SSPE/0.1% SDS at 65 °C for 30' each.

The filters were then exposed with intensifying screens at -80 °C for period from 30' (18S) up to 4 days (uPA and p53 probes). mRNA level was quantified by densitometric analysis (Molecular Dynamics, SF, ImageQuant Software).
**Results**

The human T98G cell line derived from a glioblastoma multiform tumour has been a good tool for studies involving the control of gene expression during the cell cycle. In particular these cells can become arrested in G0/G1 phase under stationary phase conditions (serum starvation) and can be induced to re-enter the cell cycle by trypsinization and dilution. A clone derived by the parental T98G, containing the wild-type p53 gene under the control of the MMTV dexamethasone-inducible promoter (GM47) is used to investigate whether uPA is regulated in the cell-cycle and in particular if the product of p53 gene plays a role in the regulation of uPA expression. For this purpose both the parental T98G and the GM47 derived clone have been grown in the presence or absence of the hormone dexamethasone and the expression of uPA mRNA has been evaluated by Northern blot analysis. As shown in Fig. 7.1 uPA is induced upon re-entry in the cell cycle with a peak after 6 hours, from replating in serum-containing medium, of exponential growth (panel A), both in the parental and the stable transfected cell lines. In the presence of dexamethasone, which induces the p53 expression at 6 hours after addition (panel B), the expression of uPA mRNA in the parental T98G cell line is maximum at 6 hours after serum and dexamethasone addition, followed by a fast decrease already at 8 hours upon hormone addition. In the GM47 derived clone however the maximum level of uPA is reached after 4 hours of growth, correspondingly to overexpression of wt p53, followed by a reduction in the uPA mRNA level similar to that of the parental cell line T98G. This result would suggest that uPA mRNA is regulated during the cell-cycle in these cells, with p53 overexpression influencing the kinetic of uPA expression restricting it to an earlier stage of the cell cycle, after which the negative effect of dexamethasone treatment plays a major influence on uPA gene expression.
Appendix 2: Negative regulation of uPA by p53

Fig. 7.1: Northern blot analysis of uPA mRNA expression

Total RNA was isolated from parental T98G and GM47 cells at the times indicated. Panel A: untreated cells; Panel B: cells treated with 1 μM dexamethasone for the times indicated. Zero time are cells trypsinized after three days of serum starvation. The arrows indicated the 2.4 Kb transcript of human uPA, the 2 Kb transcript of 18 S (panel A and B) and the 2 Kb transcript (wt) or 2.8 Kb endogenous (end) p53 mRNA in GM47 cells (panel B). Panel C: graphic representation of the densitometric analysis of the northern blot-experiment.
Discussion

uPA is synthesised by a variety of tissues, including cells of the central nervous system (Masos and Miskin, 1996), where uPA is thought to play a pivotal role in the tissue-remodelling process among which neurite outgrowth (Krystosek and Seeds, 1981).

The most common and malignant brain tumour, glioblastoma multiforme, is characterised by necrosis, vascular proliferation, and invasion in the surrounding normal brain tissues (Russel and Rubinstein, 1989). Increased uPA activity has been already observed in vitro in an astrocytoma cell line and associated with the invasive properties of glioblastoma in vivo. By comparing human astrocytoma, glioblastomas and normal brain tissues it has been evidenced that the higher uPA activity detected in malignant tumour cells correlates with increased uPA mRNA level (6 and 20 fold higher than in normal brain). Both uPA protein and mRNA are localised within astrocytomas and endothelial cells and are localised near vascular proliferation zones and at the leading edges of tumours in glioblastomas. The increased uPA expression suggest that uPA may contribute to the invasion of malignant astrocytomas into adjacent normal brain tissues and to tumour angiogenesis (Yamamoto et al., 1994).

Previous data (Grimaldi et al., 1986) have shown that increased uPA mRNA level in response to mitogens is an early event due in large part to increased transcriptional activity of the gene, restricted to the G0/G1 phase of the cell-cycle. The analysis of uPA expression in a glioblastoma cell line (T98G), bearing a mutated p53 protein, indicated that uPA is induced upon reentry of the cell in the cell-cycle, with a peak after 6-8 hours. A clone overexpressing a wild-type p53 protein under the control of the MMTV inducible promoter (GM47), showed that uPA expression is reduced and restricted to 1-4 hours after induction (Fig. 7.1,
appendix 2). The increased uPA mRNA follows the induction of p53 expression (already evident at 2 hours upon reentry in the cell-cycle; Fig. 7.1), suggesting a role of p53 in the regulation of uPA.

p53 expression is induced by treatment with dexamethasone (Mercer et al., 1990). Although the negative effect of glucocorticoids on uPA gene expression has been extensively described (Medcalf et al., 1988, Henderson and Kefferd, 1993), it is possible to exclude that the effect seen in the clone GM47 is due to dexamethasone as the parental cell line (also treated with dex) behaves as the control, in which only the solvent for dexamethasone has been added.

Recently regulation of uPA by p53 has been described and the element responsible for this effect has been mapped in the region of the enhancer. No binding sites for p53 are however present in this region, although their presence is not required for the p53 repressing activity (Kunz et al., 1995). Two binding sites for p53 have been however identified in the region of the huPA promoter upstream of the enhancer (position -5045/-5036 and -5021/-5012), recently cloned in our laboratory (Ibanez-Tallon, 1997).

It would be very interesting to understand the link between p53 expression and uPA regulation. The T98G cell line has proved useful for studies on the regulation of the cell cycle and in particular on the correlation between mutations in p53 tumour suppressor gene, loss of control of cessation of proliferation, induction of uPA expression and high invasive potential of this kind of tumour. However the induction of wild type p53 protein expression by dexamethasone, could be a limit to unravel the exact mechanisms of action of p53 on uPA downregulation, as we cannot exclude an overlapping effect due to the hormone.

Another useful system to study the p53 effect on uPA gene transcription could be a BALB/c-Val5 cell line (El-Deiry et al., 1993), a BALB/c 1 derived cell line stably expressing a p53 mutant, Val 135, carrying a temperature-sensitive
mutation at Val 135, which behaves as a mutant at 37°C and as the wild type at 32°C (Ginsberg et al., 1991). The Balb/c 1 cell line does not contain endogenous p53 (Harvey and Levine, 1991). Once an effect by p53 on uPA gene expression is identified, a transient transfection analysis, using the uPA promoter-CAT constructs, would be useful in order to identify the element mediating this effect and/or if the repression occurs through the inhibition of the assembly of a competent pre-initiation complex at the start site of transcription.
8. Conclusions and perspectives

Due to its high destructive potential, uPA expression requires tight regulation. The fact that in the organism uPA is synthesised by very few cell types and only under particular stimuli gives a major importance to the negative regulation of uPA with respect to activation. The discovery of cell lines that transcribe uPA in a constitutive or inducible manner and others in which the transcription of the gene is silenced, indicates that the control of the expression requires that one or more elements cooperate in the modulation or total inhibition of the gene, for example driving the formation of a compact chromatin structure. The studies presented in this thesis have showed that a complex array of regulatory sequences, in addition to the enhancer and minimal promoter, appear to regulate uPA transcription, some of which are cell-type specific (Table 5.1).

S1 and S3 could act as modulators of uPA gene expression, in response to different stimuli, in those cell line that express uPA. In the cell lines not expressing uPA S2 plays the major role in the silencing of transcription and probably in cooperation with S1 and S3 directs the formation a compacted chromatin structure that leads to the complete silencing of the gene.

A detailed characterisation of region S1 and S3, by the mean of deletion and mutation analysis, should clarify the role of these regions in the regulation of uPA gene and in the cooperation with S2.

Generation of clones of cell lines stable transfected with full length promoter-CAT plasmid DNA would be a good tool to study the involvement of chromatin organisation on the activity of S2 together with chromatin studies of this region of the endogenous uPA gene promoter.
In the future the purification of the factors bound to S2 and cloning of their cDNA would allow a clarification of the model according to different transcriptional status of uPA gene.
9. References


that contain HI."


References

163

of the urokinase-type plasminogen activator

gene expression." J. Biol. Chem. \textbf{265}: 13327-
13334.


inhibits transcription indirectly via a

methyl-CpG binding protein.” \textit{Cell} \textbf{64}: 1123-
1134.


modifications and the chromosome cell cycle.”

\textit{BioEssays} \textbf{14}: 9-16.


method for the quantitation of microgram

quantities of protein utilizing the principle of


“A yeast silencer contains sequences that can

promote autonomous plasmid replication and

transcriptional activation.” \textit{Cell} \textbf{70}: 709-719.


“Organization and expression of eukaryotic

dna split genes coding for proteins.” \textit{Annu. Rev.

Biochem} \textbf{50}: 349-383.


control of the yeast HO gene: cis- and trans­

Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi,
R., Edmondson, D.G., Roth, S.Y., Allis, C.D.

acetyltransferase A: a homolog to yeast Gcn5p

linking histone acetylation to gene

activation.” \textit{Cell} \textbf{54}: 843-851.

Buratowski, S., Hahn, S., Guarente, L. and

complexes in transcription initiation by RNA

polymerase II.” \textit{Cell} \textbf{56}: 549-561.


“Drosophila TFIID binds to a conserved

downstream basal promoter element that is

present in many TATA-box-deficient

promoter.” \textit{Genes Dev} \textbf{10}: 711-724.

downstream core promoter element, DPE, is

conserved from Drosophila to humans and is

recognized by TAFII60 of Drosophila.” \textit{Genes

& Dev} \textbf{11}(22): 3020-3031.


transcription factor IID (TFIID).” \textit{Annu. Rev.

Biochem} \textbf{65}: 769-799.

Burt, D.W., Nakamura, N., Kelley, P., and Dzau,
V.J. (1989). “Identification of negative and

positive regulatory elements in the human renin

Busso, N., Masur, S.K., Lazega, D., Waxman, S.,

migration by pro-urokinase binding to its

receptor: possible mechanism for signal

transduction in human epithelial cells.” \textit{J Cell

Biol} \textbf{126}: 259-270.

Cains, B.R., Kim, Y.J., Sayre, M.H., Laurent,

multisubunit complex containing the

SWI1/ADR6, SWI2/SNF2, SWI3, SNF5, and


Molecular and Cellular Biology of the Yeast

Saccharomyces cerevisiae. J. R. Broach, Pringle,
J.R., Jones, E.W. New York, Cold

Spring Harbor Laboratory Press. 1: 41-146.


“A cell-type specific and enhancer-dependent

silencer in the regulation of the expression of

the human urokinase plasminogen activator

Cao, S.X., Gutman, P.D., Dave, H.P.G., and

transcriptional silencer in the 5'-flanking


Natl. Acad. Sci. USA} \textbf{86}: 5306-5309.


“The initiator directs the assembly of a

transcription factor IID-dependent


USA} \textbf{88}: 8052-8056.


skeletal muscle: a review.” \textit{Am. J. Anat.} \textbf{132}:
119.


of the plasminogen/plasmin system in

transgenic mice.” \textit{Fibrinolysis} \textbf{5}: 269-76.


targeting and gene transfer studies of the

plasminogen/plasmin system: implication in

thrombosis, hemostasis, neointima formation

and atherosclerosis.” \textit{FASEB J} \textbf{9}: 934-938.


“Signalling, mitogenesis and the
cytoskeleton: where the action is.” \textit{BioEssays}
\textbf{17}: 171-175.

Cassady, A.I., Stacey, K.J., Nimmo, K.M.,
Murphy, K.M., von der Ahe, D., Pearson, D.,
Botteri, F.M., Nagamine, Y., Hume, D.A.

urokinase plasminogen activator gene in

murine RAW264 macrophages involves distal
and 5' non-coding sequences that are conserved between mouse and pig." *Nucl. Acids Res.* 19: 6839-6847.


protein gene is repressed in HeLa cells."


isolated mammalian nuclei." Nucleic Acid Research 11:1475-1489.


References [167]

"Isolation and characterization of the human tissue-type plasminogen activator structural gene including its 5' flanking region." J. Biol. Chem. 260: 11223-11230.


the Drosophila homeotic gene Sex combs reduced." Genetics 139: 797-814.


References


Inostroza, J.A., Mermelstein, F.H., Ilho, H., Lane, W.S. and Reinberg, D. (1992). "Dr1, a TATA-
binding protein-associated phosphoprotein and inhibitor of class II gene transcription.” Cell 70: 477-489.


an activator of pro-urokinase." J. Biol. Chem. 264: 14095-14099.


References 175


References


References


Nir, U., Walker, M.D., Rutter, W.J. (1986). "Regulation of rat insulin 1 gene expression: evidence for negative regulation in non-


References 179


References


References 182


Stelzer, G., Goppelt, A., Lottspeich, F. and Meisterernst, M. (1994). "Repression of basal transcription by HMG2 is counteracted by TFIH-associated factors in an ATP-


References


